Centre for
Economic
and Financial
Research
at
New Economic
School

February 2015

Modeling and Forecasting Realized Covariance Matrices with Accounting for Leverage

Stanislav Anatolyev
Nikita Kobotaev

Working Paper No 213

CEFIR / NES Working Paper series

Modeling and Forecasting Realized Covariance Matrices with Accounting for Leverage*

Stanislav Anatolyev[†] and Nikita Kobotaev

New Economic School

February 2015

Abstract

The existing dynamic models for realized covariance matrices do not account for an asymmetry with respect to price directions. We modify the recently proposed conditional autoregressive Wishart (CAW) model to allow for the leverage effect. In the conditional threshold autoregressive Wishart (CTAW) model and its variations the parameters governing each asset's volatility and covolatility dynamics are subject to switches that depend on signs of previous asset returns or previous market returns. We evaluate the predictive ability of the CTAW model and its restricted and extended specifications from both statistical and economic points of view. We find strong evidence that many CTAW specifications have a better in-sample fit and tend to have a better out-of-sample predictive ability than the original CAW model and its modifications.

*The article is forthcoming in *Econometric Reviews*. The expected year of publication is 2016.

[†]Corresponding author. Address: Stanislav Anatolyev, New Economic School, 100A Novaya Street, Skolkovo, Moscow, 143026 Russia. E-mail: sanatoly@nes.ru

1 Introduction

Modeling and forecasting covariance matrices of financial asset returns play an important role in asset pricing, portfolio allocation and risk management. The usual approach is to use Multivariate GARCH models, initially introduced in Bollerslev, Engle, and Wooldridge (1988). In these models the covariance matrix, which is not directly observable, is specified as a linear combination of lagged covariance matrices and lagged outer products of returns. Many versions of such models have been proposed, including popular BEKK from Engle and Kroner (1995) and DCC from Engle (2002). These models allow for flexible dynamics of the covariance matrix and guarantee its positive definiteness without imposing restrictions on model parameters. An alternative approach is to use high-frequency return data to construct realized covariance matrices for low-frequency returns. Initially Andersen, Bollerslev, Diebold, and Labys (2003) proposed an approach to estimate a daily realized variance using intradaily data. Then the idea was expanded to cover the multivariate case, see for example Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011), Hautsch, Kyj, and Oomen (2012), Lunde, Shephard, and Sheppard (2011). However, the existing literature on modeling has typically focused on the analysis of univariate realized volatilities or single realized covariances. A problem that arises in the multivariate case is that the predicted covariance matrices are not guaranteed to be positive definite. A usual solution to this problem is to apply a special transformation such that the inverse transformation would ensure positive definiteness of the covariance matrix. In this case, one needs to predict the transformed object and then re-transform it back to the covariance matrix. Andersen, Bollerslev, Diebold, and Labys (2003) proposed to use the Cholesky decomposition of the covariance matrix, Bauer and Vorkink (2011) proposed a matrix logarithmic transformation. One drawback of these approaches is involvement of nonlinear transformations which cause various biases; in addition, they bear a risk of heavy parameterization. Gourieroux, Jasiak, and Sufana (2009) and Golosnoy, Gribisch, and Liesenfeld (2012) avoid such transformations by postulating a flexible dynamic distribution for the entire matrix of realized covariance, the conditional Wishart distribution. The dynamic equation for parameters of this distribution (most of which are related to the conditional mean) allows to guarantee positive definiteness of the modeled object without imposing restrictions on parameters.

The univariate GARCH literatures incorporate the important stylized fact of stock returns that large negative returns increase future volatility while positive returns do not change or even decrease it (see, for example Hentschel, 1995). There are basically two intuitive explanations of this asymmetry called the leverage effect. The first explanation was proposed by Black (1976) and further developed by Christie (1982). According to their approach, after the price of the stock of a firm experiences an unexpected decline, its debt-to-equity ratio (financial leverage ratio) increases which in turn (assuming that the volatility of the whole firm's price remains constant) translates into an increase in the stock return volatility. Choi and Richardson (2008) point out that financial leverage plays a major role in explanation of stock volatility changes; Kroner and Ng (1998) assert that conditional covariances between stocks are affected by the leverage effect as well.

Another explanation of the asymmetric effect is developed in French, Schwert, and Stambaugh (1987), Campbell and Hentschel (1992) and Wu (2001). The authors argue that there is strong evidence of a positive relationship between the market risk premium and volatility. When the future volatility is expected to increase, the risk premium increases as well, and risk averse investors sell the stock putting a downward pressure on the stock price. Thus, an increase in expected future volatility leads to a negative stock return, which is called volatility feedback. As Campbell and Hentschel (1992) argue, the volatility feedback can also explain the asymmetric response of volatility to shocks of different sign.

