GRADUATE SCHOOL
OF DECISION SCIENCES

Universität
Konstanz

GSDS
Working Paper
No. 2017-14

A Latent Factor Model for Forecasting Realized Volatilities

Giorgio Calzolari
Roxana Halbleib
Aygul Zagidullina

July 2017

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.

GSDS – Graduate School of Decision Sciences
University of Konstanz
Box 146
78457 Konstanz

Phone: +49 (0)7531 88 3633
Fax: +49 (0)7531 88 5193
E-mail: gsds.office@uni-konstanz.de

-gsds.uni-konstanz.de

ISSN: 2365-4120
July 2017

A LATENT FACTOR MODEL FOR FORECASTING REALIZED VOLATILITIES

Giorgio CALZOLARI[1] and Roxana HALBLEIB[2] and Aygul ZAGIDULLINA[3]

July 21, 2017

Abstract

This paper proposes a new method of forecasting realized volatilities by exploiting their common dynamics within a latent factor model. The main idea is to use an additive component structure to describe the long-persistence in their autocorrelation function, where the components, extracted from high-dimensional vectors of realized volatilities, follow stationary autoregressive processes of order 1. The model we propose allows also for autoregressive structures in the idiosyncratic noises and conditional hetersokedasticity. Differently from HAR and ARFIMA, our factor model profits from the high-dimensionality of the system that provides more information of the commonality of their dynamics with direct efficiency gains in the estimates and forecasts. For estimation purposes, we use the indirect inference method that is easy to implement and provides accurate estimates. We apply the new models to vectors of up to 30 daily realized volatility series of stocks composing the Dow Jones Industrial Average index and show that they outperform standard long-memory models both in-sample and out-of-sample.

Keywords: Long Memory, Component Model, Dynamic Factor Model, Factor-GARCH Model, Indirect Inference

[1]Dipartimento di Statistica, Informatica, Applicazioni "G. Parenti", University of Firenze, Italy; email: calzolar@disia.unifi.it.

[2]Department of Economics, University of Konstanz, Germany; email: roxana.halbleib@uni-konstanz.de. Corresponding address: Roxana Halbleib, University of Konstanz, P.O. Box 124, Universitaetstrasse 10, 78464, Konstanz, Germany; phone: 0049 (0) 7531 88 5373; fax: 0049 (0) 7531 88 4450.

[3]Department of Economics, University of Konstanz, Germany; email: Aygul.Zagidullina@uni-konstanz.de.

1 Introduction

Estimating and modeling volatilities for forecasting purposes plays a central role in many financial applications. For a very long time, the volatility of financial assets was treated to be constant in time. However, financial returns exhibit clustering effects and there is significant autocorrelation in their squared series. These are clear indicators that financial volatilities are time varying and it was the seminal paper of Engle (1982) that opened a new era in researching and developing new methods to capture such empirical features by means of Generalized Autoregressive Conditional Heteroskedastic (GARCH) models. While these approaches use daily data to estimate and forecast daily volatilities, the realized measures introduced by Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold and Labys (2001) and Barndorff-Nielsen and Shephard (2002*b*) use intraday information to provide consistent estimates for daily variances[1] (known in the literature as realized variances - RVar's). However, in practice, these measures are noisy proxies of the latent variance as they are affected by the market microstructure noise (MMN) present in the observed discrete intraday prices. Most of the research on RVar has been focusing on providing accurate estimates in the presence of MMN: see Zhang et al. (2005), Barndorff-Nielsen et al. (2008), among others. The series of daily RVar's exhibit long-persistence in their autocorrelation function (ACF) and their log-transformation (log-RVar) is approximately normally distributed (Andersen, Bollerslev, Diebold and Ebens (2001) and Andersen, Bollerslev, Diebold and Labys (2001)). Risk forecasts based on RVar are mainly obtained from Autoregressive Fractional Integrated Moving Average (ARFIMA) model (Andersen, Bollerslev, Diebold and Labys (2001), among others) or from Heterogenous Autoregressive (HAR) model of Corsi (2009), both capturing their long memory.

This paper introduces an alternative approach to forecast daily RVar series by means of component aggregation techniques that aim at capturing the long dependence in their dynamics. It uses a dynamic latent factor structure, where the factors (or components) are autoregressive (AR) processes of order 1 and are extracted from large vectors of RVar series. This approach is parsimonious and flexible as it avoids the curse of dimensionality of vector autoregressive models and it can strait-forwardly allow for complex structures in order to capture further empirical features, such as conditional heteroskedasticity. The flexibility of the model is backed by the estimation method we implement, namely the indirect inference (IndInf) of Gouriéroux et al. (1993), Smith

[1] Volatility is the square root of the variance.

(1993) and Gallant and Tauchen (1996), which uses auxiliary specifications that are easy to estimate.

The new approach is motivated by the deficiencies of ARFIMA model when applied to series of daily RVar's. First, ARFIMA involves some artificial mixture of short- and long memory characteristics difficult to disentangle (Comte and Renault (1998), Harvey (2013)) and, therefore, it is treated more like a mathematical trick and less like a model with economic meaning (Corsi (2009)). Second, the ARFIMA representation involves an infinite-order lag polynomial that in practice needs to be truncated, affecting, thus, the quality of estimates and forecasts (Lanne (2006)). Third, the measurement error in the RVar estimators due to MMN requires serious attention to the MA structure, which is often neglected in ARFIMA (Hansen and Lunde (2014)). This negligence may be due to difficulties in estimating in one step all parameters of the model and to the poor estimation results when applying a heuristic method involving two-steps (Corsi (2009)). They lead to parameter estimates that are very sensitive with respect to the model specification and estimation window. This is particularly the case of the the fractional integration parameter (d) estimates, which are unstable around the threshold 0.5 that defines the frontier between stationarity and non-stationarity of the underlying variance process. Thus, Luciani and Veredas (2015) estimate d to be larger (smaller) than 0.5 on samples including (excluding) the previous financial crisis; Koopman et al. (2005) estimate d to be very close to 0.5 and the AR parameter close to 1, while Lieberman and Phillips (2008) find that estimating ARFIMA on larger samples of daily RVar's provides estimates of d smaller than 0.5. This creates some confusion in the literature in what regards the stationarity of the underlying variance process. However, Wright (1999) shows that daily variances are stationary and Hansen and Lunde (2014) reject the non-stationarity hypothesis for RVar's by using an adequate test that accounts for measurement error. Our component approach supports these findings as the estimates of AR(1) parameters describing the dynamics of factors and idiosyncratic noises are always smaller than 1. Nevertheless, despite its drawbacks, ARFIMA remains a popular choice due to its parsimony in capturing the long-memory dynamics. This advantage holds only in univariate settings and vanishes when applied to large vectors: e.g., applied on a panel of 30 series, ARFIMA(1,d,1) has around 1800 parameters.

The HAR model of Corsi (2009) is an attractive alternative to capture the long-memory of RVar series as it is easy to estimate and its AR(20) representation provides a satisfactory approximation of their long-persistence. According to Corsi (2009), HAR model can also be interpreted as a component model as it builds on an additive structure of volatilities computed over three different time horizons. However, similar to ARFIMA, the model suffers from the curse of dimensionality when applied to large vectors of RVar's (Patton and Sheppard (2015)): e.g., applied on a vector of 30 RVar series, HAR model has 2700 parameters.

The factor model we propose profits from what the alternatives suffer: the high dimensionality of the system. Thus, when the dimension of the RVar panel increases, the multivariate ARFIMA and HAR models suffer from the curse of dimensionality and efficiency loss, while the factor approach gains in efficiency. This is because the mean squared error (MSE) of vector autoregressive processes depends positively on the dimension of the dependent variable (Lütkepohl (2005)), while the MSE of a process following a dynamic factor structure is inverse proportional to it (Bai (2003), Sentana (2004)). This might be explained by the fact that increasing the dimension, one provides more information on the commonality of the underlying series, which further leads to more precise factor estimates with direct positive effects on the MSE of the system and the precision of the estimates. Our empirical results confirm these findings.

The idea of capturing long persistence by the component aggregation techniques goes back to Granger (1980). He shows that, aggregating a large number of short memory AR(1) processes with a careful choice of the parameters, one can capture the slowly-decaying pattern of ACF's. More precisely, aggregating only two components of different persistence seems to be enough to approximately capture such a pattern. This idea is exploited in the volatility literature by Ding and Granger (1996), Engle and Lee (1999) and Meddahi and Renault (2004) in the GARCH context and by Gallant et al. (1999) and Chernov et al. (2003) in the stochastic volatility framework. While two components seem to be better than one (Engle and Rosenberg (2000), Alizadeh et al. (2002), Bollerslev and Zhou (2006), Adrian and Rosenberg (2008)), in order to capture the entire persistence in the volatility dynamics, Ding and Granger (1996) suggest that the number of components should increase beyond 2.

The research on capturing the dynamics and the persistence of RVar series by means of components has so far focused mainly on continuous univariate time models, as in Barndorff-Nielsen and Shephard (2002*a*) (followed by Andersen et al. (2004), Meddahi (2003) and Bollerslev and Zhou (2006), just to name a few) that aggregate unobserved independent Lévi driven Ornstein-

Uhlenbeck processes within a state-space approach. Koopman et al. (2005) applies this model to forecast daily log-RVar's and find that two factors are sufficient to capture the persistence of the underlying series and performs similarly to ARFIMA in terms of forecasting. Barndorff-Nielsen and Shephard (2002*a*) and Koopman et al. (2005) use Quasi-Maximum Likelihood (QML) with the Kalman Filter (KF) technique to estimate the parameters of their univariate models (Harvey (1989)). However, this estimation technique becomes computationally intensive when the dimension of the dependent variable and the complexity in the model structure increases, which is our case (Koopman and Durbin (2000), Kapetanios and Marcellino (2009), Jungbacker and Koopman (2015), Dungey et al. (2000)). For this reason, we implement IndInf with a multi-step auxiliary estimation procedure (Aielli et al. (2013)) that is easy to implement: in the first step, we extract static factors and in the second step, we apply univariate dynamic models on each of the static factors in accordance with the dynamic structure of the model of interest. We validate the accuracy and the efficiency of IndInf estimators by means of Monte Carlo experiments.

The high-dimensionality of the dependent variable in our model relates it to the literature on forecasting panels of daily RVar series. The closest to our approach is the model of Luciani and Veredas (2015), which is an approximate dynamic factor model, using the non-stationary approach of Bai and Ng (2002) to extract the factors that follow a vector ARFIMA process. The approach of Luciani and Veredas (2015) is different from ours in two main ways: (1) it focuses on reducing the dimension problem of large panels of volatilities, while our focus is on using the commonality of their dynamics to capture their long-persistence and provide accurate forecasts and (2) it faces the estimation and specification problems of ARFIMA models mentioned above, while our model avoids them. Alternatively to Luciani and Veredas (2015), Atak and Kapetanios (2013) allow for the factors to follow HAR representations. A more different approach is the one of Barigozzi et al. (2014) that applies a semiparametric model to capture the common trend of panels of RVar's and the class of multiplicative error models to describe the idiosyncratic part. All these approaches use the principal component analysis (thoroughly analyzed in Ghysels (2014)) to extract the factors, which is different from our state-space representation.

To accommodate the empirical characteristics of daily RVar's, such as long-persistence and conditional heteroscedasticity, we implement three different types of dynamic factor model (DFM) specifications: (1) the standard DFM, where the factors and the idiosyncratic noises exhibit no dynamics; (2) an extended DFM, where the idiosyncratic noises have AR representations, in order to account for further serial correlation besides the one captured by the factors and (3) an extended DFM, where the factors and the idiosyncratic noises have generalized ARCH (GARCH) representations, in order to account for the conditional heteroscedasticity in daily RVar series as found by Corsi et al. (2008)).

We apply the three models to daily RVar series of 30 stocks composing the Dow Jones Industrial Average index and compare them against ARFIMA and HAR models. The empirical results show that our models outperform ARFIMA- and HAR-type representations in in-sample, but most importantly, out-of-sample when forecasting one-step and multi-step ahead. In particular, we find that, while increasing the number of factors and the structure of the model improve the goodness of fit of DFM's in-sample, it does not have the same effect out-of-sample: using DFM with white noise errors and 2 factors seems to be the best choice among all DFM's and competitive models for both 1-step and multi-step ahead forecasts, which is in line with the findings of Engle and Rosenberg (2000), Alizadeh et al. (2002), Bollerslev and Zhou (2006), Adrian and Rosenberg (2008) and Koopman et al. (2005).

The rest of the paper is organized as follows: Section 2 introduces the models of interest. Section 3 briefly describes the IndInf estimation method and its practical implementation. Section 4 presents simulation results and Section 5 provides empirical results from applying the models to real data. Section 6 concludes.

2 Dynamic Factor Models

Let

$$\mathbf{Y}_t = \mathbf{B}\mathbf{f}_t + \mathbf{u}_t, \quad (1)$$

$$\mathbf{f}_t = \mathbf{\Phi}\mathbf{f}_{t-1} + \mathbf{v}_t, \quad (2)$$

where $\mathbf{Y}_t$ is a vector of dimension $n \times 1$ composed of demeaned series of daily log-RVar's of n stocks.[2] The reasons we choose to model the log-transformation of RVar series are to avoid

[2] One can instead apply the model to daily series of log-RVar's and add a vector of intercepts in order to assure a zero mean of the vector of idiosyncratic noises. However, this would increase the computational burden of the model

parameter constraints to assure the positivity of the forecasts and to accommodate the empirical fact that log-RVar's are closer in distribution to the normal than RVar's (Andersen, Bollerslev, Diebold and Ebens (2001) and Andersen, Bollerslev, Diebold and Labys (2001)).

$\mathbf{f}_t$ is a $k \times 1$ vector of unobserved common factors, with $k \leq n$ and $\mathbf{u}_t$ is the vector of idiosyncratic noises of dimension $n \times 1$. $\mathbf{B}$ is the $n \times k$ matrix of factor loadings of rank$(\mathbf{B}) = k$. We assume that $\mathbf{f}_t$ and $\mathbf{u}_t$ as well as the factors and the idiosyncratic noises among themselves are orthogonal. Thus, $\mathbf{\Phi} = diag(\phi_1, \phi_2, \ldots, \phi_k)$ is a diagonal matrix of dimension $k \times k$ and each of the factors composing the vector $\mathbf{f}_t$ follows a univariate AR(1) process. The stationarity of the AR(1) processes is assured when $|\phi_j| < 1$ for $j = 1, \ldots, k$. Furthermore, given the empirical facts mentioned above, we assume that $\mathbf{u}_t$ and $\mathbf{v}_t$ are jointly normally distributed:

$$\begin{pmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{pmatrix} \sim N(\mathbf{0}, \mathbf{\Omega}), \tag{3}$$

where

$$\mathbf{\Omega} = \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0}_{n\times k} \\ \mathbf{0}_{k\times n} & \mathbf{\Delta} \end{pmatrix} \tag{4}$$

is a diagonal matrix of dimension $(n+k) \times (n+k)$ with $\mathbf{\Sigma} = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$ and $\mathbf{\Delta} = diag(\delta_1^2, \delta_2^2, \ldots, \delta_k^2)$ being diagonal matrices of dimension $n \times n$ and, respectively, $k \times k$, whose diagonal elements are the variances of the idiosyncratic noises and of the noises specific to the AR(1) processes describing the latent factors, respectively. The model defined in (1)-(2) is a standard dynamic factor model with idiosyncratic white noises. We denote it by DFM-WN.

To eliminate the scale indeterminacy of the common factors in DFM-WN we impose that the unconditional variances of the factors are equal to 1, i.e., $\delta_j^2 = 1 - \phi_j^2$ for all $j = 1, \ldots, k$. For further identification reasons, when the number of factors is larger or equal to two, we also impose the zero upper-triangular parametrization of $\mathbf{B}$: $b_{12} = \ldots = b_{1k} = b_{23} = \ldots = b_{k-1k} = 0$, for all $k \geq 2$ (Geweke and Zhou (1996) and Aguilar and West (2000)).

The DFM-WN approach exploits the co-movements in the dynamics of the daily log-RVar series with the aim of capturing the long persistence within a parsimonious framework. The idea is based on Granger (1980), who shows that long-memory dynamics can be approximated by the

without significant gains in the estimation results.

aggregation of short memory processes. Engle and Lee (1999), Harvey (2013), Harvey and Lange (2015), Engle and Rosenberg (2000), Alizadeh et al. (2002), Bollerslev and Zhou (2006), Adrian and Rosenberg (2008), among others, find that aggregating two components (one with close to non-stationarity and one with less persistence) is enough to mimic the slowly-decaying pattern of the ACF of a long memory process.

Figure C.1 provides some evidence in this direction. The solid (pink) line is the ACF of a simulated long memory process with $d = 0.45$, the dashed (grey) line is the ACF of an AR(1) process $X_{1,t}$ with the AR parameter equal to 0.8 and the dotted (blue) line is the ACF of a process obtained from simply aggregating two AR processes ($X_t = X_{1,t} + X_{2,t}$), with autoregressive parameters 0.8 for $X_{1,t}$ and 0.995 for $X_{2,t}$. As expected, the single AR(1) process has an ACF that decays much faster (exponentially) than the one of the long memory. However the ACF of the sum of the two AR(1)'s decays slower and it approaches the one of the long memory process.

For our purposes, we do not restrict the number of components to two. In a matter of fact, as described in the empirical application, we choose the "optimal" number of factors by means of Bayesian Information Criterion (BIC).

Although aiming at capturing the same effects, our DFM-WN specification differs from the one of Granger (1980) in two respects: (1) our model is a multivariate one, as it builds on a vector of series and (2) each element of $\mathbf{Y}_t$ builds on an aggregation of AR(1) components, but it also includes an idiosyncratic white noise component. In order to show how our DFM-WN specification given in equations (1)-(2) is related to the one of Granger (1980), we write equations (1)-(2) for each Y_i with $i = 1, \ldots, n$:

$$Y_{i,t} = \sum_{j=1}^{k} b_{i,j} f_{j,t} + u_{i,t}, \qquad i = 1, \ldots, n \tag{5}$$

$$f_{j,t} = \phi_j f_{j,t-1} + v_{j,t}, \qquad j = 1, \ldots, j, \tag{6}$$

where $V[f_{j,t}] = 1$ and $V[u_{i,t}] = \sigma_i^2$.

For comparison reasons, we also consider the approach of Granger (1980) adapted to our parameter specification:

$$X_{i,t} = \sum_{j=1}^{k} b_{i,j} f_{j,t}, \qquad i = 1, \ldots, n \tag{7}$$

$$f_{j,t} = \phi_j f_{j,t-1} + v_{j,t}, \qquad k = 1, \ldots, k. \tag{8}$$

Given that the following derivations hold for all $i = 1, \ldots, n$, for the ease of exposure, we drop for the moment the index i. Moreover, we assume in what follows that $f_{j,t}$ is stationary for all $j = 1, \ldots, k$. The following two propositions aim at outlining the main differences between the dynamic specifications of Y_t and X_t with the focus on their autocorrelation behavior.

Proposition 2.1 *Let $|\phi_j| < 1$ for all $j = 1, \ldots, k$. Then, the process Y_t defined in equations (5)-(6) is an ARMA(k,k) process of the form:*

$$(1 - \theta_1 L - \ldots - \theta_k L^k)\ Y_t = \sum_{j=1}^{k} b_j (1 - \theta_{1j} L - \ldots - \theta_{k-1,j} L^{k-1})\ v_{j,t} + (1 - \theta_1 L - \ldots - \theta_k L^k)\ u_t, \tag{9}$$

where $\theta_1, \ldots, \theta_k$ and $\theta_{1j}, \ldots, \theta_{k-1,j}$ are functions of $\phi_1, \ldots, \phi_k$ (for more detail on their specification see Appendix A).