There are examples of introducing the asymmetry into multivariate GARCH models as well, see Kroner and Ng (1998), De Goeij and Marquering (2004), Cappiello, Engle, and Shephard (2006). However, the multivariate Wishart distribution-based models for realized covariance matrices do not include leverage effects, and the purpose of this paper is to fill this gap. We modify the conditional autoregressive Wishart (CAW) model of Golosnoy, Gribisch, and Liesenfeld (2012) to allow for asymmetry in the volatility dynamics during episodes of upward and downward movements of asset prices. In our most flexible specification a change in a price direction of each asset can shift the dynamics of the this asset's volatility and its covolatilities with the other assets. We refer to this model as the conditional threshold autoregressive Wishart (CTAW) model. It guarantees positive definiteness of the predicted covariance matrices and can be straightforwardly estimated by maximum likelihood.

The general (i.e. unrestricted) CTAW model is very flexible, with the number of parameters increasing faster with the number of assets than in the CAW model because, in addition to the dimensionality of covariance matrices, the number of regimes[1] is increasing in the number of assets as well. In order to avoid possible overfitting and enhance predictive ability we consider restricted specifications of the general CTAW model. One direction of dimension reduction we exploit is to impose restrictions that some parameters do not change when regime switches occur. In particular, we consider versions where persistence parameters are kept constant across different regimes, versions where only diagonal elements of parameter matrices are subject to switches, as well as versions where one asset's reversal of price direction is able to change only parameters of its own dynamics. This leads to a variety of different specifications (diagonal, diagonal-switching, etc.) with substantially different parameter counts. Another direction of serious dimension reduction is to make regime switches be driven by some market indicator (such as the S&P index) common to all assets under consideration, with entire parameter matrices reacting to changes in price directions of this indicator. This leads to variations of what we call a Market CTAW (MCTAW) model.

In addition to a number of restricted specifications, we also consider a mixed data sampling (MIDAS, see Engle, Ghysels, and Sohn (2013)) extension of all CTAW models.

[1]From time to time we will use the terms 'regime', 'switch' and the like adapted in the literature on threshold autoregressions (Lanne and Saikkonen, 2002).

In Golosnoy, Gribisch, and Liesenfeld (2012) the MIDAS-CAW model proved to enhance the in-sample fit and out-of-sample predictive ability. Here, we too incorporate the MIDAS component assuming that only short-run fluctuations are subject to regime switching, while the long-run component is not.

We estimate all models using the dataset from Noureldin, Shephard, and Sheppard (2012) containing daily realized variances and covariances of 5 assets for nearly 9 years. Following Golosnoy, Gribisch, and Liesenfeld (2012), we employ the Bayesian Information Criterion (BIC) and LM test for residual serial correlation as tools of in-sample model selection. In doing out-of-sample forecasting experiments, we construct model confidence sets (MCS) (Hansen, Lunde, and Nason, 2011) under several loss functions, both statistical and economic, in order to make conclusions how statistically significant the differences in predictive ability of different models are.

We find strong evidence that some restricted versions of the CTAW model have a better in-sample fit and tend to have a better out-of-sample predictive abilities than the benchmark CAW models, and so do most of their MIDAS versions compared to MIDAS versions of the benchmark. These results indicate that a possibility to account for leverage effect is important for forecasting realized covariance matrices. There is mixed evidence though whether asset-wise regime switches or those driven by the market indicator deliver better in-sample and out-of-sample performance.

The rest of the paper is organized as follows. Section 2 reviews the notion of a realized covariance matrix and describes the benchmark CAW model and its variations developed in Golosnoy, Gribisch, and Liesenfeld (2012). Section 3 introduces the general CTAW model specification and briefly outlines model estimation. Section 4 describes restricted and extended specifications of the general CTAW model. Section 5 presents an extensive empirical study with details on data, estimation and forecasting strategies and empirical results. Section 6 concludes. The appendix contains numerous tables and graphs, particularly with results of the empirical study.

2 Realized Covariance Matrices and Conditional Autoregressive Wishart Model

Consider a continuous time model for the log-price vector of n assets $P(t)$:

$$dP(t) = M(t)dt + \Omega(t)^{1/2}dW(t),$$

where $M(t)$ and $\Omega(t)^{1/2}$ denote $n \times 1$ instantaneous drift vector and $n \times n$ positive definite 'square-root' of the covariance matrix, respectively, and $W(t)$ is n-dimensional vector of independent Brownian motions. Further we put $M(t) = 0$, and assume that asset returns are linearly independent, i.e. $\Omega(t)$ is positive definite.

Define the realized covariance matrix as

$$Y_t(\Delta) \equiv \sum_{j=1}^{N(\Delta)} R_{t-1+j\Delta} R'_{t-1+j\Delta},$$

where $Y_t(\Delta)$ is realized covariance matrix for day t, Δ is interval length between intradaily observations, $N(\Delta) = 1/\Delta$ is number of observations within one day, $R_{t-1+j\Delta}$ is j-th intradaily return vector for the day t. In the absence of market microstructure noise, as Δ goes to zero, $Y_t(\Delta)$ converges to the integrated covariance matrix of the continuous time stochastic volatility process for day t:

$$ICov_t = \int_{t-1}^{t} \Omega(\tau)d\tau.$$

Now we review the conditional autoregressive Wishart (CAW) model for realized covariance matrices introduced in Golosnoy, Gribisch, and Liesenfeld (2012), which we extend in sections that follow. Another possibility for a point of departure could be the Wishart autoregressive model (WAR) introduced in Gourieroux, Jasiak, and Sufana (2009). However, the CAW model is more convenient to introduce the leverage effects into the dynamics of realized covariance matrices.

Consider a positive definite matrix $Y_t = (y_{ij,t})$ of realized covariances of dimension $n \times n$ computed for periods $t = 1, \ldots, T$. Denote by $F_{t-1} = \{Y_{t-1}, Y_{t-2}, \ldots\}$ the past history of Y_t up to period $t-1$. Golosnoy, Gribisch, and Liesenfeld (2012) assume that

$$Y_t | F_{t-1} \sim W_n(K, \Sigma_t), \tag{1}$$

where $K > n$ is (scalar) number of degrees of freedom, Σ_t is $n \times n$ symmetric positive definite scale matrix, $W_n(\cdot)$ is density of the Wishart distribution. According to the properties of the Wishart distribution,

$$\mathbb{E}_{t-1}[Y_t] = K\Sigma_t,$$
$$cov_{t-1}(y_{ij,t}, y_{lm,t}) = K(\sigma_{il,t}\sigma_{jm,t} + \sigma_{im,t}\sigma_{jl,t}), \quad i, j, m = 1, \ldots, n.$$

The dynamics in the CAW(p, q) model is introduced through the dynamics of $S_t = K\Sigma_t$:

$$S_t = CC' + \sum_{i=1}^{p} B_i S_{t-i} B_i' + \sum_{i=1}^{q} A_i Y_{t-i} A_i', \tag{2}$$

where C is $n \times n$ lower-triangular matrix, and A_i and B_i are $n \times n$ parameter matrices. Further we address S_t as a scale matrix of the Wishart distribution. The equations (1)-(2) define the *general CAW(p,q)* model. The authors argue that the above-defined model is unidentified. The sufficient conditions for identification are that all the diagonal elements of matrix C and first diagonal elements of the matrices A_i and B_i are positive.[2] The total number of parameters to estimate is $(p + q)n^2 + n(n + 1)/2 + 1$, which for the CAW(1,1) specification is $2n^2 + n(n + 1)/2 + 1$. In the case of 5 assets (to be the case in Section 5), CAW(1,1) has 66 parameters to estimate.

In the so called *diagonal CAW(p, q)* model the matrices A_i and B_i are diagonal. The identification conditions in the diagonal model are the same as in the general one, but in

[2] This is a set of convenient restrictions. In fact, it is sufficient to fix a sign of any element of each of the matrices A_i and B_i, not necessarily the first diagonal one and not necessarily at 'positive', see Engle and Kroner (1995). It is, of course, natural to fix the sign to be positive.

the diagonal (1,1) model one would expect the signs of all diagonal elements of A_1 and B_1 to be the same, as both the persistence and impact coefficients are expected to be positive for all realized covolatilities. The total number of parameters to estimate in the diagonal CAW(p, q) model is $(p+q)n + n(n+1)/2 + 1$, which for the diagonal CAW(1,1) specification and 5 assets equals 26.

Finally, Golosnoy, Gribisch, and Liesenfeld (2012) consider the MIDAS (mixed data sampling, see Engle, Ghysels, and Sohn (2013)) versions of the general and diagonal CAW models which show advantage over baseline CAW models in their empirical application. The idea of MIDAS is to decompose the volatility and covolatility movements into short-run and long-run components. The *MIDAS-CAW* model reads

$$S_t = C_t S_t^* C_t', \quad M_t = C_t C_t',$$

where M_t is the long-run component specified according to the MIDAS weighted sum of lagged m-period realized volatilities:

$$M_t = \bar{C}\bar{C}' + \theta \sum_{\ell=1}^{L} \phi_\ell(\omega) \sum_{\tau=t-m\ell}^{t-m(\ell-1)-1} Y_\tau.$$

The short-run component S_t^* follows the CAW dynamics

$$S_t^* = I_n + \sum_{i=1}^{p} B_i (S_{t-i}^* - I_n) B_i' + \sum_{i=1}^{q} A_i (C_{t-i}^{-1} Y_{t-i} (C_{t-i}')^{-1} - I_n) A_i'.$$

In the diagonal MIDAS-CAW model the matrices A_i and B_i are diagonal. The identification issues regarding the matrices A_i and B_i described above still hold. Either MIDAS-CAW model has two more parameters (θ and ω) than the corresponding baseline CAW model.

3 Conditional Threshold Autoregressive Wishart Model

In this section we formulate the *conditional threshold autoregressive Wishart (CTAW)* model. Even though the primary target of our novelty is to account for leverage in the form of a non-symmetric volatility dependence, we follow the tradition and add the qualifier 'threshold' to the name of the model. As in the general CAW model we assume that

$$Y_t|F_{t-1} \sim W_n(K, \Sigma_t), \tag{3}$$

where, as previously, $\Sigma_t = S_t/K$.