The p-th autocorrelation of Y_t is given by:

$$\rho_p^{Y,WN} = \frac{\sum_{j=1}^{k} b_j^2 \phi_j^p}{\sum_{j=1}^{k} b_j^2 + \sigma^2}, \tag{10}$$

where $p = \pm 1, \pm 2, \pm 3, \ldots$.

For the proof see Appendix A. The ARMA(k,k) process describing Y_t is not a standard one (with a single white noise), but includes $k+1$ white noises: $v_{1,t}, \ldots, v_{k,t}$ and u_t. Moreover, the roots of the MA(k) part $(1 - \theta_1 L - \ldots - \theta_k L^k)\ u_t$ cancel with the ones of the AR(k) part $(1 - \theta_1 L - \ldots - \theta_k L^k)\ Y_t$.

Proposition 2.2 *Let $|\phi_j| < 1$ for all $j = 1, \ldots, k$. Then, the process X_t defined in equations*

(7)-(8) is an ARMA(k,k-1) process of the form:

$$(1 - \theta_1 L - \ldots - \theta_k L^k)\ X_t = \sum_{j=1}^{k} b_j (1 - \theta_{1j} L - \ldots - \theta_{k-1,j} L^{k-1})\ v_{j,t}, \tag{11}$$

where $\theta_1, \ldots, \theta_k$ and $\theta_{1j}, \ldots, \theta_{k-1,j}$ are functions of $\phi_1, \ldots, \phi_k$ (for more detail on their specification see Appendix A).

The p-th autocorrelation of X_t is given by:

$$\rho_p^X = \frac{\sum_{j=1}^{k} b_j^2 \phi_j^p}{\sum_{j=1}^{k} b_j^2}, \tag{12}$$

where $p = \pm 1, \pm 2, \pm 3, \ldots$.

The proof is given in Appendix A. Similar to the ARMA (k,k) describing Y_t, the ARMA(k,k-1) describing X_t is also not a standard one as it has k white noises, $v_{1,t}, \ldots, v_{k,t}$. However, different from ARMA(k,k), there is no immediate cancellation in the roots of AR and MA components, unless the AR parameters take specific values. Thus, we expect that Y_t exhibits less autocorrelation than X_t. Comparing the autocorrelations of the two processes, one may observe that because $\sigma^2 > 0$ and assuming that $\phi_j > 0$ for all j (which is empirically the case), $\rho_p^{Y,WN} < \rho_p^X$ for all p's, which means that the ACF generated by DFM-WN is smaller than the one of generated by the approach of Granger (1980). To show this, we plot in Figure C.2 the ACF of X_t and Y_t for $k = 2$ with $\phi_1 = 0.8$, $\phi_2 = 0.95$, $b_1 = b_2 = 1$ and $\sigma^2 = 1$. The ACF of X_t is always larger than the one of Y_t, however the difference between them becomes less pronounced by increasing the lags.

In order to show how DFM-WN captures the long-persistence of real data, we plot in Figure C.3 the ACF of log-RVar for IBM over the window 01.01.2001-19.12.2016 as well as the ACF's resulting from DFM-WN with $k = 1, 2, 3, 4$. As one may observe, increasing the number of factors increases the autocorrelation that approaches the one of the real data. Nevertheless, there seems to be autocorrelation left in the residuals that could be captured by adding further factors, which, however, comes at high parametrization and computational costs. Given the results above, one could avoid such costs by transforming the DFM-WN process into one in the spirit of Granger (1980): i.e., Y_t being modeled solely by sums of AR(1) processes (without an white noise component). This can be reached by allowing the idiosyncratic noises to follow AR(1) processes by themselves and define the new DFM to be given by equations (1)-(2) and the following one:

$$\mathbf{u}_t = \mathbf{\Lambda}\mathbf{u}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{13}$$

where $\mathbf{\Lambda} = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix of dimension $n \times n$ and $\boldsymbol{\varepsilon}_t$ and $\mathbf{v}_t$ are independent and jointly normally distributed

$$\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \mathbf{v}_t \end{pmatrix} \sim N(\mathbf{0}, \mathbf{\Xi}), \tag{14}$$

where

$$\mathbf{\Xi} = \begin{pmatrix} \mathbf{\Gamma} & \mathbf{0}_{n\times k} \\ \mathbf{0}_{k\times n} & \mathbf{\Delta} \end{pmatrix}, \tag{15}$$

and $\mathbf{\Gamma}$ is a diagonal matrix of dimension $n \times n$ such that $\mathbf{\Gamma} = diag(\gamma_1^2, \ldots, \gamma_n^2)$ with γ_i^2 being the variance of $\varepsilon_{i,t}$, $i = 1, \ldots, n$. The stationarity of $\mathbf{u}_t$ is assured when $|\lambda_1|, \ldots, |\lambda_n| < 1$. We impose the same scale indeterminacy and identification constraints as in DFM-WN. We denote the model defined in equations (1), (2) and (13) to be DFM-AR.

Given Proposition 2.2, DFM-AR with a total of k AR(1) processes ($k - 1$ latent factors and one noise) is an ARMA(k,k-1) process with the ACF:

$$\rho_p^{Y,AR} = \frac{\sum_{j=1}^{k-1} b_j^2 \phi_j^p + \sigma^2 \lambda^p}{\sum_{j=1}^{k-1} b_j^2 + \sigma^2}, \tag{16}$$

which is larger than the ACF of a DFM-WN with k latent factors for $\phi_j > 0, j = 1, \ldots, k$. Figure C.3 provides evidence in this direction. Besides the ACF of DFM-WN with $k = 1, 2, 3, 4$, it also depicts the ACF of DFM-AR with $k = 3$ (red dotted-dashed line). One may observe that, allowing for AR(1) idiosyncratic noises seems to provide a very good fit to real data in terms of ACF and its autocorrelation is larger than the one of DFM-WN with $k = 4$ (both models aggregate over the same number of AR(1) processes). Our empirical results presented in Section 5 support these findings as DFM-AR is preferred to DFM-WN in terms of goodness of fit criteria.

Once the parameters of DFM-WN and DFM-AR models are estimated, one can extract the factors by implementing the standard KF and an augmented KF as described by Jungbacker et al.

(2011). In both cases, the filtered factors depend on their variances as well as on the ones of the idiosyncratic noises. Corsi et al. (2008) show that daily log-RVar's exhibit clustering effects and conditionally heteroscedasticity, which is confirmed by our empirical results presented in Section 5. As a consequence, one should account for these effects in order to increase the accuracy of the filtered factors from panels of log-RVar's. For this reason, we extend the DFM-WN framework to allow for the idiosyncratic noises and the factors to follow GARCH processes. Thus, we assume that:

$$\begin{pmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{pmatrix} |\mathcal{F}_{t-1} \sim N(\mathbf{0}, \mathbf{\Omega}_t), \tag{17}$$

where $\mathcal{F}_{t-1}$ is the information up to time $t-1$, $\mathbf{\Omega}_t$ has the same diagonal representation as in Equation (4), where $\sigma^2_{i,t}$ for $i = 1, \ldots, n$ and $\delta^2_{j,t}$ for $j = 1, \ldots, k$ follow univariate GARCH(1,1) representations:

$$\sigma^2_{i,t} = w_i + a_i u^2_{i,t-1} + c_i \sigma^2_{i,t-1} \tag{18}$$

$$\delta^2_{j,t} = \omega_j + \alpha_j v_{j,t-1} + \beta_j \delta^2_{j,t-1}, \tag{19}$$

where $w_i, \omega_j, a_i, \alpha_j > 0$, $c_i, \beta_j \geq 0$. Thus, $\mathcal{F}_{t-1}$ includes all the observations on $\mathbf{Y}_t$, $\mathbf{u}_t$ and $\mathbf{v}_t$ up to $t-1$. The stationarity of the processes defined in equations (18)-(19) is assured by $a_i + c_i < 1$ and $\alpha_j + \beta_j < 1$. We denote the model defined in equations (1)-(2) and (17)-(19) by DFM-GARCH.

Similarly to DFM-WN and DFM-AR, we impose that the unconditional variances of the factors are equal to 1, i.e., $\omega_j = (1 - \phi_j^2)(1 - \alpha_j - \beta_j)$ for all $j = 1, \ldots, k$ (see Sentana et al. (2008), among others) and the zero upper-triangular parametrization of $\mathbf{B}$. Given estimates of the parameters, one can filter the factors by using the approximated KF approach described in Audrino et al. (2013) with the correction proposed by Harvey et al. (1992).

Thus, the vector of parameters to estimate is given by:

1. DFM-WN: $\boldsymbol{\theta} = (\boldsymbol{\phi}', \mathbf{b}', \boldsymbol{\sigma}')'$ with $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_k)'$, $\mathbf{b} = \text{vec}(\mathbf{B}') = (\mathbf{b_1}', \ldots, \mathbf{b_n}')'$, $\mathbf{b_i} = (b_{i1}, \ldots, b_{ik})'$ with $i = 1, \ldots, n$, and $\boldsymbol{\sigma} = (\sigma_1^2, \ldots, \sigma_n^2)'$. The total number of parameters to estimate is equal to $p = nk - \frac{k(k-1)}{2} + n + k$.

2. DFM-AR: $\boldsymbol{\theta} = (\boldsymbol{\phi}', \boldsymbol{\lambda}', \mathbf{b}', \boldsymbol{\gamma}')'$ with $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)'$ and $\boldsymbol{\gamma} = (\gamma_1^2, \ldots, \gamma_n^2)'$ while $\boldsymbol{\phi}$

and $\mathbf{b}$ are defined at point 1 above. The total number of parameters to estimate is equal to $p = nk - \frac{k(k-1)}{2} + 2n + k$.

3. DFM-GARCH: $\boldsymbol{\theta} = (\boldsymbol{\phi}', \mathbf{b}', \boldsymbol{\alpha}', \boldsymbol{\beta}', \mathbf{a}', \mathbf{c}')'$ with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)'$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)'$, $\mathbf{a} = (a_1, \ldots, a_n)'$, $\mathbf{c} = (c_1, \ldots, c_n)'$ and $\boldsymbol{\phi}$ and $\mathbf{b}$ are defined at point 1. The total number of parameters to estimate is equal to $p = nk - \frac{k(k-1)}{2} + 3n + 3k$

While DFM-WN could be theoretically estimated by QML with a standard KF, in practice, when applied to $\mathbf{Y}_t$ of large dimensions it suffers from computational burden (Jungbacker and Koopman (2015), Dungey et al. (2000)). The QML estimation of DFM-AR requires an augmentation of KF as described in Jungbacker et al. (2011) and Banbura and Modugno (2014). Besides an adequate augmentation of KF, the QML estimation of DFM-GARCH requires also an approximation of KF due to the fact that $\mathcal{F}_{t-1}$ is not available in complete form as the lagged factors and idiosyncratic noises are unobserved (see Diebold and Nerlove (1989), Harvey et al. (1992)). This induces inconsistencies in the estimation that are solved by Sentana et al. (2008) and Dungey et al. (2000) by implementing IndInf. Although theoretically appealing, the approach of Sentana et al. (2008) that uses the model of Harvey et al. (1992) as an auxiliary specification suffers from computational burden due to the relative complex structure of the scores of the auxiliary model and ignores the AR effects in the factors. Alternatively, Dungey et al. (2000) that include AR factors in their model, use two types of auxiliary models: the one of Diebold and Nerlove (1989) and a dual VAR model for levels and squares. While the implementation of the first auxiliary model suffers from the same computational limitations as the one of Sentana et al. (2008), the second one implies a large number of parameters to estimate. Alternatively, in this paper, we choose to implement a multi-step auxiliary estimation procedure as described in the following section, which can be easily applied to any dimension and structure of the model.

3 Estimation

Let $\mathbf{y}_t$ be the vector (dimension $n \times 1$) of realizations of $\mathbf{Y}_t$ at time t, where $t = 1, \ldots, T$, characterized by the probability density function (pdf) $f_0(\mathbf{y}_t, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of unknown parameters of dimension $p \times 1$ describing the true model. Let $\boldsymbol{\theta}_0$ be the true value of $\boldsymbol{\theta}$. $f_0(\mathbf{y}_t, \boldsymbol{\theta})$ is intractable or infeasible, such that the parameter vector $\boldsymbol{\theta}_0$ cannot be estimated by ML. In order to apply IndInf to alleviate this problem, one important condition is that one can easily simulate pseudo-random numbers from the model of interest.

The auxiliary model is characterized by $f^*(\mathbf{y}_t, \boldsymbol{\beta})$, where $\boldsymbol{\beta}$ is a vector of unknown parameters and let β_0 be the pseudo-true value of $\boldsymbol{\beta}$. The dimension of the vector $\boldsymbol{\beta}$ is $q \times 1$ and, in order to assure identification of the parameter vector $\boldsymbol{\theta}$, one has to impose that $q \geq p$. Different from $f_0(\mathbf{y}_t, \boldsymbol{\theta})$, $f^*(\mathbf{y}_t, \boldsymbol{\beta})$ is feasible and tractable, so that QML can be implemented to estimate the parameter vector β_0. The corresponding log-likelihood function for the observations $\mathbf{y}_t$ with $t = 1, \ldots, T$ is given by: $\mathcal{L}^*(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T; \boldsymbol{\beta})$.

The IndInf estimation procedure consists of the following steps:

Step A: Derive the QML estimator of $\boldsymbol{\beta}_0$ such that:

$$\hat{\boldsymbol{\beta}} = \arg\max_{\boldsymbol{\beta}} \mathcal{L}^*(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T; \boldsymbol{\beta}). \tag{20}$$

Step B: Choose a value for $\boldsymbol{\theta}$ and simulate H paths of length T from the model of interest: $\mathbf{y}_{h,1}(\boldsymbol{\theta}), \ldots, \mathbf{y}_{h,T}(\boldsymbol{\theta})$, with $h = 1, \ldots, H$.

Step C: Compute the IndInf estimator $\hat{\boldsymbol{\theta}}$ such that (Gallant and Tauchen (1996)):

$$\hat{\boldsymbol{\theta}}(\Xi) = \arg\min_{\boldsymbol{\theta}} \frac{\partial \mathcal{L}^*_{H,T}}{\partial \boldsymbol{\beta}'}(\boldsymbol{\theta}, \hat{\boldsymbol{\beta}}) \Psi \frac{\partial \mathcal{L}^*_{H,T}}{\partial \boldsymbol{\beta}}(\boldsymbol{\theta}, \hat{\boldsymbol{\beta}}), \tag{21}$$

where $\mathcal{L}^*_{H,T}(\boldsymbol{\theta}, \hat{\boldsymbol{\beta}}) \equiv \frac{1}{H}\sum_{h=1}^{H} \frac{1}{T}\mathcal{L}^*(\mathbf{y}_{h,1}(\boldsymbol{\theta}), \mathbf{y}_{h,2}(\boldsymbol{\theta}), \ldots, \mathbf{y}_{h,T}(\boldsymbol{\theta}); \hat{\boldsymbol{\beta}})$. Ψ is a weighting matrix that is symmetric nonnegative definite.

The IndInf estimator is consistent and asymptotically normal distributed for fixed H (see Gouriéroux et al. (1993)). Moreover, when $p = q$, the results are independent of the choice of Ψ that can be fixed to the identity matrix. The variance-covariance matrix of the IndInf estimator is given by:

$$\mathbf{W} = \Big(1 + \frac{1}{H}\Big)\Big[\frac{\partial^2 \mathcal{L}^*_{H,\infty}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\beta}'}(\boldsymbol{\theta_0}, \boldsymbol{\beta_0})\, I(\boldsymbol{\beta}_0)^{-1} \frac{\partial^2 \mathcal{L}^*_{H,\infty}}{\partial \boldsymbol{\beta} \partial \boldsymbol{\theta}'}(\boldsymbol{\theta_0}, \boldsymbol{\beta_0})\Big]^{-1}, \tag{22}$$

where $\mathcal{L}^*_{H,\infty}(\boldsymbol{\theta}, \boldsymbol{\beta}) = \lim_{T\to\infty} \mathcal{L}^*_{H,T}(\boldsymbol{\theta}, \boldsymbol{\beta})$ and $I(\boldsymbol{\beta}_0)$ is the Fisher information matrix of the auxiliary model.

In practice, one should make a choice on H. From Equation (22) one can observe that H plays an important role in the efficiency of the estimators and that large values should be preferred. However, increasing H increases also the computational burden as it implies more simulations as described in Step B above. Therefore, the common approach is to choose H to be between 1 and 10 (Monfardini (1998), Calzolari et al. (2014), among others). Calzolari and Halbleib (2016) show in their empirical application that the results do not significantly improve if H increases from 10 to 100. Therefore, both in our simulation and real data exercise presented below, we choose $H = 10$.

For each of the three DFM specifications, we choose an adequate auxiliary model that resembles the one of interest as defined in Section 2. Thus, to estimate DFM-WN we consider the following auxiliary representation:

$$\mathbf{Y}_t = \mathbf{B}^*\mathbf{f}_t^* + \mathbf{u}_t^*, \tag{23}$$

$$\mathbf{f}_t^* = \mathbf{\Phi}^*\mathbf{f}_{t-1}^* + \mathbf{v}_t^*, \tag{24}$$

where $\mathbf{B}^*$ is the matrix of loadings of dimension $n \times k$, $\mathbf{f}_t^*$ is a vector of latent factors of dimension $k \times 1$, where each component is independent of each other, $\mathbf{f}_t^*$ and $\mathbf{u}_t^*$ are orthogonal, $\mathbf{u}_t^* \sim N(\mathbf{0}, \mathbf{\Sigma}^*)$ where $\mathbf{\Sigma}^* = diag(\sigma_1^{*2}, \ldots, \sigma_n^{*2})$ and $\mathbf{v}_t^* \sim N(\mathbf{0}, \mathbf{\Delta}^*)$, where $\mathbf{\Delta}^* = diag(\delta_1^{*2}, \ldots, \delta_k^{*2})$, $\mathbf{\Phi}^* = diag(\phi_1^*, \ldots, \phi_k^*)$. The univariate AR(1) processes describing the dynamics of the components of $\mathbf{f}_t^*$ are stationary when $|\phi_j^*| < 1$ for $j = 1, \ldots, k$.

To estimate DFM-AR, the auxiliary specification includes besides equations (23) and (24) also:

$$\mathbf{u}_t^* = \mathbf{\Lambda}^*\mathbf{u}_{t-1}^* + \boldsymbol{\varepsilon}_t^*, \tag{25}$$

where $\boldsymbol{\varepsilon}_t^* \sim N(0, \mathbf{\Gamma}^*)$, with $\mathbf{\Gamma}^* = diag(\gamma_1^{*2}, \ldots, \gamma_n^{*2})$, $\mathbf{\Lambda}^* = diag(\lambda_1^*, \ldots, \lambda_n^*)$ and $\boldsymbol{\varepsilon}_t^*$ is independent of $\mathbf{v}_t^*$ and $\mathbf{v}_t^* \sim N(\mathbf{0}, \mathbf{\Delta}^*)$, where $\mathbf{\Delta}^*$ is defined above. The AR(1) processes of the idiosyncratic noises are stationary when $|\lambda_i^*| < 1$ for $i = 1, \ldots, n$.

In order to estimate DFM-GARCH, we use an auxiliary specification where the factors and the noises follow univariate GARCH(1,1) processes. Thus, we assume that $\mathbf{u}_t^*|\mathcal{F}_{t-1}^* \sim N(\mathbf{0}, \mathbf{\Sigma}_t^*)$ and $\mathbf{v}_t^*|\mathcal{F}_{t-1}^* \sim N(\mathbf{0}, \mathbf{\Delta}_t^*)$, where $\mathcal{F}_{t-1}^*$ includes all the information on $\mathbf{u}_t^*$, $\mathbf{f}_t^*$ and $\mathbf{Y}_t$ up to time $t-1$, $\mathbf{\Sigma}_t^* = diag(\sigma_{1,t}^{*2}, \ldots, \sigma_{n,t}^{*2})$ and $\mathbf{\Delta}_t^* = diag(\delta_{1,t}^{*2}, \ldots, \delta_{k,t}^{*2})$ and $\sigma_{i,t}^{*2}$ and $\delta_{j,t}^{*2}$ have the following representations for $i = 1, \dots, n$ and $j = 1, \dots, k$, respectively:

$$\sigma_{i,t}^{*2} = w_i^* + a_i^* u_{i,t-1}^{*2} + c_i^* \sigma_{i,t-1}^{*2} \quad (26)$$

$$\delta_{j,t}^{*2} = \omega_j^* + \alpha_j^* v_{j,t-1}^* + \beta_j^* \delta_{j,t-1}^{*2}, \quad (27)$$

where $w_i^*, \omega_j^*, a_i^*, \alpha_j^* > 0$, $c_i^*, \beta_j^* \geq 0$. The processes defined in equations (26) and (27) are stationary when $a_i^* + c_i^* < 1$ and $\alpha_j^* + \beta_j^* < 1$, respectively. As a results, based on equations (23), (24), (26) and (27), $\mathbf{Y}_t$ conditional on $\mathcal{F}_{t-1}^*$ is also normally distributed with mean zero and time-varying variance-covariance matrix.

To eliminate the scale indeterminacy of the common factors, we impose that the unconditional variance of each component of $\mathbf{f}_t^*$ is equal to one, i.e. $\delta_j^{*2} = 1 - \phi_j^{*2}$ and $\omega_j^* = (1 - \phi_j^{*2})(1 - \alpha_j^* - \beta_j^*)$. Moreover, similar to the models presented in Section 2, when $k \geq 2$, we impose an upper-triangular parametrization of $\mathbf{B}^*$. The auxiliary models have the same number of parameters as the ones in Section 2.

As already mentioned above, while the auxiliary models for DFM-WN and DFM-AR could be estimated in one step by means of QML, the one of the DFM-GARCH model is infeasible as $\mathcal{F}_{t-1}^*$ is not available. To solve this problem, we implement a sequential estimation procedure that consists of several steps[3]. The sequential procedure is applied to all three DFM specifications and consists of:

Step 1: Use QML on the static factor model given in Equation (23) to estimate $\mathbf{B}^*$ and $\mathbf{\Sigma}^*$.

Step 2: Extract k "approximated" common factors $\hat{\mathbf{f}}_t^* = (\hat{f}_{1t}^*, \dots, \hat{f}_{kt}^*)$ and corresponding residuals $\hat{\mathbf{u}}_t^* = (\hat{u}_{1t}^*, \dots, \hat{u}_{nt}^*)$ based on Lawley and Maxwell (1962).

Step 3: Use QML to estimate the rest of the parameter vector by imposing the adequate dynamics on each of the factors and idiosyncratic noises extracted in Step 2 above. In particular:

Step 3.1: To estimate DFM-WN, use QML on univariate AR(1) processes for each of the components of $\hat{\mathbf{f}}_t^*$ in order to get estimates of $\mathbf{\Phi}^*$.

Step 3.2: To estimate DFM-AR, additionally to 3.1., use QML on univariate AR(1) processes for each of the components of $\hat{\mathbf{u}}_t^*$ in order to get estimates of $\mathbf{\Lambda}^*$ The unconditional

[3] Some early experiments on this approach are undertaken by Aielli et al. (2013) and Calzolari and Halbleib (2016)

variances of $\hat{\mathbf{u}}_t^*$ can be strait-forwardly derived from $\mathbf{\Sigma}^*$ and $\mathbf{\Lambda}^*$.

Step 3.3: To estimate DFM-GARCH, additionally to 3.1., use QML on univariate GARCH processes for each component of $\hat{\mathbf{u}}_t^*$ and $\hat{\mathbf{f}}_t^*$ in order to estimate the parameters of equations (26) and (27).

This three-step estimation procedure provides a huge simplification in terms of computation as it can be applied to any dimension of $\mathbf{Y}_t$ and to any structure of the factors and idiosyncratic noises. This shows the advantages of IndInf that consistently estimates parameters of very complex models by implementing simple auxiliary specifications. Moreover, in our case, the IndInf provides an unified estimation procedure for all three models avoiding, thus, further transformations/approximations necessary in order to implement QML.

The only additional cost involved by the three-step procedure is the unavailability of the variance-covariance matrix of the IndInf estimators given by Equation (22) as the Fisher information matrix of the auxiliary model cannot be estimated. We replace it by a consistent estimator given by the sample variance-covariance matrix of 1000 independent simulated score vectors of the auxiliary model, which are computed after the last iteration, upon convergence (see Calzolari and Halbleib (2016)).

4 Monte Carlo Study

In this section we provide evidence on the performance of IndInf to accurately estimate the parameters of the models introduced in this paper as well as on its performance compared to the standard ML when estimating the parameters of DFM-WN. Moreover, we provide evidence on the advantages of using large versus small dimensions of $\mathbf{Y}_t$ on the properties of the estimates.

To keep the computation simple, we set the maximum number of factors to 3. Because the results between the three choices of $k = 1, 2, 3$ do not significantly differ, we focus here on presenting simulation results only for the 2-factor case, while the results for $k = 1$ and $k = 3$ can be obtained from the authors upon request. We focus on simulating data that have properties as close as possible to the real data used in the empirical application presented in the following section. Thus, we choose the total number of observations to be given by $T = 3773$, the total number of assets to be equal to $n = 30$ and the parameter values to be close to the estimates obtained from the real data (see Section 5). In all simulation experiments, we set the total number of replications to be equal to $R = 1000$.

To avoid the curse of dimensionality, we implement some restricted versions of DFM-AR and DFM-GARCH as it follow: (1) we impose that all 30 idiosyncratic noises are driven by the same AR parameter, i.e. impose that $\lambda_1 = \ldots = \lambda_n$ and (2) we impose that only the first factor has a (G)ARCH representation, while the others and the idiosyncratic noises are (conditional) homoskedastic. Although this might seem a restrictive assumption, the conditional heteroskedasticity of $\mathbf{Y}_t$ is still captured by $V(\mathbf{Y}_t|\mathcal{F}_{t-1}) = \mathbf{B}V(\mathbf{f}_t|\mathcal{F}_{t-1})\mathbf{B}' + V(\mathbf{u}_t|\mathcal{F}_{t-1}) = \mathbf{B}\mathbf{\Phi}\mathbf{\Delta}_t\mathbf{\Phi}'\mathbf{B}' + \mathbf{\Sigma}$, where $\mathbf{\Delta}_t$ is diagonal and $\mathbf{\Delta}_t = diag(\delta_{1t}^2, \delta_2^2, \ldots, \delta_k^2)$ where δ_{1t}^2 is defined in Equation (19). Within the DFM-GARCH framework, we implement both DFM-ARCH(1) and DFM-GARCH(1,1) models. Given that the simulation results for both models are very similar, we focus here on presenting the ones for DFM-ARCH(1), while the ones for DFM-GARCH(1,1) can be obtained from the authors upon request.

Tables B.1 and B.2 present the simulation results from estimating the DFM-WN by IndInf and ML and DFM-AR and DFM-ARCH by IndInf. The estimates appear in general unbiased (differences to the true parameter values are observable after two digits), especially in what regards the AR and ARCH parameters and the variances of the idiosyncratic noises. This is due to the very large number of observations used for the estimation of these parameters ($30 \times 3773 \approx 110000$) as well as to the relatively high ratio between the unconditional variance of the factors ($\delta_1^2 = \ldots = \delta_k^2 = 1$) and the unconditional variance of the idiosyncratic noise ($\sigma_1^2 = \ldots = \sigma_n^2 = 0.2$) (for similar results, see also Sentana et al. (2008)). While the introduction of AR effects in the idiosyncratic noises worsens in general the precision of the estimates, this is not the case of ARCH effects, which provide in general more precise estimates.

When comparing IndInf and ML, one may observe that when estimating the AR parameters and the variances of idiosyncratic noises, the two approaches perform equally well in terms of bias and efficiency. The difference between them becomes obvious when estimating $\mathbf{B}$: as expected, the ML estimates are more efficient than the IndInf ones[4].

To understand the effects of the dimensionality of $\mathbf{Y}_t$ on the properties of the estimates, we

[4]The same pattern is observed in Table B.3, which presents results for a smaller panel of RVar series.

run a simulation experiment on a set of 8 variables and compare the estimation results (presented in Table B.3) to the ones from tables B.1 and B.2: as expected, the estimates stemming from the panel of 30 series are in general more precise than the ones derived from the panel of 8 series. Thus, increasing the dimensionality $\mathbf{Y}_t$ one provides more information on the commonality of the series with positive effects on the efficiency of the estimates.

5 Empirical Application

In this section we provide empirical evidence on the performance of the models described in Section 2 when applied to real data. Because the focus of the paper is forecasting, the evidence is mainly out-of-sample. The performance of the models is compared to existing long memory ones, such as ARFIMA and HAR.

The data we consider are daily series of RVar's of 30 stocks composing the Dow Jones Industrial Average index[5] for the period starting on 01.11.2001 and ending on 19.12.2016 ($T = 3773$ observations). The daily RVar's series are computed from 5-minute returns stemming from Trade and Quotations (TAQ) database[6] by using a subsampling estimator in the spirit of Zhang et al. (2005)(see Halbleib and Voev (2016)). Table B.4 in Appendix B reports the descriptive statistics of the series and of their logarithm transformation. While the RVar series exhibit right skewness and overkurtosis, the logarithm transformation smoothes these features: the empirical standard deviation decreases, the skewness gets closer to zero and the kurtosis gets closer to the one of the standard normal distribution (Andersen, Bollerslev, Diebold and Ebens (2001) and Andersen, Bollerslev, Diebold and Labys (2001)). As already pointed in Section 2, this is why we choose to model the series of log-RVar's instead of RVar's.

Figure C.4 in Appendix C plots the line graphs the daily log-RVar series for the 30 stocks and for the period under consideration. From the graphs, one may observe that the series exhibit

[5]The stocks are: Alcoa Inc. (AA), American Express Company (AXP), Boeing Corporation (BA), Bank of America Corporation (BAC), Citigroup Inc. (C), Caterpillar Inc. (CAT), Chevron Corporation (CVX), Dupont (DD), Walt Disney Company (DIS), General Electric Company (GE), The Goldman Sachs Group, Inc. (GS), Home Depot Inc. (HD), Honeywell International Inc. (HON), Hewlett-Packard Company (HPQ), International Business Machines (IBM), International Paper Company (IP), Johnson & Johnson (JNJ), J.P. Morgan Chase & Company (JPM), Coca-Cola Company Kraft Foods Inc. (KO), McDonald's Corporation (MCD), 3M Company (MMM), Altria Group Inc. (MO), Merck & Company Inc. (MRK), Nike, Inc. (NKE), Pfizer Inc. (PFE), Procter & Gamble Company (PG), United Technologies Corporation (UTX), Verizon Communications Inc. (VZ), Wal-Mart Stores Inc. (WMT) and Exxon Mobil Corporation (XOM).

[6]We acknowledge Sebastian Bayer for preparing the high-frequency data.

common behavior, especially during turbulent financial times, such as the previous financial crises from 2007/2008. This empirical fact motivates our choice of capturing their dynamics by means of latent common factors. Moreover, all series have clustering effects, which indicate that they are conditionally heteroskedastic. Applying an ARCH-LM test, we reject the H_0 of homoskedasticity for all series at all meaningful significance levels, which is in line with the findings of Corsi et al. (2008). However, after filtering GARCH(1,1) effects from each of the series, the H_0 of the ARCH-LM tests are not rejected at 5% significance level.[7] Figure C.5 plots the ACF of the 30 daily log-RVar series up to lag 200. The graphs show the long persistence in the ACF's that decay slowly.

The models we consider for comparison are long-memory ones, such as ARFIMA-type models (as applied by Andersen, Bollerslev, Diebold and Ebens (2001)) and HAR-type models (as introduced by Corsi (2009)). Given the multivariate character of our approach, we implement vector counterparts of the univariate approaches as follows:

1. The Vector ARFIMA (VARFIMA) model we consider is of the form:

$$\Psi(L)D(L)\mathbf{Y}_t = \Pi(L)\boldsymbol{\xi}_t, \tag{28}$$

   where $\Psi(L) = I_n - \Psi_1 L - \Psi_2 L^2 - \ldots - \Psi_P L^P$, $\Pi(L) = I_n - \Pi_1 L - \Pi_2 L^2 - \ldots - \Pi_S L^S$ are matrix lag polynomials with Ψ_l, $l = 1, \ldots, P$ and Π_s, $s = 1, \ldots, S$ being coefficient matrices of dimension $n \times n$ and $D(L)$ is a diagonal matrix of dimension $n \times n$, $D(L) = diag((1-L)^{d_1}, \ldots, (1-L)^{d_n})'$, where $d_1, \ldots, d_n$ are the degrees of fractional integration of each of the elements of the vector $\mathbf{Y}_t$. The process is stationary if the roots of $\Psi(L)$ and $\Pi(L)$ are outside the unit circle and $d_i < 0.5$ for $i = 1, \ldots, n$. The model given in Equation (28) is known as the VARFIMA(P,d,S). The dimension of the parameter vector describing the model is equal to $(P + S)n^2 + n$. For our purposes we consider three specifications, namely VARFIMA(0,d,0), VARFIMA(1,d,0) and VARFIMA(1,d,1). Given the large number of parameters involved (30, 930, and 1830, respectively), we choose to implement two restricted specifications, namely one where Ψ_1 and Π_1 are scalars and $d_1 = \ldots = d_n = d$ (denoted by sARFIMA) and one where Ψ_1 and Π_1 are diagonal matrices

[7] The results on ARCH-LM tests can be obtained from the authors upon request.

and the elements of $\boldsymbol{\xi}_t$ are orthogonal to each other (denoted by dARFIMA). This reduces the total number of parameters to 1, 2 and 3, respectively, in the sARFIMA framework and to 30, 60 and 90, respectively, in the dARFIMA one. The parameters of dARFIMA are obtained from estimating univariate ARFIMA models on each of the series by means of non-linear least squares as described in Beran (1995). While this estimation technique is also used for sARFIMA, it becomes infeasible when applied to full matrix VARFIMA models. To estimate such models one has to use QML on the final equation form as described in Lütkepohl (2005). However, in practice, it becomes infeasible when the dimension of the depended variable is large, as it is in our case.

2. The Vector HAR (VHAR) model is given by:

$$\mathbf{Y}_t = \Gamma^{(d)}\mathbf{Y}_{t-1} + \Gamma^{(w)}\mathbf{Y}_{t-1}^{(w)} + \Gamma^{(m)}\mathbf{Y}_{t-1}^{(m)} + \boldsymbol{\omega}_t, \tag{29}$$

where $\mathbf{Y}_{t-1}^{(w)}$ and $\mathbf{Y}_{t-1}^{(m)}$ average over the past 5 and 20 past values of $\mathbf{Y}_t$ and $\Gamma^{(d)}$, $\Gamma^{(w)}$, and $\Gamma^{(m)}$ are matrices of dimension $n \times n$. The VHAR model is heavily parameterized as it is has $3n^2 = 2700$ number of parameters. For this reason, as in the VARFIMA case, we consider two restrictive specifications: the scalar VHAR (denoted as sHAR), where the the matrices $\Gamma^{(d)}$, $\Gamma^{(w)}$ and $\Gamma^{(m)}$ are restricted to be scalars and the diagonal VHAR (denoted as dHAR) where the matrices $\Gamma^{(d)}$, $\Gamma^{(w)}$ and $\Gamma^{(m)}$ have a diagonal representation and the elements of $\boldsymbol{\omega}_t$ are orthogonal on each other. Thus, the total number of parameters is reduced to 3 and 90, respectively. The HAR models are estimated by means of ordinary least squares (OLS). Given the autoregressive design of HAR models and their simple estimation, as shown by Lütkepohl (2005), the full VHAR model can also be estimated by applying OLS to each of the 30 series composing $\mathbf{Y}_t$. Thus, differently from the VAFRIMA, here we implement besides the sHAR and dHAR, also the full matrix VHAR (denoted vHAR).

The diagonal representations are a common practice in the literature on modeling multivariate volatilities (see Chiriac and Voev (2011)) and are comparable to our DFM approaches in terms of parametrization. Although very restrictive, the scalar representations prove to be valuable choices in terms of forecasting as shown by Chiriac and Voev (2011).

For comparison reasons and to avoid the curse of dimensionality, in the present empirical exercise, we implement the three DFM approaches in the following forms: (1) DFM-WN without any parameter constraints except for the ones for identification, (2) DFM-AR by imposing that all 30 idiosyncratic noises are driven by the same AR parameter, i.e. impose that $\lambda_1 = \ldots = \lambda_n$ and (3) DFM-(G)ARCH by imposing an ARCH(1) and a GARCH(1,1) structure, respectively, only to the first factor. We impose the same structures in the corresponding auxiliary models. Moreover, we set the maximum number of factors to three and choose the optimal one by means of BIC. Tables B.5 and B.8 give the total number of parameters of the three representations.

Before focusing on the out-of-sample performance, we briefly outline some in-sample results. Tables B.5, B.6 and B.7 in Appendix B report the estimation results for the DFM-WN and DFM-AR models, while Tables B.8, B.9 and B.10 present the results for the DFM-ARCH and DFM-GARCH representations. From the tables, one may see that almost all estimates are significantly different from zero, except for some loading parameters for the 3-factor case. The factors of the DFM-WN model exhibit different persistence: the additional factor in the 2-component case has a larger persistence than the first factor that is almost equal the one of the 1-factor model. The third factor in the 3-component case has a lower persistence than the other two, which are almost equal to the 2-component case. This indicates that the factor of the 1-component model inherits the ”average” persistence, while adding new factors induces some variation in their persistence, especially within the DFM-WN and DFM-(G)ARCH representations. In DFM-(G)ARCH representations, the estimates of the (G)ARCH parameters describing the conditional heteroskedactic process of the first factor remain unaffected by the addition of homoskedastic factors.

The estimated values of the AR coefficients and of the (G)ARCH parameters indicate that the underlying series are stationary. Although not necessary (see the discussion below), we impose, in the estimation routine, stationarity constraints due to the specific implementation of IndInf. As we usually want to avoid long estimation times and deal with bad starting values, we allow for relatively large step changes in the parameter estimates that lead to jumps in the non-stationary zone during the optimization routine. While this might be no problem in classical estimation techniques (as long as this effect is short-timed), for the IndInf it might become a problem given that for any new value of the parameter estimates, new data is simulated. Simulating data with dynamics that are totally different (e.g., non-stationary) from the ones of the underlying series (e.g., stationary) troubles the estimation routine so severely that in most of the cases it breaks down before convergence. To avoid this, we impose that AR(1), (G)ARCH parameters and the sum of GARCH parameters are all strictly smaller than 1.

However, these constraints are not necessary if choosing an adequate estimation routine. We prove this by starting, in two of the cases (1- and 2-factor DFM-WN), with ”risky” starting values (at around 0.999) and imposing a very small maximum step change in the parameters during the optimization routine (of about 0.00001). After numerous iterations, the estimates converge to the values reported in tables B.5, B.6 and B.7. Thus, the reported AR(1) estimates, being significantly smaller than 1, can be replicated by dropping the stationarity constraints and using a ”slower” convergence routine.

Tables B.5 and B.8 report the BIC’s of the DFM models considered. They indicate that increasing the number of components and allowing for AR in the idiosyncratic noises or GARCH effects in the factors improves the goodness of fit. Moreover a GARCH(1,1) representation in the factors provides lower BIC’s compared to ARCH(1). Although the GARCH(1,1) structure in the factors improves the goodness-of-fit compared to the case of (conditional) homoskedasticity, the highest improvement is reached by allowing for an AR structure in the idiosyncratic noises. This is in line with the fact that the dynamics of the RVar series are mainly driven by long-persistance and less by conditional heteroskedasticity.