We change the volatility equation by allowing for separate changes of the volatility dynamics for different assets depending on signs of their daily returns. That is, the coefficients in the dynamic equation for S_t are different for different directions of past price changes for each asset. However, the 'satiated' model that takes into account all combinations of directions for all assets would have 2^n variations of matrices A_i and B_i, which in the case of CAW(1,1) and 5 assets would result in $2^{5+1}n^2+n(n+1)/2+1 = 1,616$ parameters which is practically infeasible. To achieve a much higher degree of parsimony at the same time keeping the degree of generality high, we invoke a priori restrictions that a change in price direction (or equivalently, a change in sign of return) for a particular asset ('individual leverage effect') changes only those coefficients in A_i and B_i that impact only those realized volatility and covolatilities that are related to this asset. This results in the following *general AB-flexible CTAW* specification:

$$\begin{aligned} S_t = CC' &+ \sum_{i=1}^{p} \left(B_i + \sum_{j=1}^{n} H_{i,j} I_{j,t-i} \right) S_{t-i} \left(B_i + \sum_{j=1}^{n} H_{i,j} I_{j,t-i} \right)' \\ &+ \sum_{i=1}^{q} \left(A_i + \sum_{j=1}^{n} G_{i,j} I_{j,t-i} \right) Y_{t-i} \left(A_i + \sum_{j=1}^{n} G_{i,j} I_{j,t-i} \right)', \end{aligned} \tag{4}$$

where

$$I_{j,t} = \mathbb{I}_{\{r_{j,t}<0\}},\ j = 1 \ldots n$$

are asset-wise directional indicators[3] indicating whether the price of asset j went down in the previous periods; $H_{i,j}$ and $G_{i,j}$ are $n \times n$ matrices of parameters that contain zeros except for the elements on the j-th column and j-th row. We call the state when the past sign is negative *regime 2* and refer to the opposite as *regime 1*. The qualifier 'AB-flexible' emphasizes that all A and B matrices are subject to switches. This specification allows for flexible changes in dynamics for different assets, guarantees positive definiteness of the predicted realized variance matrices, and adds only $(2n-1)(p+q)n+n$ parameters to the original CAW model. In case of (1,1) model for 5 assets this results in a total of $2n^2 + n(n+1)/2 + 2n(2n-1) + 1 = 156$ parameters which is practically feasible to estimate.[4]

As in the case of the general CAW model, it is sufficient for identification are that all diagonal elements of the matrix C and first diagonal elements of the matrices A_i and B_i are positive. No extra restrictions on elements of matrices $H_{i,j}$ and $G_{i,j}$ are needed for identification provided that no two regimes collapse into one (which is highly unlikely with the current definition of indicators and a long enough sample).

In order to gain some insight into the dynamics implied by equation (4), consider the first sum in the equation (denoted $\tilde{S}_t$) in the 3×3, $p = q = 1$ case, when only the first asset switches its regime and the other two stay in regime 1. The dynamics of the second sum in equation (4) is the same up to notation. From now on we consider only (1,1) specifications, so in order to simplify notation we omit indices of B_1 and A_1, the first

[3] In Section 4 we consider a model with another indicator variable – a market-wide directional indicator.

[4] For comparison, in Golosnoy, Gribisch, and Liesenfeld (2012) the Schwarz criterion preferred specification is the unrestricted CAW(2,2) model having 116 parameters; the leading one-step-ahead predicting model in the subprime crisis is the unrestricted CAW(3,2) model having 141 parameters.

index of $H_{1,1}$, and time indexes of $I_{1,t-1}$ and of elements of S_{t-1}. Thus consider

$$\begin{aligned} \tilde{S}_t &= (B + H_1 I_1)\, S_{t-1}\, (B + H_1 I_1)' \\ &= B S_{t-1} B' + (H_1 S_{t-1} H_1' + B S_{t-1} H_1' + H_1 S_{t-1} B')\, I_1, \end{aligned} \tag{5}$$

where

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},\ H_1 = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & 0 & 0 \\ h_{31} & 0 & 0 \end{pmatrix},\ S_{t-1} = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix}.$$

Now we analyze how the dynamics of the elements of $\tilde{S}_t$ change when the first asset switches from regime 1 to regime 2 and all other assets stay in regime 1. Denote the (i,j)-th element of matrix $\tilde{S}_t$ by $\tilde{s}_{i,j}$. The equation for $\tilde{s}_{11}$ can be written as

$$\tilde{s}_{11} = \sum_{k=1}^{3} b_{1k} \sum_{m=1}^{3} s_{km} b_{1m} + \left(\sum_{k=1}^{3} h_{1k} \sum_{m=1}^{3} s_{km} h_{1m} + \sum_{k=1}^{3} b_{1k} \sum_{m=1}^{3} s_{km} h_{1m} + \sum_{k=1}^{3} h_{1k} \sum_{m=1}^{3} s_{km} b_{1m} \right) I_1. \tag{6}$$

The first term would describe the dynamics if all assets were in regime 1. The next three terms appear due to the switch of the first asset to regime 2. Therefore, the coefficients on each element $s_{km},\ k,m = 1,2,3$ change due to the switch of the first asset.