When compared to ARFIMA- and HAR-type models (see Table B.11), the DFM-AR model with three components outperforms all others, including the vHAR model, which is the most parameterized one and, consequently, expected to capture at best the dynamics of the underlying series. The diagonal and the scalar VARFIMA and HAR perform similarly in terms of in-sample fit, which is in line with Chiriac and Voev (2011), but are outperformed by DFM-WM with 3 factors and all DFM-AR models. Although in some cases slightly worse, the BIC’s of DFM-WN and DFM-GARCH with 2 factors are comparable to the ones of the diagonal and scalar ARFIMA and HAR models. Table B.11 also reports the BIC of DFM-WN model with 2 and 3 factors estimated by QML: compared to the IndInf counterparts, the BIC’s stemming from the QML approach are larger, indicating a worse in-sample goodness of fit.

Estimating the DFM models on smaller panels of RVar series faces various problems. Our Monte Carlo evidence provided in the previous section shows that estimating these models on a panel of 8 series leads to less efficient estimates. On real data, the estimation on panels of dimension $n = 5$, $n = 6$ and $n = 9$ faces convergence problems as the AR coefficients approach the non-stationary threshold. While DFM-WN and DFM-AR work satisfactory for panels of 7 and 8 series, we encountered difficulties to get reliable results from DFM-(G)ARCH. Tables B.12 and B.13 present estimates and standard errors of DFM-WN and DFM-AR when applied on panels including the first 7 and 8 RVar series, respectively, (namely of Alcoa Inc., American Express Company, Boeing Corporation, Bank of America Corporation, Citigroup Inc., Caterpillar Inc., Chevron Corporation and Dupont) along with the estimates and standard errors of the corresponding series from applying the same models to the panel of 30 series. The results confirm the findings of the Monte Carlo exercise: using smaller panels to extract and estimate the factors leads in general to less precise parameter estimates than using large panels.

Summing up the in-sample results, we find that (1) the DFM models perform remarkably well compared to the competitive ones, (2) increasing the number of factors and the structure of DFM model improve the in-sample goodness of fit, (3) accounting for the long persistance in the series increases the goodness-of fit more than accounting for conditional heterskedasticity, (4) the DFM-WN with 2 components, which has been established in literature to be a good fit for univariate conditional volatilities (Barndorff-Nielsen and Shephard (2002*a*), Koopman et al. (2005)), perform similarly to the restricted HAR and VARFIMA models, (5) the mostly parameterized HAR model, which is expected to perform best in-sample, is outperformed by the 3-factors DFM-AR model in terms of BIC and (6) increasing the dimension of the panel improves the precision of the estimates.

The out-of-sample evaluation of the models is done by comparing their performance when forecasting 1-, 5- and 10-step ahead during four different periods including: 1000, 750, 500 and 250 observations, respectively. For this, we divide the whole sample of 3773 observations in an in-sample window including the first 2773 observations and out-of-sample windows including the rest of 1000 observations, the last 750, 500 and 250 observations, respectively. The forecasts are done by rolling-window and the multi-step ahead forecasts are computed by cumulating the 1-step ahead ones on non-overlapping windows. Thus, the total number of 5 (10)-step ahead forecasts is about 200 (100), 150 (75), 100 (50) and 50 (25), respectively for each of the 4 out-of-sample windows. For the unobserved true (log-)variances, we use the (log-)RVar series we have at hand (Patton (2011)). The reason we consider, besides 1-step ahead forecasts also multi-step ahead ones is to exploit the ability of the models under consideration to capture the long-persistence in the dynamic of RVar series. As shown by Chiriac and Voev (2011), among others, these types of models are particulary useful to forecast the risks over longer periods compared to the short-memory approaches (such as GARCH models on daily returns or ARMA models on daily RVar series) with practical relevance for risk management according to Basel Committee (1996) and the follow-up accords.

Table B.14 presents the MSE for 1-step ahead forecasts averaged over the 30 stocks for both log-RVar and RVar series. While the models implemented forecast log-RVar's, in order to compute the forecasts of RVar, one has to take the exponential transformation. Based on the model assumptions in Section 2, the forecast errors for log-RVar are normally distributed with mean zero and variance V. Due to the log-transformation, the forecast errors for RVar are normally distributed with the mean given by $e^{\frac{V}{2}}$. Thus, the forecasts of RVar are biased and need to be corrected as described in Bianchi and Calzolari (1980) and Oomen (2001). The entries in bold correspond to the forecasts building the 95% Model Confidence Set (MCS) as described by Hansen et al. (2011).[8]

From Table B.14, one may observe that regardless of the out-of-sample window size, the DFM models provide in almost all the cases smaller MSE's than the competitive models. More precisely, when forecasting log-RVar's, while for the windows of size 1000 and 750, the 2-factor DFM-WN has the smallest MSE among all models considered, for the other two window choices, the best is the 1-factor DFM-WN. The 2-factor DFM-WN is part of the 95% MCS for all window sizes. The second best models, after DFM-WN, are the DFM-(G)ARCH. Among the competitive models, only for the evaluation window of 500 days, two out of nine models enter the 95% MCS. In general it seems that they perform equally well among themselves as documented by Chiriac and Voev (2011). When forecasting RVar's, 2-factor DFM-WN is the most reliable choice as it enters the 95% MCS for all the window sizes and it provides the smallest MSE among all models for two out of four out-of-sample windows. For two evaluation windows, the 95% MCS includes also DFM-(G)ARCH models. In three out of the four cases, the dARFIMA(0,d,0) seems to be also a good choice, as it enters the 95% MCS. When forecasting RVar on the window of length 1000, dARFIMA(1,d,1) becomes part to the 95% MCS although its MSE is much larger than the rest. These forecasts are, however, not very precise, as they have an almost double variance

[8]To derive the MCS's we implement the open-source toolbox arch.bootstrap.mcs written by Kevin Sheppard and available on http://arch.readthedocs.io/en/latest/multiple-comparison/multiple-comparison-reference.html

compared to the others.[9] As expected, due to the huge parametrization, the vHAR model is not an attractive choice for forecasting purposes: it has in general the largest MSE for all window sizes and forecasting objectives.

Tables B.15 and B.16 provide MSE results for 5- and 10-step ahead forecasts for log-RVar and RVar series, respectively. The entries of both tables reveal that 2-factor DFM-WN is a good choice for forecasting multi-step ahead variances as it always enters the 95% MCS. Allowing for AR(1) idiosyncratic noises or (G)ARCH effects does not necessarily improve the forecasting ability of DFM as they are inferior to DFM-WN. The ARFIMA and HAR models do not enter in general the 95% MCS. Allowing for AR(1) idiosyncratic noises or (G)ARCH effects does not necessarily improve the forecasting ability of DFM as they are inferior to DFM-WN. The ARFIMA and HAR models do not enter in general the 95% MCS, with the exception of dARFIMA when forecasting the variance over 5 periods ahead in one single choice of out-of-sample windows. However, similar to the 1-step ahead forecasts, these forecasts are little reliable as they are very imprecise compared to the others.

Before concluding, we present some out-of-sample forecasts from working with small panels compared to large ones. For this reason, we run a simple forecast exercise for the first 8 series by applying the 2-factor DFM-WN model. We compare the MSE and the variance of the corresponding forecasts stemming from the panel of 30 series. The out-of-sample results confirm the Monte Carlo and in-sample-findings: the forecasts stemming from the smaller panel have higher average MSE's (0.9294 and 3.6879 when forecasting log-RVar and RVar, respectively) and average variances (0.7672 and 1.4246 when forecasting log-RVar and RVar, respectively) than their counterparts stemming from the large panel of 30 series (MSE: 0.7944 and 3.0786 , respectively and variance: 0.6910 and 1.0819, respectively).

To sum up the out-of sample forecasts, we find that (1) DFM models provide the best forecasts in terms of MSE compared to the standard approaches; (2) the 2-factor DFM-WN is an attractive choice in forecasting both 1-step and multi-step ahead (log-) variances as it is included in all 95% MCS's, (3) increasing the number of factors or the structure of the model does not necessarily pays off in terms of forecasting accuracy, especially multi-step ahead, as, in general, it adds model/estimation noise to the forecasts and (4) the forecasts become more efficient if derived from

[9] Bias and variance values of the forecasts can be obtained from the authors upon request.

larger panels of RVar series.

6 Conclusions

This paper introduces a new way of forecasting large panels of daily realized variance series by using a latent dynamic factor structure. This approach is in the spirit of Granger (1980) and aims at capturing the long-memory in the underlying series by aggregating latent components/factors with short memory. We also allow for conditional heteroskedasticity in the factor and/or idiosyncratic noises as well as for autoregressive structure in the idiosyncratic noises. These approaches take advantage of the common dynamics in the series of daily realized variance series to extract the factors within a state-space model representation and alleviate the curse of dimensionality and other shortcomings of standard long-memory models, such as ARFIMA and HAR-type models, when applied to large vectors of series.

Given the latent structure and the complexity of the model due to conditional heteroskedasticity, we use for estimation purposes the indirect inference technique with a multi-step auxiliary estimation procedure. This technique is very easy to implement regardless of the number of factors or the complexity of the model and provides accurate estimates. Within a comprehensive empirical application on daily series of realized variances of 30 stocks composing Dow Jones Industrial Average index, we show that the dynamic factor models introduced here provide better in-sample and out-of-sample results than HAR and ARFIMA models. In particular, we show that increasing the number of factors and the complexity of the specification, increases the in-sample goodness of fit, while a relatively simple specification with 2 factors outperforms all models when forecasting 1-step and multi-step ahead.

Acknowledgements

For helpful comments we would like to thank Torben Andersen, Ralf Brüggemann, Fulvio Corsi, Andrew Harvey, Ostap Okhrin, Winfried Pohlmeier, Esther Ruiz, Enrique Sentana, members of the Chair of Econometrics and of the Chair of Statistics at the Department of Economics at University of Konstanz, Germany, participants at CFE 2014 in Pisa, CFE 2016 in Seville and 10th

SoFiE Conference in New York. All remaining errors are ours. Giorgio Calzolari gratefully acknowledges financial support of the project MIUR PRIN MISURA - Multivariate models for risk assessment. Roxana Halbleib gratefully acknowledges financial support from the European Social Fund, Ministry of Science, Research and Arts of Baden–Württemberg, Germany, from Zukunftskolleg University of Konstanz, from the Baden-Württemberg Stiftung, Germany through the Eliteprogramm für Postdoktorandinnen und Postdoktoranden. This work was primarily conducted using the computational resource bwUniCluster funded by the Ministry of Science, Research and Arts of Baden-Württemberg and the universities of the State of Baden-Württemberg, Germany, within the bwHPC framework program.

References

Adrian, T. and Rosenberg, J. (2008), 'Stock returns and volatility: Pricing the short-run and long-run components of market risk', *Journal of Finance* **63**, 2997–3030.

Aguilar, O. and West, M. (2000), 'Bayesian dynamic factor models and portfolio allocation', *Journal of Business and Economic Statistics* **18**, 338–357.

Aielli, G. P., Calzolari, G. and Fiorentini, G. (2013), Fast indirect estimation of latent factor models with conditional heteroskedasticity, *in* E. Brentari and M. Carpita, eds, 'Advances in Latent Variables.', Milano: Vita e Pensiero. Proceedings of the SIS2013 Statistical Conference, Brescia, 19-21 June, 2013. ISBN 978-88-343-2556-8, http://meetings.sis-statistica.org/index.php/sis2013/ALV/paper/viewFile/2579/335.

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

Andersen, T., Bollerslev, T. and Meddahi, N. (2004), 'Analytic evaluation of volatility forecasts', *International Economic Review* **45**, 1079–1110.

Andersen, T. G. and Bollerslev, T. (1998), 'Answering the skeptics: Yes, standard volatility models do provide accurate forecasts', *International Economic Review* **39**, 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001), 'The distribution of realized exchange rate volatility', *Journal of the American Statistical Association* **96**, 42–55.

Atak, A. and Kapetanios, G. (2013), 'A factor approach to realized volatility forecasting in the presence of finite jumps and cross-sectional correlation in pricing errors', *Economics Letters* **120**, 224–228.

Audrino, F., Corsi, F. and Filipova, K. (2013), 'Bond risk premia forecasting: A simple approach for extracting macroeconomic information from a panel of indicators'.
URL: *10.1080/07474938.2013.833809*

Bai, J. (2003), 'Inferential theory for factor models of large dimensions', **71**, 135–171.

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

Banbura, M. and Modugno, M. (2014), ‘Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data’, *Journal of Applied Econometrics* **29**, 133–160.

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

Barndorff-Nielsen, O. E., Hansen, P., Lunde, A. and Shephard, N. (2008), ‘Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise’, *Econometrica* **76**(6), 14811536.

Barndorff-Nielsen, O. E. and Shephard, N. (2002*a*), ‘Econometric analysis of realized volatility and its use in estimating stochastic volatility models’, *Journal of the Royal Statistical Society Series B* **64**(2), 253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2002*b*), ‘Estimating quadratic variation using realized variance’, *Journal of Applied Econometrics* **17**, 457–477.

Basel Committee (1996), Overview over the amendment to the capital accord to incorporate market risks, Technical report, Basel Committee on Banking Supervision.

Beran, J. (1995), ‘Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models’, *Journal of the Royal Statistical Society, Series B* **57**, 659–672.

Bianchi, C. and Calzolari, G. (1980), ‘The one-period forecast errors in nonlinear econometric models’, **21**, 201–208.

Bollerslev, T. and Zhou, H. (2006), ‘Volatility puzzles: A simple framework for gauging return volatility regressions’, *Journal of Econometrics* **131**, 123–150.

Calzolari, G. and Halbleib, R. (2016), ‘Estimating stable latent factor models by indirect inference’, *Journal of Econometrics* . forthcoming.

Calzolari, G., Halbleib, R. and Parrini, A. (2014), ‘Estimating garch-type models with symmetric stable innovations: Indirect inference versus maximum likelihood’, *Computational Statistics & Data Analysis* **71**, 945–973.

Chernov, M., Gallant, A. R., Ghysels, E. and Tauchen, G. (2003), ‘Alternative models for stock price dynamics’, *Journal of Econometrics* **116**, 225–257.

Chiriac, R. and Voev, V. (2011), ‘Modelling and forecasting multivariate realized volatility’, *Journal of Applied Econometrics* **26**, 922–947.

Comte, F. and Renault, E. (1998), ‘Long memory in continuous-time stochastic volatility models’, *Mathematical Finance* **8**, 291–323.

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

Corsi, F., Kretschmer, U., Mittnik, S. and Pigorsch, C. (2008), ‘Volatility of realized volatility’, *Econometric Reviews* **27**, 46–78.

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

Ding, Z. and Granger, C. W. J. (1996), ‘Modeling volatility persistence of speculative returns: A new approach’, *Journal of Econometrics* **73**, 185–215.

Dungey, M., Martin, V. and Pagan, A. (2000), ‘A multivariate latent factor decomposition of international bond yield spreads’, **15**, 697–715.

Engle, R. F. (1982), ‘Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation’, *Econometrica* **50**(4), 987–1008.

Engle, R. F. and Lee, G. J. (1999), A permanent and transitory component model of stock return volatility, *in* R. Engle and H. White, eds, ‘Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger’, Oxford: Oxford University Press, pp. 475–497.

Engle, R. and Rosenberg, J. (2000), ‘Testing the volatility term structure using option hedging criteria’, *Journal of Derivatives* **8**, 10–28.

Gallant, A. R., Hsu, C.-T. and Tauchen, G. (1999), ‘Using daily range data to calibrate volatility diffusions and extract the forward integrated variance’, *Review of Economics and Statistics* **81**, 617–631.

Gallant, A. R. and Tauchen, G. (1996), ‘Which moments to match?’, *Econometric Theory* **12**, 657–681.

Geweke, J. and Zhou, G. (1996), ‘Measuring the pricing error of the arbitrage pricing theory’, *The Review of Financial Studies* **89**, 557–587.

Ghysels, E. (2014), Factor analysis with large panels of volatility proxies. CEPR Discussion Paper No. 10034.

Gouriéroux, C., Monfort, A. and Renault, E. (1993), ‘Indirect inference’, *Journal of Applied Econometrics* **8**, 85–118.

Granger, C. W. J. (1980), ‘Long-memory relationships and the aggregation of dynamic models’, *Journal of Econometrics* **14**, 227–238.

Halbleib, R. and Voev, V. (2016), ‘Forecasting covariance matrices: A mixed approach’, *Journal of Financial Econometrics* **14**.

Hansen, P. R. and Lunde, A. (2014), ‘Estimating the persistence and the autocorrelation function of a time series that is measured with error’, *Econometric Theory* **30**, 60–93.

Hansen, P. R., Lunde, A. and Nason, J. M. (2011), ‘The Model Confidence Set’, *Econometrica* **79**(2), 453–497.

Harvey, A. (2013), *Dynamic Models for Volatility and Heavy Tails*, Econometric Society Monograph. Cambridge University Press.

Harvey, A. C. (1989), *Forecasting structural time series models and the Kalman filter*, Cambridge University Press.

Harvey, A. C., Ruiz, E. and Sentana, E. (1992), ‘Unobservable component time series models with arch disturbances’, *Journal of Econometrics* **52**, 129–158.

Harvey, A. and Lange, R.-J. (2015), Modeling the interactions between volatility and returns. CWPE Discussion Paper nr. 1518.

Jungbacker, B. and Koopman, S. (2015), ‘Likelihood-based dynamic factor analysis for measurement and forecasting’, *The Econometrics Journal* **18**, C1–C21.

Jungbacker, B., Koopman, S. and van der Wel, M. (2011), ‘Maximum likelihood estimation for dynamic factor models with missing data’, *Journal of Economic Dynamics & Control* **35**, 1358–1368.

Kapetanios, G. and Marcellino, M. (2009), ‘A parametric estimation method for dynamic factor models of large dimensions’, *Journal of Time Series Analyisis* **30**, 208–238.

Koopman, S. J. and Durbin, J. (2000), ‘A parametric estimation method for dynamic factor models of large dimensions’, *Journal of Time Series Analyisis* **21**, 281–296.

Koopman, S. J., Jungbacker, B. and Hol, E. (2005), ‘Forecasting daily variability of the s&p 100 stock index using historical, realised and implied volatility measurements’, *Journal of Empirical Finance* **12**, 445–475.

Lanne, M. (2006), ‘A mixture multiplicative error model for realized volatility’, *Journal of Financial Econometrics* **4**, 594–616.

Lawley, D. N. and Maxwell, A. E. (1962), ‘Factor analysis as a statistical method’, *Journal of the Royal Statistical Society. Series D (The Statistician)* **12**, 209–229.

Lieberman, O. and Phillips, P. C. B. (2008), ‘Refined inference on long memory in realized volatility’, *Econometric Reviews* **27**, 254–267.

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

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

Meddahi, N. (2003), ‘Arma representation of integrated and realized variances’, *The Econometrics Journal* **6**, 334–355.

Meddahi, N. and Renault, E. (2004), ‘Temporal aggregation of volatility models’, *Journal of Econometrics* **119**, 355–379.

Monfardini, C. (1998), ‘Estimating stochastic volatility models through indirect inference’, *The Econometrics Journal* **1**, 113–128.

Oomen, R. C. A. (2001), Using high frequency stock market index data to calculate, model and forecast realized return variance. European University, Economics Discussion Paper No. 2001/6.

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

Patton, A. and Sheppard, K. (2015), 'Good volatility, bad volatility: Signed jumps and the persistence of volatility', **97**, 683–697.

Sentana, E. (2004), 'Factor representing portfolios in large asset markets', **119**, 257–289.