Consider the dynamics for the other elements in, say, the first row of $\tilde{S}_t$ The equations for the elements $\tilde{s}_{1,j},\ j = 2,3$ are

$$\tilde{s}_{1j} = \sum_{k=1}^{3} b_{1k} \sum_{m=1}^{3} s_{km} b_{jm} + \left(\sum_{k=1}^{3} h_{1k} s_{k1} h_{j1} + \sum_{k=1}^{3} b_{1k} s_{k1} h_{j1} + \sum_{k=1}^{3} h_{1k} \sum_{m=1}^{3} s_{km} b_{jm} \right) I_1, \quad j = 2,3.$$

Again, the coefficients on all elements $s_{km},\ k,m = 1,2,3$ change due to the switch of the first asset, but in a different manner than in equation (6).

The fact that in case the daily return for the first asset in period $t-1$ is negative the dynamics of all conditional covariances associated with the first asset changes was documented in De Goeij and Marquering (2004) where the leverage effect was studied in the framework of the MGARCH model. Apart from unobservable vs. observable definition of multivariate volatility, there are two more differences between the CTAW model and leverage augmented MGARCH model of De Goeij and Marquering (2004). First, in the CTAW each element of the matrix modeled depends on all lagged elements of this matrix, while in the leverage augmented MGARCH each element of the matrix modeled depends only on its own lags. Second, the coefficients in different equations of the CTAW model are not fully independent, while in the leverage augmented MGARCH model they are.

Consider the dynamics for the elements that are neither in the first row nor in the first column of $\tilde{S}_t$. The equations for elements $\tilde{s}_{i,j},\ i,j = 2,3$ are

$$\tilde{s}_{ij} = \sum_{k=1}^{3} b_{ik} \sum_{m=1}^{3} s_{km} b_{jm} + \left(h_{i1} s_{11} h_{j1} + \sum_{k=1}^{3} b_{ik} s_{k1} h_{j1} + h_{i1} \sum_{m=1}^{3} s_{1m} b_{jm} \right) I_1, \quad i,j = 2,3.$$

As one can see only the coefficients on covariances with the first asset change.[5] This is reasonable since if the first asset switches to regime 2 and its dynamics change, then its impact on the other assets should also change.

The dynamics of the second sum in equation (4) is the same as the dynamics of the first sum with a change in notation. In case the other assets are in regime 2, a similar analysis may be carried out, while a simultaneous switch can be modeled as a combination of such cases.

The CTAW model may be estimated by maximum likelihood with the following log-likelihood function:

$$\mathcal{L}(\theta) = \sum_{t=1}^{T} \left(-\frac{Kn}{2} \log 2 - \frac{n(n-1)}{4} \log \pi - \sum_{i=1}^{n} \log \Gamma \left(\frac{K+1-i}{2} \right) \right.$$
$$\left. - \frac{K}{2} \log \left| \frac{S_t}{K} \right| + \left(\frac{K-n-1}{2} \right) \log |Y_t| - \frac{1}{2} \operatorname{tr}(K S_t^{-1} Y_t) \right),$$

where K is the number of degrees of freedom, $\Gamma(\cdot)$ is the gamma function, $|\cdot|$ is the

[5] Covariances with the first asset are represented by the elements $s_{k1},\ k = 1,2,3$ and $s_{1m},\ m = 1,2,3$.

determinant operator, and $\mathrm{tr}(\cdot)$ is the trace operator. The set of parameters θ contains K and parameter matrices.

We also consider the *diagonal AB-flexible CTAW* model which introduces leverage effects into the diagonal CAW(p, q) model. In the diagonal AB-flexible CTAW specification the form of the dynamic equation for the scale matrix is given by equation (4) but the parameter matrices $A_i,\ i = 1, \ldots, p,\ B_i,\ i = 1, \ldots, q$ are diagonal, while the parameter matrices $H_{i,j}, i = 1, \ldots, p,\ j = 1, \ldots, n,\ G_{i,j},\ i = 1, \ldots, q,\ j = 1, \ldots, n$ now have zero elements except for the j-th element on the main diagonal:

$$(H_{i,j})_{mk} = \delta_{mk}\delta_{jm}h_{i,j}, \quad i = 1, \ldots, p,\ j, m, k = 1, \ldots, n \tag{7}$$

$$(G_{i,j})_{mk} = \delta_{mk}\delta_{jm}g_{i,j}, \quad i = 1, \ldots, p,\ j, m, k = 1, \ldots, n \tag{8}$$

where $(A)_{mk}$ denotes the (m, k)th element of the matrix A, and δ_{mk} is Kronecker's delta. The identification conditions in the diagonal model are the same as in the general one, but in the diagonal (1,1) model one would expect the signs of all diagonal elements of A_1, $A_1 + \sum_{j=1}^{n} G_{1,j}$, B_1 and $B_1 + \sum_{j=1}^{n} H_{1,j}$ to be the same, as both the persistence and impact coefficients are expected to be positive for all realized covolatilities in all regimes.