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

Smith, A. (1993), 'Estimating non–linear time series models using simulated vector autoregressions: Two approaches', *Journal of Applied Econometrics* **8**, 63–84.

Wright, J. H. (1999), 'Testing for a unit root in the volatility of asset returns', *Journal of Applied Econometrics* **14**, 309–318.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005), 'A tale of two time scales: Determining integrated volatility with noisy high frequency data', *Journal of the American Statistical Association* **100**, 1394–1411.

Appendix A: Mathematical Derivations

A.1. Proof of Proposition 2.1

Consider the DFM-WN model given in equations (5)-(6), where we drop the index i:

$$Y_t = \sum_{j=1}^{K} b_j f_{j,t} + u_t \tag{A.30}$$

$$f_{j,t} = \phi_j f_{j,t-1} + v_{j,t}, \tag{A.31}$$

and $V[u_t] = \sigma^2$, $V[f_{j,t}] = 1$ and $|\phi_j| < 1$ for $j = 1, \ldots, k$. Then, we can writte $f_{jt} = \frac{v_{jt}}{1-\phi_j L}$. Replacing this in Equation (A.30), one gets:

$$Y_t = \sum_{j=1}^{k} \frac{b_j v_{jt}}{1-\phi_j L} + u_t = \frac{\sum_{j=1}^{k} b_j \prod_{s\neq j}^{k}(1-\phi_s L)\, v_{jt}}{\prod_{j=1}^{k}(1-\phi_j L)} + u_t. \tag{A.32}$$

Rewriting Equation (A.32) yields

$$\prod_{j=1}^{k}(1-\phi_j L)\; y_t = \sum_{j=1}^{k} b_j \prod_{s\neq j}^{k}(1-\phi_s L)\; v_{jt} + \prod_{j=1}^{k}(1-\phi_j L)\; u_t, \tag{A.33}$$

which is the same as writing

$$(1-\phi_1 L)\cdots(1-\phi_k L)\; y_t = \sum_{j=1}^{k} b_j\; (1-\phi_1 L)\cdots(1-\phi_{j-1}L)(1-\phi_{j+1}L)\cdots(1-\phi_k L)\; v_{jt} + $$
$$(1-\phi_1 L)\cdots(1-\phi_k L)\; u_t. \tag{A.34}$$

One can use the fact that:

$$\prod_{j=1}^{k}(1-\phi_j L) = 1 - \sum_{l=1}^{k}\theta_l L^l, \tag{A.35}$$

where θ_l's with $l = 1, \ldots, k$ are functions of $\phi_1, \ldots, \phi_k$. They can be obtained by multiplying (A.35) by L^{-k} and defining $z = L^{-1}$ yielding

$$(z-\phi_1)(z-\phi_2)\cdots(z-\phi_k) = z^k - \theta_1 z^{k-1} - ... - \theta_k. \tag{A.36}$$

Setting sequentially in Equation (A.36) $z = \phi_i$ for $i = 1, ..., k$ (making the left hand side of the equation equal to zero), we obtain the system

$$\phi_1^k - \theta_1\phi_1^{k-1} - ... - \theta_k = 0 \tag{A.37}$$

$$\vdots$$

$$\phi_k^k - \theta_1\phi_k^{k-1} - ... - \theta_k = 0, \tag{A.38}$$

from which we derive the θ_l's.

In the same manner one can write:

$$\prod_{s\neq j}^{k}(1-\phi_s L) = 1 - \sum_{l=1}^{k-1}\theta_{lj} L^l, \tag{A.39}$$

where θ_{lj} are functions of $\phi_1, \phi_2, \ldots, \phi_{j-1}, \phi_{j+1}, \ldots, \phi_k$ and that can be derived from a system of $k-1$ equations similar to the one given in (A.37)-(A.38). Thus, (A.32) can be written as:

$$\left[1-\sum_{l=1}^{k}\theta_l L^l\right] Y_t = \sum_{j=1}^{k} b_j \left[1-\sum_{l=1}^{k-1}\theta_{lj} L^l\right] v_{jt} + \left[1-\sum_{l=1}^{k}\theta_l L^l\right] u_t, \tag{A.40}$$

or as:

$$(1-\theta_1 L - ... - \theta_k L^k)\; Y_t = \sum_{j=1}^{k} b_j (1-\theta_{1j} L - ... - \theta_{k-1,j} L^{k-1})\; v_{jt} + \quad (1-\theta_1 L - ... - \theta_k L^k)\; u_t, \tag{A.41}$$

indicating that Y_t is an ARMA(k, k) process composed of one AR(k), k MA(k) and one MA(k+1) components.

In order to derive the autocorrelation function of Y_t, it is more appropriate to write it as an MA(∞) given that $\frac{1}{1-\phi L} = 1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \ldots$. Thus,

$$Y_t = \sum_{j=1}^{k} b_j \frac{v_{j,t}}{1-\phi_j L} + u_t \tag{A.42}$$

$$= \sum_{j=1}^{k} b_j v_{j,t}(1 + \phi_j L + \phi_j^2 L^2 + \ldots) + u_t \tag{A.43}$$

$$= \sum_{j=1}^{k} (b_j v_{j,t} + b_j \phi_j v_{j,t-1} + b_j \phi_j^2 v_{j,t-2} + \ldots) + u_t \tag{A.44}$$

$$= \sum_{m=0}^{\infty}\sum_{j=1}^{k} b_j \phi_j^m L^m v_{j,t} + u_t \tag{A.45}$$

Thus, the unconditional variance of Y_t is given by:

$$\gamma_0^Y \equiv E[Y_t^2] = \sum_{j=1}^{k} b_j^2 (1 + \phi_j^2 + \phi_j^4 + \ldots) V[v_{j,t}] + V[u_t]$$

$$= \sum_{j=1}^{k} b_j^2 \frac{1}{(1-\phi_j^2)}(1-\phi_j^2) + \sigma^2 = \sum_{j=1}^{k} b_j^2 + \sigma^2, \tag{A.46}$$

as $E[Y_t] = 0$, $V[v_{j_t}] = (1-\phi_j^2)$ and $V[u_t] = \sigma^2$.

The autocovariance computed at lag p is given by:

$$\gamma_p^Y \equiv E[Y_t Y_{t-p}] = E[(\sum_{j=1}^{k}(b_j v_{j,t} + b_j\phi_j v_{j,t-1} + \cdots + b_j \phi_j^p v_{j,t-p} + b_j \phi_j^{p+1} v_{j,t-p-1} + \ldots) + u_t)$$

$$\times (\sum_{j=1}^{k}(b_j v_{j,t-p} + b_j \phi_j v_{j,t-p-1} + \ldots) + u_{t-p})]$$

$$= \sum_{j=1}^{k} b_j^2 (\phi_j^p + \phi_j^{p+2} + \phi_j^{p+4} + \ldots) V[v_{j,t}]$$

$$= \sum_{j=1}^{k} b_j^2 \phi_j^p (1 + \phi_j^2 + \phi_j^4 + \ldots) V[v_{j,t}]$$

$$= \sum_{j=1}^{k} b_j^2 \frac{\phi_j^p}{(1-\phi_j^2)}(1-\phi_j^2) = \sum_{j=1}^{k} b_j^2 \phi_j^p \tag{A.47}$$

Thus, the autocorrelation function at lag p is given by:

$$\rho_p^Y \equiv \frac{\gamma_p^Y}{\gamma_0^Y} = \frac{\sum_{j=1}^k b_j^2 \phi_j^p}{\sum_{j=1}^k b_j^2 + \sigma^2}. \tag{A.48}$$

A.2. Proof of Proposition 2.2

X_t is given by the aggregation of only AR(1) components:

$$X_t = \sum_{j=1}^k b_j f_{j,t} \tag{A.49}$$

$$f_{j,t} = \phi_j f_{j,t-1} + v_{j,t} \tag{A.50}$$

Then,

$$X_t = \sum_{j=1}^k \frac{b_j v_{jt}}{1-\phi_j L} \tag{A.51}$$

$$= \frac{\sum_{j=1}^k b_j \prod_{s\neq j}^k (1-\phi_{sL})\ v_{jt}}{\prod_{j=1}^k (1-\phi_j L)}. \tag{A.52}$$

Thus,

$$\prod_{j=1}^k (1-\phi_j L)\ X_t = \sum_{j=1}^k b_j \prod_{s\neq j}^k (1-\phi_s L)\ v_{jt}, \tag{A.53}$$

which is equivalent to:

$$(1-\sum_{l=1}^k \theta_l L^l)\ X_t = \sum_{j=1}^k b_j (1-\sum_{l=1}^{k-1} \theta_{lj} L^l)\ v_{jt}, \tag{A.54}$$

where θ_l and θ_{lj} are defined in the previous proof. Consequently, X_t is an ARMA(k,k-1) process composed by one AR(k) and k MA(k-1) components.

In order to derive the ACF of X_t, we re-write it as an MA(∞) process as it follows:

$$X_t = \sum_{j=1}^k b_j \frac{v_{j,t}}{1-\phi_j L} = \sum_{j=1}^k b_j v_{j,t}(1+\phi_j L + \phi_j^2 L^2 + \phi_j^3 L^3 + \dots) = \sum_{m=0}^{\infty}\sum_{j=1}^k b_j \phi_j^m L^m v_{j,t} \tag{A.55}$$

Thus, the variance of X_t is given by:

$$\gamma_0^X \equiv E[X_t^2] = \sum_{j=1}^k b_j^2 V[v_{j,t}](1+\phi_j^2+\phi_j^4+\dots) = \sum_{j=1}^k b_j^2 \frac{1}{1-\phi_j^2}(1-\phi_j^2) = \sum_{j=1}^k b_j^2, \tag{A.56}$$

as $E[X_t] = 0$ and $V[v_{j,t}] = 1-\phi_j^2$.

The p-th autocovariance is given by:

$$\gamma_p^X \equiv E[X_t X_{t-p}] = E[(\sum_{j=1}^k (b_j v_{j,t} + \dots + b_j \phi_j^p v_{j,t-p} + b_j \phi_j^{p+1} v_{j,t-p-1} + \dots)]$$
$$\times [(\sum_{j=1}^k (b_j v_{j,t-p} + b_j \phi_j v_{j,t-p-1} + \dots)]$$

$$= \sum_{j=1}^{k} b_j^2 \phi_j^p (1 + \phi_j^2 + \phi_j^4 \dots) V[v_{j,t}]$$

$$= \sum_{j=1}^{k} b_j^2 \phi_j^p \frac{1}{1-\phi_j^2}(1-\phi_j^2) = \sum_{j=1}^{k} b_j^2 \phi_j^p \tag{A.57}$$

The p-th autocorrelation is given by:

$$\rho_p^X \equiv \frac{\gamma_p^X}{\gamma_0^X} = \frac{\sum_{j=1}^{k} b_j^2 \phi_j^p}{\sum_{j=1}^{k} b_j^2}. \tag{A.58}$$

Appendix B: Tables

Table B.1: **Monte Carlo results from estimating DFM-WN with 2 factors by Indirect Inference (IndInf) and Maximum Likelihood (ML) and DFM-AR and DFM-ARCH by Indirect Inference**: $n = 30$ series, $T = 3773$ observations, $R = 1000$ replications.

			DFM-WN				DFM-AR			DFM-ARCH		
			ML		IndInf		IndInf			IndInf		
		True	Mean	Std. dev	Mean	Std. dev	True	Mean	Std. dev	True	Mean	Std. dev
$\mathbf{\Phi}$	ϕ_1	0.920	0.919	0.007	0.920	0.007	0.920	0.920	0.007	0.930	0.929	0.006
	ϕ_2	0.970	0.969	0.004	0.969	0.004	0.920	0.919	0.007	0.980	0.978	0.006
$\mathbf{\Lambda}$	λ						0.500	0.499	0.003			
$\boldsymbol{\alpha}$	α									0.160	0.157	0.031
$\mathbf{\Sigma}$	σ_1^2	0.200	0.200	0.005	0.200	0.005	0.200	0.199	0.007	0.180	0.180	0.004
	σ_2^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_3^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_4^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_5^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_6^2	0.200	0.199	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_7^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_8^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_9^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{10}^2	0.200	0.199	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{11}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{12}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{13}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{14}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.199	0.006	0.180	0.180	0.005
	σ_{15}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{16}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{17}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{18}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{19}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.006	0.180	0.180	0.004
	σ_{20}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{21}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{22}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{23}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{24}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{25}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.004
	σ_{26}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{27}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{28}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{29}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.200	0.007	0.180	0.180	0.005
	σ_{30}^2	0.200	0.200	0.005	0.200	0.005	0.200	0.199	0.007	0.180	0.180	0.005

Table B.2: **Cnt'd Monte Carlo results from estimating DFM-WN with 2 factors by Indirect Inference (IndInf) and Maximum Likelihood (ML) and DFM-AR and DFM-ARCH by Indirect Inference**: $n = 30$ series, $T = 3773$ observations, $R = 1000$ replications.

			DFM-WN				DFM-AR		DFM-ARCH		
			ML		IndInf		IndInf		IndInf		
		True	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev	True	Mean	Std. dev
	$b_{1,1}$	1.000	0.996	0.040	0.999	0.043	0.999	0.044	0.740	0.738	0.035
	$b_{1,2}$	1.000	0.995	0.043	1.002	0.082	0.998	0.074	0.740	0.739	0.039
	$b_{1,3}$	1.000	0.996	0.042	0.996	0.083	1.001	0.076	0.740	0.739	0.039
	$b_{1,4}$	1.000	0.995	0.043	1.002	0.082	0.997	0.074	0.740	0.739	0.039
	$b_{1,5}$	1.000	0.996	0.042	0.996	0.083	1.001	0.076	0.740	0.738	0.039
	$b_{1,6}$	1.000	0.995	0.042	1.002	0.083	0.997	0.075	0.740	0.738	0.039
	$b_{1,7}$	1.000	0.996	0.041	0.996	0.083	1.001	0.075	0.740	0.738	0.039
	$b_{1,8}$	1.000	0.995	0.042	1.002	0.083	0.996	0.074	0.740	0.738	0.038
	$b_{1,9}$	1.000	0.996	0.042	0.996	0.083	1.001	0.075	0.740	0.738	0.039
	$b_{1,10}$	1.000	0.995	0.042	1.002	0.082	0.997	0.075	0.740	0.738	0.039
	$b_{1,11}$	1.000	0.996	0.041	0.996	0.083	1.000	0.075	0.740	0.738	0.038
	$b_{1,12}$	1.000	0.995	0.042	1.002	0.083	0.996	0.075	0.740	0.738	0.039
	$b_{1,13}$	1.000	0.996	0.041	0.996	0.083	1.001	0.075	0.740	0.738	0.039
	$b_{1,14}$	1.000	0.995	0.042	1.002	0.083	0.997	0.075	0.740	0.739	0.039
	$b_{1,15}$	1.000	0.996	0.042	0.996	0.083	1.001	0.075	0.740	0.738	0.039
	$b_{1,16}$	1.000	0.995	0.043	1.002	0.082	0.997	0.074	0.740	0.738	0.039
	$b_{1,17}$	1.000	0.996	0.042	0.996	0.083	1.001	0.075	0.740	0.738	0.039
	$b_{1,18}$	1.000	0.995	0.042	1.002	0.082	0.998	0.074	0.740	0.739	0.039
	$b_{1,19}$	1.000	0.997	0.042	0.997	0.083	1.002	0.074	0.740	0.739	0.039
	$b_{1,20}$	1.000	0.995	0.043	1.002	0.083	0.997	0.075	0.740	0.738	0.039
	$b_{1,21}$	1.000	0.996	0.041	0.996	0.083	1.002	0.075	0.740	0.739	0.039
	$b_{1,22}$	1.000	0.995	0.043	1.002	0.082	0.997	0.074	0.740	0.738	0.039
B	$b_{1,23}$	1.000	0.996	0.041	0.996	0.083	1.001	0.074	0.740	0.738	0.038
	$b_{1,24}$	1.000	0.995	0.042	1.002	0.083	0.996	0.075	0.740	0.738	0.039
	$b_{1,25}$	1.000	0.996	0.041	0.996	0.083	1.000	0.075	0.740	0.738	0.039
	$b_{1,26}$	1.000	0.995	0.042	1.002	0.082	0.997	0.073	0.740	0.738	0.039
	$b_{1,27}$	1.000	0.995	0.042	0.997	0.083	1.002	0.076	0.740	0.739	0.039
	$b_{1,28}$	1.000	0.995	0.043	1.002	0.082	0.998	0.075	0.740	0.739	0.039
	$b_{1,29}$	1.000	0.996	0.042	0.996	0.083	1.002	0.075	0.740	0.739	0.038
	$b_{1,30}$	1.000	0.995	0.042	1.002	0.083	0.998	0.075	0.740	0.739	0.039
	$b_{2,2}$	-1.000	-0.996	0.065	-0.997	0.070	-1.000	0.044	0.220	0.218	0.022
	$b_{2,3}$	1.000	0.996	0.065	0.997	0.070	0.999	0.044	-0.220	-0.218	0.022
	$b_{2,4}$	-1.000	-0.996	0.065	-0.997	0.070	-1.000	0.045	-0.220	-0.219	0.022
	$b_{2,5}$	1.000	0.996	0.065	0.996	0.070	0.998	0.044	0.220	0.218	0.021
	$b_{2,6}$	-1.000	-0.996	0.065	-0.997	0.070	-0.999	0.045	-0.220	-0.219	0.022
	$b_{2,7}$	1.000	0.995	0.065	0.997	0.070	0.999	0.043	-0.220	-0.219	0.022
	$b_{2,8}$	-1.000	-0.996	0.065	-0.997	0.070	-1.000	0.045	0.220	0.219	0.022
	$b_{2,9}$	1.000	0.996	0.065	0.997	0.070	0.999	0.044	-0.220	-0.219	0.022
	$b_{2,10}$	-1.000	-0.996	0.065	-0.997	0.070	-0.999	0.046	-0.220	-0.219	0.022
	$b_{2,11}$	1.000	0.996	0.065	0.997	0.069	0.999	0.043	0.220	0.218	0.021
	$b_{2,12}$	-1.000	-0.996	0.065	0.997	0.070	-0.999	0.045	-0.220	-0.219	0.022
	$b_{2,13}$	1.000	0.996	0.064	0.997	0.070	0.998	0.044	-0.220	-0.219	0.022
	$b_{2,14}$	-1.000	-0.996	0.065	0.997	0.070	-1.001	0.045	0.220	0.218	0.022
	$b_{2,15}$	1.000	0.995	0.065	0.997	0.070	0.999	0.044	-0.220	-0.219	0.022
	$b_{2,16}$	-1.000	-0.996	0.065	0.997	0.071	-0.999	0.045	-0.220	-0.219	0.022
	$b_{2,17}$	1.000	0.995	0.065	0.997	0.069	0.998	0.044	0.220	0.218	0.021
	$b_{2,18}$	-1.000	-0.996	0.065	0.997	0.070	-0.999	0.045	-0.220	-0.219	0.022
	$b_{2,19}$	1.000	0.995	0.065	0.997	0.070	0.998	0.044	-0.220	-0.219	0.022
	$b_{2,20}$	-1.000	-0.996	0.065	0.997	0.070	-1.000	0.045	0.220	0.219	0.021
	$b_{2,21}$	1.000	0.996	0.065	0.997	0.070	0.999	0.044	-0.220	-0.219	0.022
	$b_{2,22}$	-1.000	-0.995	0.064	0.997	0.070	-1.000	0.045	-0.220	-0.219	0.022
	$b_{2,23}$	1.000	0.996	0.065	0.997	0.070	0.998	0.045	0.220	0.218	0.022
	$b_{2,24}$	-1.000	-0.996	0.065	0.997	0.070	-1.000	0.044	-0.220	-0.219	0.022
	$b_{2,25}$	1.000	0.996	0.065	0.997	0.070	0.999	0.044	-0.220	-0.219	0.022
	$b_{2,26}$	-1.000	-0.995	0.065	0.997	0.070	-1.000	0.045	0.220	0.218	0.022
	$b_{2,27}$	1.000	0.995	0.065	0.996	0.069	0.999	0.044	-0.220	-0.219	0.022
	$b_{2,28}$	-1.000	-0.996	0.065	0.997	0.070	-0.999	0.045	-0.220	-0.219	0.022
	$b_{2,29}$	1.000	0.996	0.065	0.997	0.070	0.999	0.043	0.220	0.218	0.022
	$b_{2,30}$	-1.000	-0.996	0.064	-0.997	0.070	-1.000	0.045	-0.220	-0.219	0.022

Table B.3: **Monte Carlo results from estimating DFM-WN with 2 factors by Indirect Inference (IndInf) and Maximum Likelihood (ML)** : $n = 8$ series, $T = 3773$ observations, $R = 1000$ replications.