Consider the dynamics implied by this model again in the 3×3, $p = q = 1$ case, when only the first asset switches its regime and the other two stay in regime 1. Denote by h the only element in the matrix H_1. Consider $\tilde{S}_t$, given by the equation (5). The equation for the element $\tilde{s}_{11}$ reads

$$\tilde{s}_{11} = b_{11}s_{11}b_{11} + h\left[h(s_{11} + s_{12} + s_{13}) + b_{11}s_{11} + s_{11}b_{11}\right] I_1.$$

The equations for the other elements from the first row of the matrix $\tilde{S}_t$ read

$$\tilde{s}_{1j} = b_{11}s_{1j}b_{jj} + hs_{1j}b_{jj}I_1, \quad j = 2, 3.$$

The equations for the elements $\tilde{s}_{i,j},\ i,j = 2,3$ read

$$\tilde{s}_{ij} = b_{ii} s_{ij} b_{jj}, \quad i,j = 2,3.$$

As one can see only the dynamics for the covariances with the first asset change, and the coefficients only on the lagged covariances with the first asset change, and these changes are characterized by one parameter h.

The total number of parameters in the diagonal AB-flexible CTAW specification is $(p+q)n + n + (p+q)n + 1$. In case of the (1,1) model for 5 assets this results in $2n + n + 2n + 1 = 26$ parameters.

4 Restricted and extended CTAW models

The general CTAW model is very flexible, and it is possible that it would overfit the in-sample data. In order to mitigate this possibility we also consider restricted versions of the general AB-flexible CTAW model. We also consider MIDAS extensions of all models.

4.1 A-flexible CTAW

In the *general A-flexible CTAW* specification the coefficients on the lagged scale matrices S_{t-i} in equation (4) do not depend on the regimes of assets so that the dynamics of the scale matrix are as follows:

$$\begin{aligned} S_t = CC' &+ \sum_{i=1}^{p} B_i S_{t-i} B_i' \\ &+ \sum_{i=1}^{q} \left(A_i + \sum_{j=1}^{n} G_{i,j} I_{j,t-i} \right) Y_{t-i} \left(A_i + \sum_{j=1}^{n} G_{i,j} I_{j,t-i} \right)'. \end{aligned} \tag{9}$$

Such specification can be motivated by the fact that realized covariances Y_t are observable and can be used as proxies for news about the assets. Investors first of all react on news about the stocks, so the leverage effect should affect the coefficients on Y_t. On the contrary, S_t are unobservable and therefore investors cannot react to its changes, thus the coefficients on S_t need not be affected.

The total number of parameters to estimate in this specification is $(p+q)n^2 + n(n+1)/2 + qn(2n-1) + 1$. In case of the (1,1) model for 5 assets this results in $2n^2 + n(n+1)/2 + n(2n-1) + 1 = 111$ parameters.

Likewise, in the *diagonal A-flexible CTAW* specification we impose the restrictions on the diagonal AB-flexible CTAW model that the parameter matrices B_i are the same in all regimes. That is, the form of the dynamic equation for the scale matrix is given by equation (9) but the parameter matrices $A_i,\ i = 1,\dots,p$ and $B_i,\ i = 1,\dots,q$ are diagonal, and the parameter matrices $H_{i,j}, i = 1,\dots,p,\ j = 1,\dots,n,\ G_{i,j},\ i = 1,\dots,q,\ j = 1,\dots,n$ are given by equations (7) and (8). The total number of parameters in this specification is $(p+q)n + n + pn + 1$. In case of the (1,1) model for 5 assets this implies $2n + n + n + 1 = 21$ parameters.

4.2 Diagonal-switching CTAW

In the *diagonal-switching AB-flexible CTAW* specification the form of the dynamic equation for the scale matrix is the same as for the general AB-flexible CTAW model given by equation (4), but the parameter matrices $H_{i,j}, i = 1,\dots,p,\ j = 1,\dots,n,\ G_{i,j},\ i = 1,\dots,q,\ j = 1,\dots,n$ are given by equations (7) and (8).

Denote the only element in the matrix H_1 by h. Consider $\tilde{S}_t$ given by equation (5). The equation for its element $\tilde{s}_{11}$ reads

$$\tilde{s}_{11} = \sum_{k=1}^{3} b_{1k} \sum_{m=1}^{3} s_{km} b_{1m} + h \left[h(s_{11} + s_{12} + s_{13}) + \sum_{k=1}^{3} b_{1k} s_{k1} + \sum_{m=1}^{3} s_{1m} b_{m1} \right] I_1.$$

The equations for the other elements from the first row of the matrix $\tilde{S}_t$ read

$$\tilde{s}_{1j} = \sum_{k=1}^{3} b_{1k} \sum_{m=1}^{3} s_{km} b_{jm} + h \sum_{m=1}^{3} s_{1m} b_{mj} I_1, \quad j = 2, 3.$$

The equations for the elements $\tilde{s}_{i,j},\ i,j = 2,3$ read

$$\tilde{s}_{ij} = \sum_{k=1}^{3} b_{ik} \sum_{m=1}^{3} s_{km} b_{jm}, \quad i,j = 2,3.$$

As one can see, only the dynamics for the covariances with the first asset change, and the coefficients only on lagged covariances with the first asset change. In addition, these changes are characterized by only one parameter h.