Estimation			ML		IndInf	
Parameters		True	Mean	Std. dev	Mean	Std. dev
$\mathbf{\Phi}$	ϕ_1	0.920	0.920	0.007	0.919	0.007
	ϕ_2	0.970	0.970	0.004	0.969	0.004
$\mathbf{\Sigma}$	σ_1^2	0.200	0.200	0.005	0.200	0.006
	σ_2^2	0.200	0.200	0.005	0.200	0.006
	σ_3^2	0.200	0.200	0.006	0.200	0.007
	σ_4^2	0.200	0.200	0.006	0.200	0.007
	σ_5^2	0.200	0.200	0.006	0.200	0.007
	σ_6^2	0.200	0.200	0.006	0.200	0.006
	σ_7^2	0.200	0.200	0.006	0.200	0.008
	σ_8^2	0.200	0.200	0.006	0.200	0.006
$\mathbf{B}$	$b_{1,1}$	1.000	1.000	0.041	0.999	0.043
	$b_{1,2}$	1.000	1.000	0.044	1.000	0.085
	$b_{1,3}$	1.000	1.000	0.044	1.000	0.083
	$b_{1,4}$	1.000	1.000	0.043	0.999	0.085
	$b_{1,5}$	1.000	1.000	0.043	1.000	0.083
	$b_{1,6}$	1.000	1.000	0.044	0.999	0.084
	$b_{1,7}$	1.000	1.000	0.044	1.000	0.083
	$b_{1,8}$	1.000	1.000	0.044	0.999	0.085
	$b_{2,2}$	1.000	1.002	0.065	0.998	0.070
	$b_{2,3}$	-1.000	-1.001	0.064	0.998	0.070
	$b_{2,4}$	1.000	1.002	0.064	0.998	0.070
	$b_{2,5}$	-1.000	-1.001	0.065	0.999	0.070
	$b_{2,6}$	1.000	1.002	0.064	0.998	0.070
	$b_{2,7}$	-1.000	-1.001	0.064	0.998	0.070
	$b_{2,8}$	1.000	1.002	0.064	0.998	0.070

Table B.4: **Descriptive statistics of daily RVar's and log-RVar's:** from 01.11.2001 until 19.12.2016, $T = 3773$ observations. The daily RVar's are upscaled by 10^4.

Stock	***RVar***				***log-RVar***			
	Mean	Std. dev	Skewness	Kurtosis	Mean	Std. dev	Skewness	Kurtosis
AA	4.709	9.039	12.059	257.669	1.059	0.854	0.899	4.365
AXP	3.392	8.505	11.752	262.116	0.358	1.138	0.870	3.603
BA	2.309	3.546	6.888	74.425	0.386	0.857	0.660	3.661
BAC	5.104	17.677	11.367	206.801	0.514	1.215	1.003	4.430
C	6.231	27.222	16.425	394.282	0.666	1.203	1.070	4.677
CAT	2.807	4.810	9.086	147.407	0.572	0.843	0.783	4.184
CVX	1.974	4.639	20.068	649.448	0.187	0.851	0.795	4.531
DD	2.217	3.859	10.416	203.740	0.321	0.864	0.760	3.908
DIS	2.328	4.466	14.565	425.553	0.295	0.930	0.775	3.629
GE	2.623	6.953	10.117	152.161	0.183	1.050	0.922	4.222
GS	3.521	12.342	19.801	537.667	0.557	0.927	1.244	5.645
HD	2.512	5.243	16.881	516.306	0.379	0.911	0.812	3.870
HON	2.458	5.494	25.150	1038.538	0.348	0.951	0.524	3.515
HPQ	3.183	5.587	11.182	201.656	0.703	0.849	0.683	3.975
IBM	1.509	2.911	10.171	173.764	-0.079	0.843	0.986	4.612
IP	3.566	6.882	7.159	82.480	0.683	0.931	0.944	4.214
JNJ	1.076	2.336	15.259	407.614	-0.451	0.874	0.928	4.422
JPM	4.109	11.080	10.675	170.016	0.549	1.119	0.905	3.846
KO	1.142	2.185	15.254	427.196	-0.309	0.813	0.850	4.562
MCD	1.599	3.347	22.978	904.561	-0.050	0.920	0.592	3.474
MMM	1.481	4.708	36.617	1762.923	-0.107	0.844	0.770	4.749
MO	1.479	2.804	8.542	106.393	-0.097	0.840	0.970	4.774
MRK	2.064	4.616	15.134	370.526	0.207	0.867	0.875	4.534
NKE	2.066	3.214	8.966	160.169	0.291	0.825	0.859	3.868
PFE	1.824	2.908	10.101	195.893	0.184	0.812	0.744	4.066
PG	1.074	2.919	25.163	840.958	-0.371	0.776	0.964	5.459
UTX	1.751	3.396	12.816	275.795	0.095	0.838	0.790	4.407
VZ	1.890	3.926	13.817	360.298	0.076	0.914	0.894	4.084
WMT	1.409	2.946	20.122	687.566	-0.122	0.836	0.830	4.266
XOM	1.798	4.484	23.189	851.313	0.085	0.852	0.833	4.637

Table B.5: **Estimation results for DFM-WN and DFM-AR models**: $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations.

Parameters \ Model		DFM-WN			DFM-AR		
		1 Factors	2 Factors	3 Factors	1 Factors	2 Factors	3 Factors
$\boldsymbol{\Phi}$	ϕ_1	0.9202	0.9253	0.9155	0.9135	0.9200	0.9109
	ϕ_2		0.9757	0.8610		0.9287	0.8217
	ϕ_3			0.9794			0.9340
$\boldsymbol{\Lambda}$	ρ				0.5670	0.4960	0.4579
$\boldsymbol{\Sigma}$	σ_1^2	0.2666	0.2492	0.2237	0.2660	0.2485	0.2239
	σ_2^2	0.2783	0.1771	0.1743	0.2774	0.1763	0.1761
	σ_3^2	0.1701	0.1662	0.1624	0.1690	0.1657	0.1630
	σ_4^2	0.6684	0.1268	0.1269	0.6683	0.1268	0.1260
	σ_5^2	0.4644	0.1135	0.1067	0.4645	0.1135	0.1059
	σ_6^2	0.1720	0.1704	0.1631	0.1727	0.1723	0.1609
	σ_7^2	0.2517	0.2576	0.0234	0.2552	0.2591	0.0234
	σ_8^2	0.1459	0.1484	0.1437	0.1468	0.1493	0.1438
	σ_9^2	0.1836	0.1617	0.1514	0.1850	0.1602	0.1530
	σ_{10}^2	0.2119	0.1829	0.1713	0.2091	0.1842	0.1724
	σ_{11}^2	0.2271	0.1698	0.1633	0.2290	0.1714	0.1629
	σ_{12}^2	0.1699	0.1658	0.1663	0.1720	0.1666	0.1662
	σ_{13}^2	0.1809	0.1551	0.1420	0.1823	0.1563	0.1418
	σ_{14}^2	0.3371	0.3246	0.3195	0.3400	0.3249	0.3212
	σ_{15}^2	0.1520	0.1468	0.1459	0.1522	0.1456	0.1451
	σ_{16}^2	0.3077	0.2364	0.2413	0.3050	0.2329	0.2410
	σ_{17}^2	0.2561	0.2174	0.2061	0.2509	0.2176	0.2070
	σ_{18}^2	0.2361	0.1146	0.1018	0.2367	0.1139	0.1013
	σ_{19}^2	0.1789	0.1561	0.1483	0.1783	0.1559	0.1485
	σ_{20}^2	0.3174	0.2080	0.2049	0.3162	0.2073	0.2053
	σ_{21}^2	0.1432	0.1443	0.1425	0.1432	0.1440	0.1428
	σ_{22}^2	0.2772	0.2546	0.2550	0.2769	0.2541	0.2538
	σ_{23}^2	0.2410	0.2246	0.2270	0.2407	0.2234	0.2272
	σ_{24}^2	0.1715	0.1710	0.1694	0.1703	0.1704	0.1693
	σ_{25}^2	0.1972	0.1904	0.1921	0.1962	0.1914	0.1900
	σ_{26}^2	0.1679	0.1545	0.1583	0.1687	0.1550	0.1582
	σ_{27}^2	0.1261	0.1193	0.1190	0.1267	0.1188	0.1196
	σ_{28}^2	0.2403	0.2045	0.1895	0.2387	0.2051	0.1889
	σ_{29}^2	0.2089	0.1755	0.1761	0.2097	0.1761	0.1750
	σ_{30}^2	0.1932	0.1870	0.0645	0.1939	0.1872	0.0645
Total nr. of parameters		61	91	121	62	92	122
BIC		-55.45	-58.99	-63.66	-62.44	-65.80	-69.90

Table B.6: **Cont'd estimation results for DFM-WN and DFM-AR models**: $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations.

Parameters \ Model		DFM-WN			DFM-AR		
		1 Factors	2 Factors	3 Factors	1 Factors	2 Factors	3 Factors
$\mathbf{B}$	$b_{1,1}$	0.6673	0.6740	0.7097	0.6709	0.6726	0.7081
	$b_{1,2}$	0.9827	1.0312	1.0107	0.9813	1.0356	1.0095
	$b_{1,3}$	0.7370	0.7046	0.6916	0.7397	0.7000	0.6895
	$b_{1,4}$	0.8828	1.0077	1.0077	0.8840	1.0192	1.0190
	$b_{1,5}$	0.9611	1.0689	1.0513	0.9602	1.0775	1.0584
	$b_{1,6}$	0.7154	0.7176	0.7411	0.7160	0.7191	0.7412
	$b_{1,7}$	0.6750	0.6412	0.7662	0.6796	0.6387	0.7643
	$b_{1,8}$	0.7550	0.7474	0.7699	0.7570	0.7450	0.7689
	$b_{1,9}$	0.8063	0.7689	0.7436	0.8067	0.7645	0.7374
	$b_{1,10}$	0.9230	0.9360	0.9052	0.9259	0.9382	0.9020
	$b_{1,11}$	0.7764	0.8052	0.8213	0.7803	0.8063	0.8215
	$b_{1,12}$	0.7917	0.7664	0.7706	0.7912	0.7598	0.7669
	$b_{1,13}$	0.8297	0.7780	0.7621	0.8276	0.7694	0.7544
	$b_{1,14}$	0.5999	0.5767	0.5714	0.5976	0.5731	0.5677
	$b_{1,15}$	0.7297	0.6957	0.7078	0.7312	0.6903	0.7030
	$b_{1,16}$	0.7265	0.7642	0.7645	0.7262	0.7673	0.7660
	$b_{1,17}$	0.6919	0.6465	0.6307	0.6932	0.6407	0.6235
	$b_{1,18}$	0.9882	1.0355	1.0116	0.9911	1.0396	1.0136
	$b_{1,19}$	0.6779	0.6400	0.6296	0.6803	0.6367	0.6260
	$b_{1,20}$	0.7170	0.6375	0.6393	0.7222	0.6278	0.6327
	$b_{1,21}$	0.7375	0.7047	0.7185	0.7386	0.6997	0.7145
	$b_{1,22}$	0.6379	0.5989	0.6084	0.6376	0.5940	0.6017
	$b_{1,23}$	0.6965	0.6606	0.6740	0.6984	0.6572	0.6675
	$b_{1,24}$	0.6974	0.6899	0.6997	0.7003	0.6878	0.6980
	$b_{1,25}$	0.6623	0.6286	0.6355	0.6623	0.6231	0.6320
	$b_{1,26}$	0.6414	0.6118	0.6187	0.6412	0.6096	0.6140
	$b_{1,27}$	0.7438	0.7064	0.7031	0.7476	0.7025	0.6966
	$b_{1,28}$	0.7553	0.7083	0.6843	0.7580	0.7024	0.6756
	$b_{1,29}$	0.6898	0.6389	0.6518	0.6919	0.6335	0.6452
	$b_{1,30}$	0.7138	0.6681	0.7603	0.7156	0.6612	0.7554
	$b_{2,2}$		0.0793	-0.2999		0.0808	-0.3012
	$b_{2,3}$		-0.2298	-0.1766		-0.2307	-0.1759
	$b_{2,4}$		0.5049	-0.3137		0.5096	-0.3190
	$b_{2,5}$		0.3457	-0.3724		0.3493	-0.3776
	$b_{2,6}$		-0.0745	-0.0421		-0.0754	-0.0414
	$b_{2,7}$		-0.1837	0.3482		-0.1866	0.3476
	$b_{2,8}$		-0.1123	-0.0647		-0.1135	-0.0661
	$b_{2,9}$		-0.2952	-0.1978		-0.2953	-0.1966
	$b_{2,10}$		-0.0377	-0.2946		-0.0383	-0.2938
	$b_{2,11}$		0.0629	-0.1250		0.0652	-0.1248
	$b_{2,12}$		-0.2195	-0.1233		-0.2209	-0.1240
	$b_{2,13}$		-0.3302	-0.1910		-0.3317	-0.1909
	$b_{2,14}$		-0.2184	-0.1202		-0.2231	-0.1178
	$b_{2,15}$		-0.2316	-0.0750		-0.2355	-0.0728

Table B.7: **Cont'd estimation results for DFM-WN and DFM-AR models**: $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations. The entries without any font correspond to p-values smaller than 0.01, the ones in italics to p-values between 0.01 and 0.05, the ones in bold to p-values between 0.05 and 0.10 and the ones in italics and bold to p-values larger than 0.10.

Parameters \ Model		DFM-WN			DFM-AR		
		1 Factors	2 Factors	3 Factors	1 Factors	2 Factors	3 Factors
	$b_{2,16}$		0.1017	-0.1777		0.1031	-0.1822
	$b_{2,17}$		-0.3332	-0.1381		-0.3341	-0.1383
	$b_{2,18}$		0.0940	-0.3407		0.0969	-0.3418
	$b_{2,19}$		-0.2886	-0.1446		-0.2912	-0.1444
	$b_{2,20}$		-0.4601	*-0.0658*		-0.4678	-0.0636
	$b_{2,21}$		-0.2161	-0.0797		-0.2178	-0.0799
	$b_{2,22}$		-0.2692	*-0.0468*		-0.2726	-0.0447
	$b_{2,23}$		-0.2674	***-0.0258***		-0.2691	-0.0243
	$b_{2,24}$		-0.0955	-0.1054		-0.0933	-0.1074
	$b_{2,25}$		-0.2301	-0.0899		-0.2323	-0.0890
	$b_{2,26}$		-0.2411	*-0.0430*		-0.2426	-0.0428
	$b_{2,27}$		-0.2498	-0.1121		-0.2517	-0.1131
	$b_{2,28}$		-0.3406	-0.1691		-0.3427	-0.1700
	$b_{2,29}$		-0.3114	***-0.0348***		-0.3146	-0.0342
	$b_{2,30}$		-0.2613	0.2307		-0.2638	0.2302
	$b_{3,3}$			-0.2993			-0.2958
	$b_{3,4}$			0.4452			0.4419
	$b_{3,5}$			0.2669			0.2643
	$b_{3,6}$			-0.0906			-0.0909
	$b_{3,7}$			-0.1115			-0.1084
	$b_{3,8}$			-0.1399			-0.1406
$\mathbf{B}$	$b_{3,9}$			-0.3734			-0.3710
	$b_{3,10}$			-0.1248			-0.1241
	$b_{3,11}$			0.0290			0.0275
	$b_{3,12}$			-0.2667			-0.2640
	$b_{3,13}$			-0.4117			-0.4074
	$b_{3,14}$			-0.2745			-0.2727
	$b_{3,15}$			-0.2680			-0.2643
	$b_{3,16}$			0.0608			0.0617
	$b_{3,17}$			-0.3980			-0.3925
	$b_{3,18}$			***0.0057***			0.0066
	$b_{3,19}$			-0.3485			-0.3426
	$b_{3,20}$			-0.5063			-0.4995
	$b_{3,21}$			-0.2568			-0.2562
	$b_{3,22}$			-0.3020			-0.2984
	$b_{3,23}$			-0.2936			-0.2902
	$b_{3,24}$			-0.1315			-0.1274
	$b_{3,25}$			-0.2684			-0.2611
	$b_{3,26}$			-0.2724			-0.2708
	$b_{3,27}$			-0.3033			-0.3006
	$b_{3,28}$			-0.4068			-0.3988
	$b_{3,29}$			-0.3471			-0.3440
	$b_{3,30}$			-0.2280			-0.2263

Table B.8: **Estimation results for DFM-ARCH and DFM-GARCH models**: $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations.

Parameters \ Model		DFM-ARCH			DFM-GARCH		
		1 Factors	2 Factors	3 Factors	1 Factors	2 Factors	3 Factors
$\mathbf{\Phi}$	ϕ_1	0.9220	0.9267	0.9168	0.9188	0.9241	0.9130
	ϕ_2		0.9757	0.9794		0.9757	0.9794
	ϕ_3			0.8610			0.8610
$\boldsymbol{\alpha}$	α	0.1806	0.1637	0.1245	0.1079	0.0803	0.0799
$\boldsymbol{\beta}$	β				0.8141	0.8576	0.8481
$\mathbf{\Sigma}$	σ_1^2	0.2666	0.2492	0.2237	0.2666	0.2492	0.2237
	σ_2^2	0.2783	0.1771	0.1743	0.2783	0.1771	0.1743
	σ_3^2	0.1701	0.1662	0.1624	0.1701	0.1662	0.1624
	σ_4^2	0.6684	0.1268	0.1269	0.6684	0.1268	0.1269
	σ_5^2	0.4644	0.1135	0.1067	0.4644	0.1135	0.1067
	σ_6^2	0.1720	0.1704	0.1631	0.1720	0.1704	0.1631
	σ_7^2	0.2517	0.2577	0.0234	0.2517	0.2577	0.0234
	σ_8^2	0.1459	0.1484	0.1437	0.1459	0.1484	0.1437
	σ_9^2	0.1836	0.1617	0.1514	0.1836	0.1617	0.1514
	σ_{10}^2	0.2119	0.1829	0.1713	0.2119	0.1829	0.1713
	σ_{11}^2	0.2271	0.1698	0.1633	0.2272	0.1698	0.1633
	σ_{12}^2	0.1699	0.1658	0.1663	0.1699	0.1658	0.1663
	σ_{13}^2	0.1809	0.1551	0.1420	0.1809	0.1551	0.1420
	σ_{14}^2	0.3371	0.3246	0.3195	0.3371	0.3246	0.3195
	σ_{15}^2	0.1520	0.1468	0.1459	0.1520	0.1468	0.1459
	σ_{16}^2	0.3077	0.2364	0.2413	0.3077	0.2364	0.2413
	σ_{17}^2	0.2561	0.2174	0.2061	0.2561	0.2174	0.2061
	σ_{18}^2	0.2361	0.1146	0.1018	0.2361	0.1146	0.1018
	σ_{19}^2	0.1789	0.1561	0.1483	0.1789	0.1561	0.1483
	σ_{20}^2	0.3174	0.2080	0.2049	0.3174	0.2080	0.2049
	σ_{21}^2	0.1432	0.1443	0.1425	0.1432	0.1443	0.1425
	σ_{22}^2	0.2772	0.2546	0.2550	0.2772	0.2546	0.2550
	σ_{23}^2	0.2410	0.2246	0.2270	0.2410	0.2246	0.2270
	σ_{24}^2	0.1715	0.1710	0.1694	0.1715	0.1710	0.1694
	σ_{25}^2	0.1972	0.1904	0.1921	0.1972	0.1904	0.1921
	σ_{26}^2	0.1679	0.1545	0.1583	0.1679	0.1545	0.1583
	σ_{27}^2	0.1261	0.1193	0.1190	0.1261	0.1193	0.1190
	σ_{28}^2	0.2403	0.2045	0.1895	0.2403	0.2045	0.1895
	σ_{29}^2	0.2089	0.1755	0.1761	0.2089	0.1755	0.1761
	σ_{30}^2	0.1932	0.1870	0.0645	0.1932	0.1870	0.0645
Total nr. of parameters		62	92	121	63	93	122
BIC		-55.38	-58.92	-63.60	-56.31	-60.00	-64.40

Table B.9: **Cont'd estimation results for DFM-ARCH and DFM-GARCH models**: $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations.

	Model	DFM-ARCH			DFM-GARCH		
Parameters		1 Factors	2 Factors	3 Factors	1 Factors	2 Factors	3 Factors
B	$b_{1,1}$	0.6679	0.6749	0.7104	0.6700	0.6759	0.7125
	$b_{1,2}$	0.9843	1.0330	1.0115	0.9856	1.0353	1.0150
	$b_{1,3}$	0.7385	0.7061	0.6926	0.7400	0.7062	0.6947
	$b_{1,4}$	0.8843	1.0087	1.0070	0.8853	1.0122	1.0118
	$b_{1,5}$	0.9634	1.0698	1.0510	0.9643	1.0736	1.0560
	$b_{1,6}$	0.7164	0.7185	0.7419	0.7182	0.7194	0.7444
	$b_{1,7}$	0.6758	0.6422	0.7669	0.6767	0.6424	0.7683
	$b_{1,8}$	0.7560	0.7488	0.7705	0.7571	0.7499	0.7727
	$b_{1,9}$	0.8077	0.7707	0.7450	0.8093	0.7712	0.7468
	$b_{1,10}$	0.9246	0.9376	0.9062	0.9264	0.9394	0.9093
	$b_{1,11}$	0.7778	0.8063	0.8218	0.7794	0.8082	0.8251
	$b_{1,12}$	0.7930	0.7679	0.7717	0.7946	0.7689	0.7741
	$b_{1,13}$	0.8310	0.7798	0.7639	0.8329	0.7796	0.7658
	$b_{1,14}$	0.6008	0.5778	0.5725	0.6014	0.5776	0.5746
	$b_{1,15}$	0.7308	0.6968	0.7091	0.7318	0.6975	0.7109
	$b_{1,16}$	0.7270	0.7655	0.7651	0.7287	0.7670	0.7686
	$b_{1,17}$	0.6934	0.6483	0.6322	0.6943	0.6483	0.6342
	$b_{1,18}$	0.9902	1.0371	1.0119	0.9921	1.0396	1.0161
	$b_{1,19}$	0.6792	0.6416	0.6309	0.6807	0.6417	0.6322
	$b_{1,20}$	0.7184	0.6391	0.6411	0.7198	0.6393	0.6423
	$b_{1,21}$	0.7385	0.7062	0.7199	0.7406	0.7068	0.7219
	$b_{1,22}$	0.6389	0.6004	0.6096	0.6401	0.6008	0.6106
	$b_{1,23}$	0.6978	0.6626	0.6754	0.6994	0.6630	0.6774
	$b_{1,24}$	0.6991	0.6915	0.7005	0.7006	0.6923	0.7025
	$b_{1,25}$	0.6638	0.6304	0.6363	0.6649	0.6305	0.6378
	$b_{1,26}$	0.6430	0.6136	0.6197	0.6439	0.6135	0.6213
	$b_{1,27}$	0.7455	0.7078	0.7044	0.7465	0.7087	0.7060
	$b_{1,28}$	0.7564	0.7104	0.6861	0.7585	0.7101	0.6882
	$b_{1,29}$	0.6914	0.6401	0.6529	0.6922	0.6403	0.6547
	$b_{1,30}$	0.7147	0.6695	0.7614	0.7161	0.6702	0.7628
	$b_{2,2}$		0.0793	-0.2999		0.0793	-0.2999
	$b_{2,3}$		-0.2298	-0.1766		-0.2298	-0.1767
	$b_{2,4}$		0.5049	-0.3137		0.5049	-0.3137
	$b_{2,5}$		0.3457	-0.3725		0.3456	-0.3724
	$b_{2,6}$		-0.0745	-0.0421		-0.0745	-0.0421
	$b_{2,7}$		-0.1837	0.3482		-0.1837	0.3481
	$b_{2,8}$		-0.1123	-0.0647		-0.1123	-0.0647
	$b_{2,9}$		-0.2952	-0.1978		-0.2952	-0.1978
	$b_{2,10}$		-0.0378	-0.2946		-0.0377	-0.2946
	$b_{2,11}$		0.0629	-0.1250		0.0629	-0.1250
	$b_{2,12}$		-0.2195	-0.1233		-0.2195	-0.1233
	$b_{2,13}$		-0.3302	-0.1910		-0.3302	-0.1910
	$b_{2,14}$		-0.2184	-0.1202		-0.2184	-0.1202
	$b_{2,15}$		-0.2316	-0.0750		-0.2316	-0.0750

Table B.10: **Cont'd estimation results for DFM-ARCH and DFM-GARCH models**: $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations. The entries without any font correspond to p-values smaller than 0.01, the ones in italics to p-values between 0.01 and 0.05 and the ones in bold to p-values between 0.05 and 0.10.

Parameters \ Model		DFM-ARCH 1 Factors	DFM-ARCH 2 Factors	DFM-ARCH 3 Factors	DFM-GARCH 1 Factors	DFM-GARCH 2 Factors	DFM-GARCH 3 Factors
$\mathbf{B}$	$b_{2,16}$		0.1017	-0.1777		0.1017	-0.1777
	$b_{2,17}$		-0.3332	-0.1381		-0.3332	-0.1381
	$b_{2,18}$		0.0940	-0.3407		0.0940	-0.3407
	$b_{2,19}$		-0.2886	-0.1446		-0.2886	-0.1446
	$b_{2,20}$		-0.4601	*-0.0657*		-0.4601	*-0.0658*
	$b_{2,21}$		-0.2161	-0.0797		-0.2161	-0.0797
	$b_{2,22}$		-0.2692	*-0.0468*		-0.2692	*-0.0468*
	$b_{2,23}$		-0.2674	**-0.0258**		-0.2674	**-0.0258**
	$b_{2,24}$		-0.0955	-0.1054		-0.0955	-0.1054
	$b_{2,25}$		-0.2301	-0.0899		-0.2301	-0.0899
	$b_{2,26}$		-0.2411	-0.0430		-0.2411	-0.0430
	$b_{2,27}$		-0.2498	-0.1121		-0.2498	-0.1121
	$b_{2,28}$		-0.3406	-0.1690		-0.3405	-0.1690
	$b_{2,29}$		-0.3115	**-0.0348**		-0.3114	**-0.0348**
	$b_{2,30}$		-0.2613	0.2307		-0.2613	0.2307
	$b_{3,3}$			-0.2993			-0.2993
	$b_{3,4}$			0.4453			0.4452
	$b_{3,5}$			0.2669			0.2669
	$b_{3,6}$			-0.0906			-0.0906
	$b_{3,7}$			-0.1115			-0.1115
	$b_{3,8}$			-0.1399			-0.1399
	$b_{3,9}$			-0.3734			-0.3734
	$b_{3,10}$			-0.1248			-0.1248
	$b_{3,11}$			0.0290			0.0290
	$b_{3,12}$			-0.2667			-0.2667
	$b_{3,13}$			-0.4117			-0.4117
	$b_{3,14}$			-0.2745			-0.2745
	$b_{3,15}$			-0.2680			-0.2680
	$b_{3,16}$			0.0608			0.0608
	$b_{3,17}$			-0.3981			-0.3981
	$b_{3,18}$			**0.0057**			**0.0057**
	$b_{3,19}$			-0.3485			-0.3485
	$b_{3,20}$			-0.5064			-0.5064
	$b_{3,21}$			-0.2568			-0.2568
	$b_{3,22}$			-0.3020			-0.3020
	$b_{3,23}$			-0.2936			-0.2936
	$b_{3,24}$			-0.1315			-0.1315
	$b_{3,25}$			-0.2684			-0.2684
	$b_{3,26}$			-0.2725			-0.2724
	$b_{3,27}$			-0.3034			-0.3033
	$b_{3,28}$			-0.4068			-0.4068
	$b_{3,29}$			-0.3472			-0.3471
	$b_{3,30}$			-0.2281			-0.2280

Table B.11: **BIC values and total number of parameters of alternative models and estimation technique**: $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations.

Model	BIC	Total number of parameters
QML DFM-WN 2 factors	-58.79	91
QML DFM-WN 3 factors	-61.60	120
sARFIMA(1,d,0)	-60.04	2
sARFIMA(0,d,0)	-60.05	1
sARFIMA(1,d,1)	-60.30	3
dARFIMA(1,d,0)	-59.77	60
dARFIMA(0,d,0)	-59.86	30
dARFIMA(1,d,1)	-59.67	90
sHAR	-60.30	3
dHAR	-60.12	90
vHAR	-68.94	2700

Table B.12: **Estimates and standard errors** for the parameters corresponding to the first 7 and 8 series (Alcoa Inc., American Express Company, Boeing Corporation, Bank of America Corporation, Citigroup Inc., Caterpillar Inc., Chevron Corporation and Dupont) when applying DFM-WN on panels including the first $n = 7$, $n = 8$ and all $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations and $k = 2$.

		$n = 7$		$n = 8$		$n = 30$	
Parameters		Estimate	Std. dev	Estimate	Std. dev	Estimate	Std. dev
$\mathbf{\Phi}$	ϕ_1	0.9444	0.0057	0.9366	0.0059	0.9253	0.0060
	ϕ_2	0.9762	0.0164	0.9593	0.0138	0.9757	0.0047
$\mathbf{\Sigma}$	σ_1^2	0.2033	0.0059	0.2148	0.0059	0.2492	0.0060
	σ_2^2	0.2114	0.0058	0.1976	0.0053	0.1771	0.0046
	σ_3^2	0.2315	0.0070	0.2163	0.0063	0.1662	0.0042
	σ_4^2	0.2143	0.0090	0.2031	0.0092	0.1268	0.0050
	σ_5^2	0.0459	0.0083	0.0566	0.0080	0.1135	0.0040
	σ_6^2	0.1167	0.0048	0.1253	0.0040	0.1704	0.0043
	σ_7^2	0.2103	0.0072	0.2171	0.0064	0.2576	0.0063
	σ_8^2			0.1049	0.0038	0.1484	0.0037
$\mathbf{B}$	$b_{1,1}$	0.7302	0.0354	0.7231	0.0331	0.6740	0.0280
	$b_{1,2}$	1.0101	0.0530	1.0313	0.0489	1.0312	0.0422
	$b_{1,3}$	0.7124	0.0346	0.7189	0.0331	0.7046	0.0328
	$b_{1,4}$	0.9856	0.0680	1.0019	0.0625	1.0077	0.0553
	$b_{1,5}$	1.0389	0.0708	1.0606	0.0637	1.0689	0.0503
	$b_{1,6}$	0.7732	0.0366	0.7669	0.0342	0.7176	0.0300
	$b_{1,7}$	0.6988	0.0358	0.6902	0.0334	0.6412	0.0294
	$b_{1,8}$			0.7993	0.0359	0.7474	0.0312
	$b_{2,2}$	-0.2529	0.0229	0.2150	0.0194	0.0793	0.0159
	$b_{2,3}$	0.0579	0.0133	-0.0904	0.0133	-0.2298	0.0192
	$b_{2,4}$	-0.5387	0.0396	0.5413	0.0337	0.5049	0.0376
	$b_{2,5}$	-0.5664	0.0426	0.5360	0.0341	0.3457	0.0282
	$b_{2,6}$	0.0617	0.0122	-0.0661	0.0119	-0.0745	0.0125
	$b_{2,7}$	0.1498	0.0162	-0.1562	0.0150	-0.1837	0.0169
	$b_{2,8}$			-0.0872	0.0125	-0.1123	0.0137

Table B.13: **Estimates and standard errors** for the parameters corresponding to the first 7 and 8 series (Alcoa Inc., American Express Company, Boeing Corporation, Bank of America Corporation, Citigroup Inc., Caterpillar Inc., Chevron Corporation and Dupont) when applying DFM-AR on panels including the first $n = 7$, $n = 8$ and all $n = 30$ series of demeaned daily log-RVar's from 01.11.2001 until 19.12.2016, $T = 3773$ observations and $k = 2$.

		$n=7$		$n=8$		$n=30$	
Parameters		Estimate	Std. dev	Estimate	Std. dev	Estimate	Std. dev
$\mathbf{\Phi}$	ϕ_1	0.9230	0.0064	0.9217	0.0064	0.9200	0.0062
	ϕ_2	0.8313	0.0129	0.8390	0.0123	0.9287	0.0063
$\mathbf{\Lambda}$	λ	0.5404	0.0061	0.5105	0.0057	0.4960	0.0028
$\mathbf{\Sigma}$	σ_1^2	0.2002	0.0080	0.2132	0.0077	0.2485	0.0078
	σ_2^2	0.2097	0.0076	0.1963	0.0067	0.1763	0.0058
	σ_3^2	0.2321	0.0092	0.2166	0.0082	0.1657	0.0054
	σ_4^2	0.2138	0.0121	0.2021	0.0120	0.1268	0.0065
	σ_5^2	0.0465	0.0109	0.0577	0.0102	0.1135	0.0053
	σ_6^2	0.1174	0.0061	0.1254	0.0052	0.1723	0.0056
	σ_7^2	0.2113	0.0097	0.2178	0.0084	0.2591	0.0084
	σ_8^2			0.1054	0.0048	0.1493	0.0047
$\mathbf{B}$	$b_{1,1}$	0.7352	0.0329	0.7264	0.0323	0.6726	0.0289
	$b_{1,2}$	1.0178	0.0447	1.0343	0.0442	1.0356	0.0418
	$b_{1,3}$	0.7143	0.0324	0.7209	0.0321	0.7000	0.0313
	$b_{1,4}$	0.9965	0.0517	1.0046	0.0515	1.0192	0.0510
	$b_{1,5}$	1.0494	0.0530	1.0634	0.0522	1.0775	0.0479
	$b_{1,6}$	0.7761	0.0327	0.7699	0.0323	0.7191	0.0305
	$b_{1,7}$	0.6969	0.0322	0.6897	0.0318	0.6387	0.0309
	$b_{1,8}$			0.8000	0.0334	0.7450	0.0351
	$b_{2,2}$	0.2624	0.0254	0.2198	0.0248	0.0808	0.0237
	$b_{2,3}$	-0.0545	0.0198	-0.0923	0.0190	-0.2307	0.0208
	$b_{2,4}$	0.5535	0.0289	0.5489	0.0289	0.5096	0.0313
	$b_{2,5}$	0.5777	0.0283	0.5417	0.0279	0.3493	0.0278
	$b_{2,6}$	-0.0543	0.0186	-0.0642	0.0178	-0.0754	0.0191
	$b_{2,7}$	-0.1449	0.0206	-0.1558	0.0198	-0.1866	0.0210
	$b_{2,8}$			-0.0875	0.0181	-0.1135	0.0190

Table B.14: **1-step ahead out of sample mean squared errors:** averaged over the 30 assets. 'Window' gives the number of out of sample (daily) observations. The entries in bold stand for the forecasts composing the 95% MCS. KF-DFM-WN stands for DFM-WN estimated by QML.

Model	Window	forecasting log-Rvar 1000	750	500	250	forecasting RVar 1000	750	500	250
KF-DFM-WN	2 factors	1.4657	1.0072	0.6738	0.7381	21.5418	**8.6486**	**5.5578**	5.3217
DFM-WN	1 factor	1.4296	**0.9673**	**0.6370**	**0.6562**	21.7777	8.8556	5.8694	5.2920
	2 factors	**1.3359**	**0.9298**	**0.6601**	**0.7141**	**20.8604**	**8.3141**	**5.5483**	**4.9955**
	3 factors	1.4644	1.0166	0.6990	0.7534	21.3972	**8.7523**	**5.5482**	5.4122
DFM-AR	1 factor	1.4784	1.0342	0.7189	0.7648	21.6196	8.8588	5.8913	5.5968
	2 factors	1.4847	1.0384	0.7186	0.7806	21.5449	8.7699	5.7573	5.5300
	3 factors	1.4847	1.0425	0.7305	0.7898	21.5380	8.7663	5.7721	5.6013
DFM-ARCH	1 factor	1.4316	0.9687	**0.6371**	**0.6562**	21.8451	8.9398	5.9921	5.2938
	2 factors	1.4574	0.9993	**0.6670**	0.7306	21.5136	**8.6385**	**5.5652**	5.2998
	3 factors	1.4651	1.0168	0.6988	0.7535	21.4125	**8.5930**	**5.5738**	5.4159
DFM-GARCH	1 factor	1.4330	0.9703	0.6396	0.6584	21.8826	8.9888	6.0624	5.3048
	2 factors	1.4592	1.0011	0.6693	0.7327	21.5455	**8.6775**	**5.6191**	5.3134
sARFIMA(1,d,0)		1.5140	1.0479	0.7048	0.7920	21.1739	8.3368	5.1401	5.6718
dARFIMA(1,d,0)		1.5069	1.0411	0.6996	0.7864	21.1349	8.3016	5.1072	5.6297
sARFIMA(0,d,0)		1.5072	1.0408	0.6993	0.7862	21.1380	8.3065	5.1115	5.6382
dARFIMA(0,d,0)		1.5010	1.0355	**0.6952**	0.7815	**21.1105**	**8.2817**	**5.0886**	5.6107
sARFIMA(1,d,1)		1.5322	1.0588	0.7097	0.8026	21.2775	8.3970	5.1726	5.7392
dARFIMA(1,d,1)		1.5262	1.0536	0.7072	0.7966	**23.4628**	8.3908	5.1758	5.7355
sHAR(1,5,20)		1.5199	1.0539	0.7156	0.8073	21.2718	8.3992	5.1923	5.7957
dHAR(1,5,20)		1.5164	1.0515	**0.7133**	0.8043	21.2657	8.3909	5.1861	5.7856
vHAR(1,5,20)		1.6191	1.1434	0.7855	0.8807	21.7520	8.8006	5.5761	6.2321

Table B.15: **5- and 10-step ahead out of sample mean squared errors:** averaged over the 30 assets and over the number of steps ahead. The forecasts are for the log-RVar series. 'Window' gives the number of out of sample (daily) observations. The entries in bold stand for the forecasts composing the 95% MCS. KF-DFM-WN stands for DFM-WN estimated by QML.