The total number of parameters in this specification is $(p+q)n^2 + n(n+1)/2 + (p+q)n+1$. In case of the (1,1) model for 5 assets this results in $2n^2+n(n+1)/2+2n+1 = 76$ parameters.

Likewise, in the *diagonal-switching A-flexible CTAW* specification we impose the restrictions on the diagonal switching AB-flexible CTAW model that the parameter matrices B_i are the same in all regimes. That is, the form of the dynamic equation for the scale matrix is given by equation (9) but the matrices $G_{i,j}$ are given by equation (8).

4.3 Market CTAW

Finally, we consider a restriction on the switching mechanism. In the baseline CTAW model and its variations the changes of volatility dynamics of all assets under consideration are driven by changes of all assets' price directions. In the *Market CTAW* (MCTAW) specification the changes in volatility dynamics of all assets are driven by changes in price directions of a single portfolio representing the market. This may prove useful if individual volatilities and covolatilities are largely driven by the whole market condition which may be characterized by the market portfolio return, and not by idiosyncracies of individual prices. If individual assets often change their direction along with directional changes of the market portfolio,[6] the model parameters describing volatility dynamics in some states may be estimated imprecisely, and pooling these states

[6]Indeed, in our dataset the correlation between the signs of the market return and individual returns varies from 0.46 to 0.55; the correlation between the sign of the market return and the indicator that most of individual returns have the same sign is 0.69.

may have a beneficial effect on precision of estimation and quality of forecasting.

Of course, because such restriction in the switching mechanism deems a lot of parameters in the matrices $H_{i,j}$ and $G_{i,j}$ in (4) separately unidentifiable, appropriate restrictions must be placed on their elements.

Denote the return on the portfolio representing the market at t by $r_{m,t}$, and let $I_{m,t} = \mathbb{I}_{\{r_{m,t}<0\}}$ be the corresponding directional indicator. As with the other CTAW models we consider a number of specifications. Again, all specifications may be broadly divided into two classes: AB-switching and A-switching. In the *AB-flexible MCTAW* model the dynamics for the scale matrix is driven by the following equation:

$$\begin{aligned} S_t = CC' &+ \sum_{i=1}^{p} \left(B_i + H_{i,m} I_{m,t-i}\right) S_{t-i} \left(B_i + H_{i,m} I_{m,t-i}\right)' \\ &+ \sum_{i=1}^{q} \left(A_i + G_{i,m} I_{m,t-i}\right) Y_{t-i} \left(A_i + G_{i,m} I_{m,t-i}\right)'. \end{aligned} \tag{10}$$

In the *A-flexible MCTAW* model the parameter matrices B_i are not subject to switching:

$$\begin{aligned} S_t = CC' &+ \sum_{i=1}^{p} B_i S_{t-i} B_i' \\ &+ \sum_{i=1}^{q} \left(A_i + G_{i,m} I_{m,t-i}\right) Y_{t-i} \left(A_i + G_{i,m} I_{m,t-i}\right)'. \end{aligned} \tag{11}$$

For both classes we consider the general specification, where no restrictions are placed on the parameters (except for the identification restrictions, see below), the diagonal-switching specification, where matrices $G_{i,m},\ i = 1, \ldots, p$ and $H_{i,m},\ i = 1, \ldots, q$ (if applicable) are diagonal, and finally, the diagonal specification, where matrices $A_i,\ G_{i,m},\ i = 1, \ldots, p$ and $B_i,\ H_{i,m},\ i = 1, \ldots, q$ (whichever applicable) are diagonal.

As in the case of the CTAW models, it is sufficient for identification that all diagonal elements of the matrix C and first diagonal elements of the matrices A_i and B_i are positive, and no extra restrictions on elements of matrices $H_{i,m}$ and $G_{i,m}$ are needed for identification provided that the market portfolio exhibits both positive and negative price changes. Analogously to the diagonal (1,1) CTAW, in the diagonal (1,1) MCTAW specification one would expect the signs of all diagonal elements of A_1, $A_1 + G_{1,m}$, B_1 and $B_1 + H_{1,m}$ to be the same.

The total number of parameters to estimate in the general AB-switching specification is $2(p+q)n^2 + n(n+1)/2 + 1$, in the general A-switching specification – $(p+2q)n^2 + n(n+1)/2 + 1$. In case of the (1,1) model for 5 assets this results in 116 and 91 parameters, respectively.