Model	Window	5-step ahead 1000	750	500	250	10-step ahead 1000	750	500	250
KF-DFM-WN	2 factors	3.6734	3.3718	2.5258	2.7980	5.2992	5.0967	3.9465	4.6005
DFM-WN	1 factor	3.4852	3.1942	2.3467	**2.4489**	4.9453	4.7849	**3.6222**	**3.9511**
	2 factors	**3.1264**	**2.8946**	**2.1779**	**2.3742**	**4.5474**	**4.4372**	**3.4603**	**3.9285**
	3 factors	3.6651	3.3869	2.5992	2.8768	5.3044	5.1462	4.1041	4.7827
DFM-AR	1 factor	3.5619	3.2970	2.5036	2.7431	5.0884	4.9577	3.9060	4.5233
	2 factors	3.6278	3.3495	2.5448	2.8359	5.2148	5.0521	3.9829	4.6971
	3 factors	3.6236	3.3549	2.5739	2.8678	5.2198	5.0790	4.0570	4.7828
DFM-ARCH	1 factor	3.4958	3.2043	2.3533	2.4566	4.9636	4.8027	3.6346	3.9649
	2 factors	3.6447	3.3399	2.4982	2.7650	5.2532	5.0433	3.9001	4.5404
	3 factors	3.6704	3.3919	2.6038	2.8841	5.3132	5.1553	4.1132	4.7961
DFM-GARCH	1 factor	3.4900	3.1980	2.3490	2.4531	4.9525	4.7905	**3.6265**	3.9573
	2 factors	3.6422	3.3364	2.4955	2.7632	5.2478	5.0360	3.8950	4.5360
sARFIMA(1,d,0)		3.9059	3.5757	2.7511	3.1144	5.8000	5.5994	4.5683	5.3868
dARFIMA(1,d,0)		3.8815	3.5519	2.7320	3.0907	5.7600	5.5602	4.5360	5.3453
sARFIMA(0,d,0)		3.8828	3.5505	2.7312	3.0911	5.7629	5.5581	4.5352	5.3469
dARFIMA(0,d,0)		3.8625	3.5325	2.7168	3.0724	5.7303	5.5292	4.5119	5.3151
sARFIMA(1,d,1)		3.9973	3.6543	2.8051	3.2101	5.9700	5.7569	4.6904	5.5858
dARFIMA(1,d,1)		3.9609	3.6263	2.7909	3.1721	5.8976	5.7020	4.6586	5.5090
sHAR(1,5,20)		3.9301	3.6001	2.7845	3.1944	5.8625	5.6703	4.6646	5.5674
dHAR(1,5,20)		3.9129	3.5877	2.7747	3.1807	5.8310	5.6479	4.6469	5.5421
vHAR(1,5,20)		4.4316	4.1191	3.2678	3.7051	6.6148	6.4595	5.4091	6.3771

Table B.16: **5- and 10-step ahead out of sample mean squared errors:** averaged over the 30 assets and over the number of steps ahead. The forecasts are for the RVar series. 'Window' gives the number of out of sample (daily) observations. The entries in bold stand for the forecasts composing the 95% MCS. KF-DFM-WN stands for DFM-WN estimated by QML.

		5-step ahead				10-step ahead			
Model \ Window		1000	750	500	250	1000	750	500	250
KF-DFM-WN	2 factors	89.8835	35.3152	20.4462	19.7320	165.3026	67.1811	37.2264	36.9659
DFM-WN	1 factor	90.5201	35.6117	20.9046	19.6241	166.3041	67.4075	37.5916	36.8373
	2 factors	**85.8771**	**32.8043**	**19.1051**	**18.1535**	**1597.1810**	**61.9942**	**34.2132**	**33.9450**
	3 factors	89.1610	34.9550	**20.4153**	20.0295	163.8360	66.4397	37.1454	37.5090
DFM-AR	1 factor	89.5982	35.5412	20.9698	20.4522	164.5118	67.3545	37.8953	38.0541
	2 factors	89.4839	35.4171	20.7519	20.2778	164.3753	67.2334	37.6385	37.7418
	3 factors	89.4769	35.4229	20.8560	20.5746	164.3437	67.2188	37.8142	38.3377
DFM-ARCH	1 factor	90.6150	35.7097	**21.0289**	19.6452	166.4308	67.5190	37.7102	36.8831
	2 factors	89.6748	35.1783	**20.3528**	19.6251	164.8920	66.9191	37.0538	36.7484
	3 factors	89.1864	34.9764	20.4415	20.0476	163.8789	66.4727	37.1827	37.5449
DFM-GARCH	1 factor	90.6464	35.7506	21.0883	19.6589	166.4596	67.5519	37.7618	36.9040
	2 factors	89.7160	35.2199	20.4054	19.6468	164.9451	66.9607	37.1030	36.7823
sARFIMA(1,d,0)		89.2690	35.3187	20.6044	21.3548	165.0264	68.4330	39.2269	40.2531
dARFIMA(1,d,0)		89.1013	35.1752	20.4806	21.1853	164.7047	68.1610	38.9984	39.9245
sARFIMA(0,d,0)		89.1145	35.1991	**20.5033**	21.2241	164.7317	68.2111	39.0463	40.0087
dARFIMA(0,d,0)		89.0008	35.1014	20.4225	21.1210	164.5163	68.0279	38.8997	39.8106
sARFIMA(1,d,1)		89.8747	35.7221	20.9003	21.7933	166.2719	69.2752	39.8637	41.1525
dARFIMA(1,d,1)		93.9817	35.6875	20.9105	21.7877	169.9623	69.1874	39.8639	41.1151
sHAR(1,5,20)		89.8238	35.7053	20.9624	22.0265	166.1373	69.2011	39.9595	41.6208
dHAR(1,5,20)		89.7938	35.6650	20.9314	21.9863	166.0766	69.1173	39.8925	41.5265
vHAR(1,5,20)		92.2547	37.7609	22.9585	24.3275	170.4402	72.6568	43.1927	45.5970

Appendix C: Figures

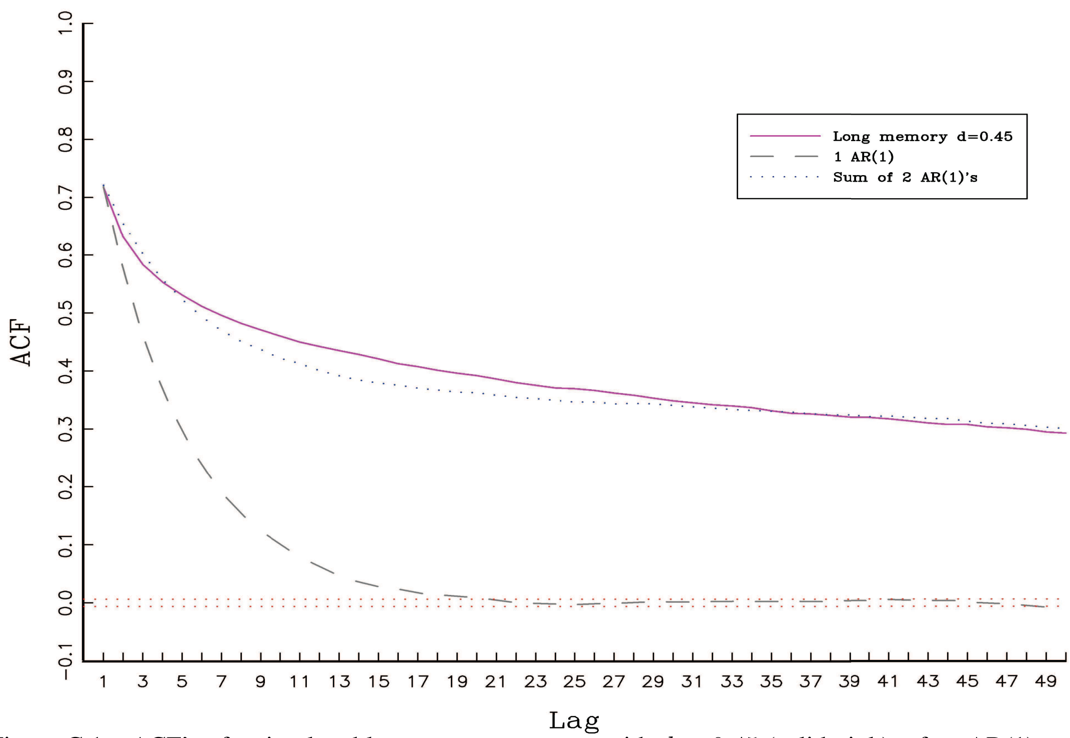

Figure C.1: ACF's of a simulated long memory process with $d = 0.45$ (solid pink), of an AR(1) process with AR parameter equal to 0.8 (dashed grey) and of the sum of two AR(1) processes with AR parameters 0.8 and 0.995, respectively (dotted blue). The length of the simulated series is 100000. The red dotted lines give the 95% confidence interval. The dotted and dashed lines are drawn in such a way that AR(1) and the sum of the AR(1)'s have the same first autocorrelation as the long memory.

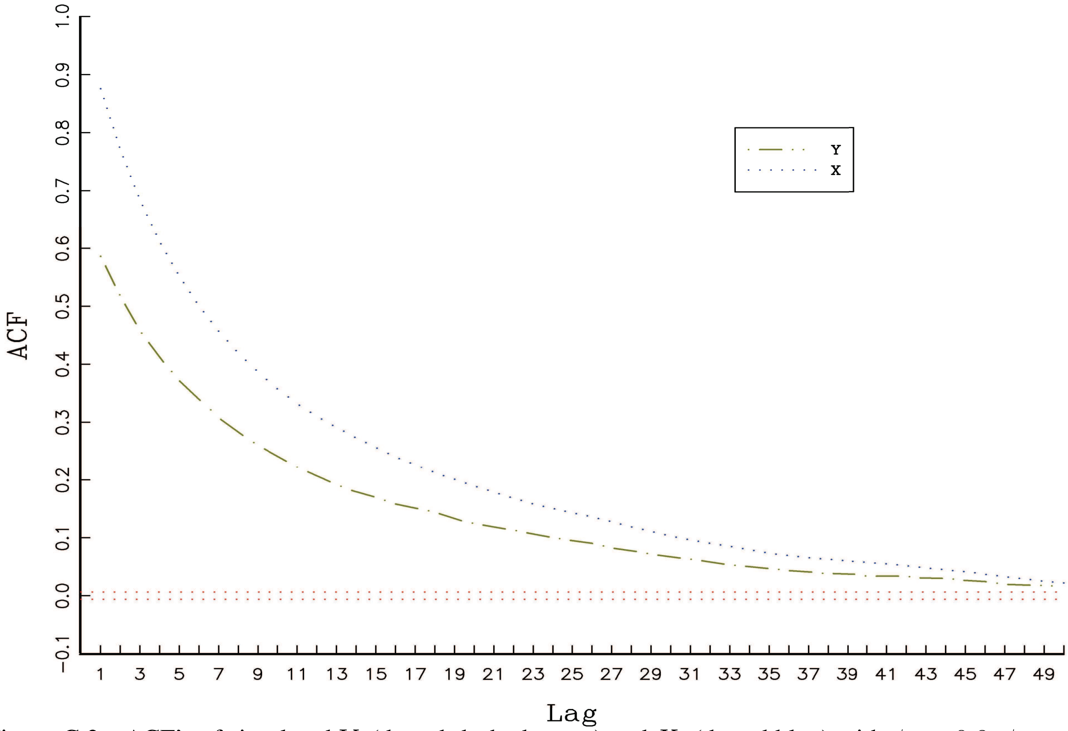

Figure C.2: ACF's of simulated Y_t (dotted-dashed green) and X_t (dotted blue) with $\phi_1 = 0.8$, $\phi_2 = 0.95$, $b_1 = b_2 = 1$ and $\sigma^2 = 1$. The length of the simulated series is 100000. The red dotted lines give the 95% confidence interval.

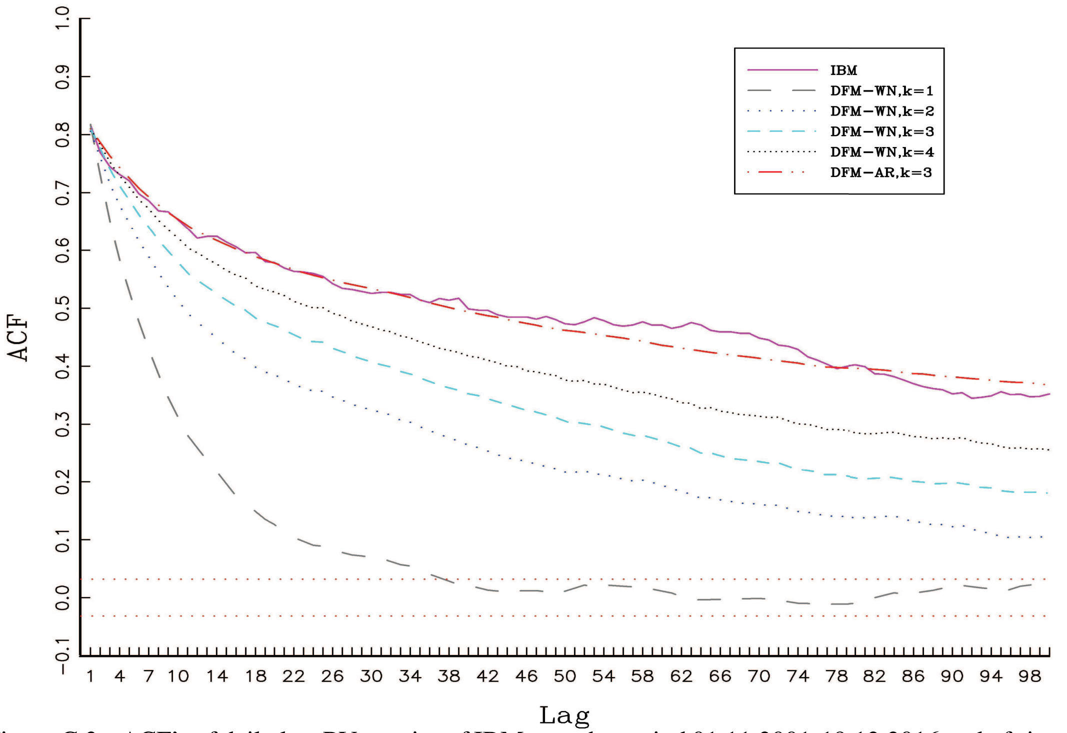

Figure C.3: ACF's of daily log-RVar series of IBM over the period 01.11.2001-19.12.2016 and of simulated DFM-WN processes with $k = 1, 2, 3, 4$ and DFM-AR with $k = 3$. The length of the simulated series is 100000. The red dotted lines give the 95% confidence interval of the ACF of the real data. The non-solid lines are drawn in such a way that the corresponding simulated processes have the same first autocorrelation as the log-RVar of IBM.

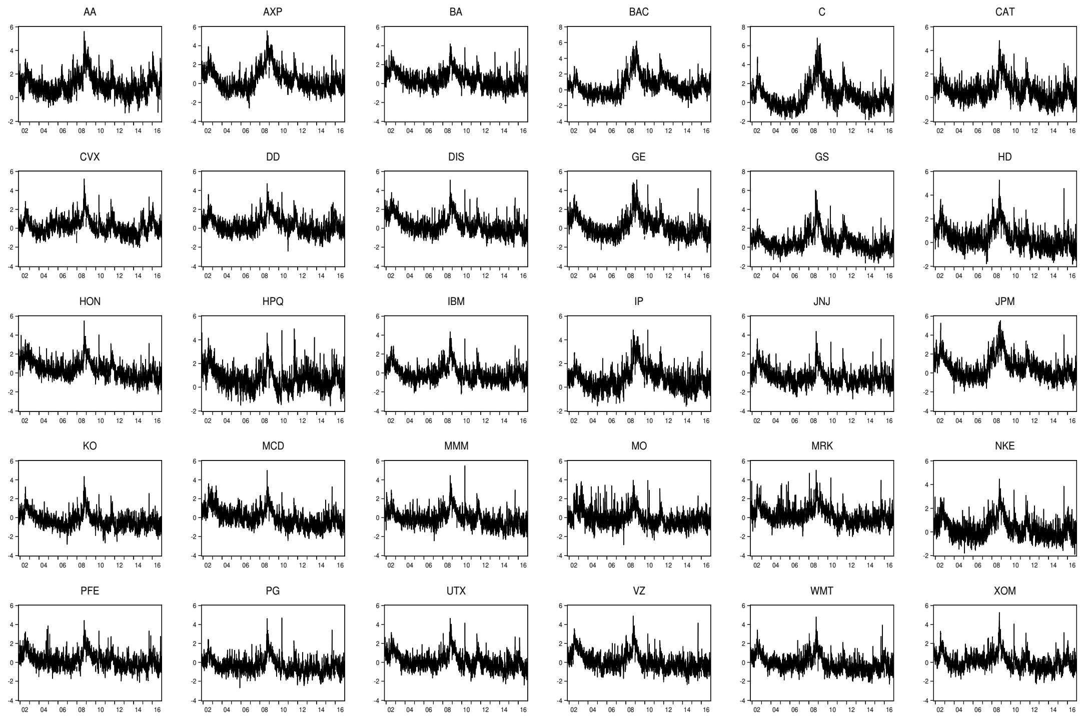

Figure C.4: **Line graph of daily log-RVar's over the period 01.11.2001 – 19.12.2016** ($T = 3773$ trading days). The log-RVar's are scaled up by 10^4. On X-axis we plot the years.

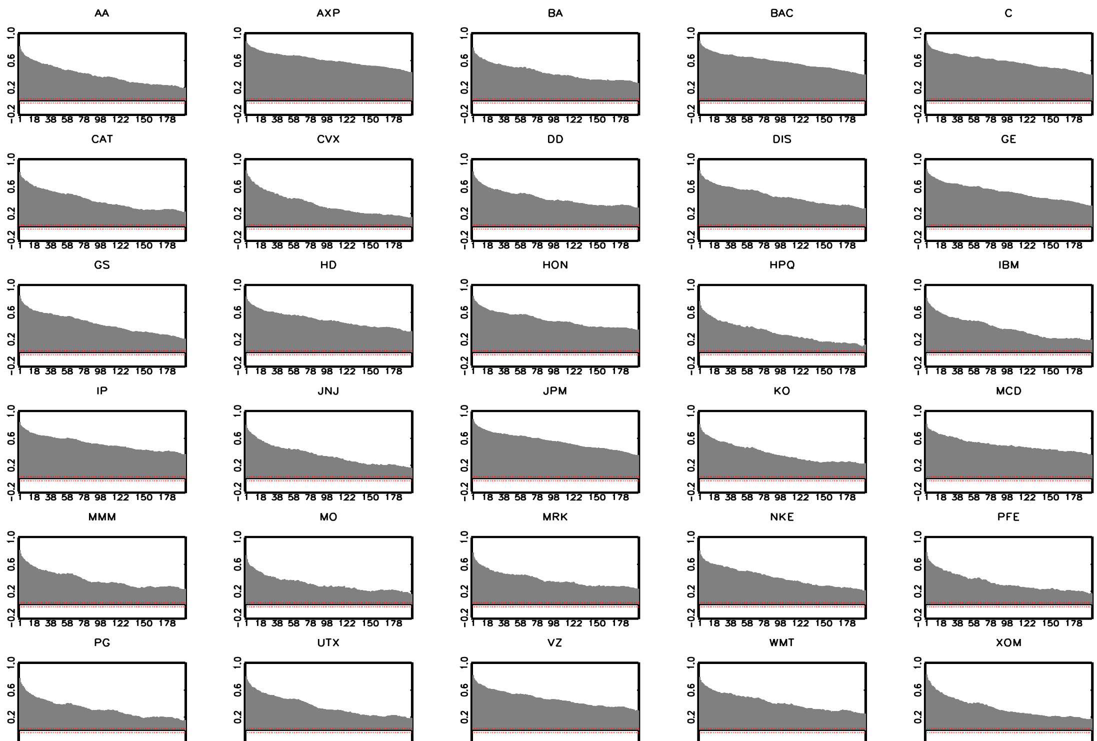

Figure C.5: **ACF of daily log-RVar's up to 200 lags over the period 01.11.2001 – 19.12.2016** ($T = 3773$ trading days). The log-RVar's are scaled up by 10^4. On X-axis we plot the lags. The dotted-line is the 95% confidence interval.