4.4 MIDAS-CTAW

Golosnoy, Gribisch, and Liesenfeld (2012) combine their CAW model with the MIDAS approach to volatility modeling of Engle, Ghysels, and Sohn (2013) and empirically show that the MIDAS-CAW model is able to yield a better in-sample fit and out-of-sample predictability than the baseline CAW model. Motivated by these findings, we also consider MIDAS extensions of all specifications of CTAW models (which we call 'baseline', as opposed to their MIDAS versions), reasonably assuming that regime switches driven by daily returns manifest themselves only in short-run dynamics. We modify equation (2) so that the matrices A_i, $i = 1, \ldots, p$ and possibly B_i, $i = 1, \ldots, q$ are subject to the same switches and restrictions as in the baseline versions. We set the auxiliary MIDAS parameters as in Golosnoy, Gribisch, and Liesenfeld (2012): $m = 20$, $L = 12$ and $\phi_\ell(\omega) \propto (1 - \ell/L)^{\omega-1}$ subject to $\sum_{\ell=1}^{L} \phi_\ell(\omega) = 1$. Any MIDAS-CTAW model has two more parameters (θ and ω) than the corresponding baseline CTAW model.

For readers' convenience, Table 1 summarizes the restrictions placed on model parameters in the baseline versions of the CAW, CTAW and MCTAW models. The restrictions placed on their MIDAS versions are analogous.

5 Empirical Application

5.1 Data Description

In order to compare predictive abilities of the models described above we exploit the dataset constructed by Noureldin, Shephard, and Sheppard (2012). These are 5-min realized covariance matrices for returns of five stocks: American Express (AXP), JPMorgan (JPM), General Electric (GE), DuPont (DP) and International Business Machines (IBM). Realized covariance matrices are calculated with 1-min offset and then averaged in order to cope with microstructure noise and to fully exploit the data. In directional indicators, we use daily returns on the same assets and the S&P500 return (SPY) as a return on the market portfolio. The sample period is from February 2001 to December 2009. The total number of observations is 2,242. Summary statistics on realized variances and covariances and returns are shown in Table 2. As one can see, all variances and covariances are positive-skewed and leptokurtic. The asset returns exhibit a typical pattern, with light skewness and serious excess kurtosis. There are approximately as many positive returns as there are negative ones for each individual asset, while for the market portfolio positive returns exceed negative ones by about 13%.

We divide the whole sample into two parts: in-sample and out-of-sample. The in-sample part is from February 1, 2001 to January 1, 2006. The total number of observations[7] in this part is 1,236. The out-of-sample part is from January 2, 2006 to December 31, 2009. The total number of observations in this part is 1,006. The realized variances for AXP and realized covariances for AXP-JPM are plotted on Figure 1; the out-of-sample part is shaded. As one can see, during the sample period there are both volatile and non-volatile subperiods, the patterns for variance and covariance being quite similar. The out-of-sample part also has volatile and non-volatile episodes. To be able to compare forecasting results during periods of different degree of turbulence, we separately consider two out-of-sample subperiods: the calm one (from January 4, 2006 to August

[7] Each observation is a 5×5 realized covariance matrix.

28, 2008) and the highly volatile one (from August 29, 2008 to June 6, 2009), both are shaded in Figure 1 with different intensity.

5.2 Estimation methodology

We estimate the general specification of the CAW and CTAW model and all restricted and extended specifications listed in Section 4, for orders (p, q) from the list $(0,1), (1,1), (1,2), (2,1), (2,2), (2,3), (3,2), (3,3)$. For each model we compute values of the Bayesian Information Criterion (BIC), as well as minimal and median (across 15 different volatilities and covolatilities) p-values for the Lagrange multiplier test for serial correlation in the residuals with 20 lags corrected for conditional heteroskedasticity (20_{min} and 20_{med}).[8] Recall that the MIDAS versions in contrast to the baseline versions, employ many lags ($mL = 20 \times 12 = 240$). To make comparison of all models on the same footing, we estimate them and compute corresponding values of BIC over the last $T - mL = 1,236 - 240 = 996$ in-sample observations.

In order to guarantee that the number of degrees of freedom, diagonal elements of matrices C, and first elements on main diagonals of matrices A_i and B_i are all positive, we use their square roots as parameters for ML estimation. In a similar way, for MIDAS models we constrain θ to be positive and ω to be greater than 2.

We use a consecutive estimation procedure similar to the bottom-up procedure proposed in Golosnoy, Gribisch, and Liesenfeld (2012). We refine the latter by exploiting multiple starting parameter values in order to exclude dependence on them. We start with CAW(0,1) for which we use diagonal matrices C, A_i and B_i with different sets of diagonal elements as starting parameter values; the model with the highest likelihood is selected in the end. Next we estimate CAW(1,1) for which we try, in addition to diagonal matrices C, A_i and B_i, the optimal solution from CAW(0,1) as well as its value scaled by weights varying from 0.5 to 1.0. Again, the model with the highest likelihood is selected

[8] Golosnoy, Gribisch, and Liesenfeld (2012) instead report p-values for the Ljung–Box test; these values, in contrast to ours, are very close to zero.

in the end. Higher order models are processes in a similar way. For CTAW models the procedure is analogous except that to estimate CTAW(0,1) we additionally try the optimal solution from CAW(0,1) and its scaled versions. Not only does such a strategy explore the regions around optimal solutions of lower order models but also it exploits independent starting parameter values to inspect other areas where the optimal solution can lie.[9]

All standard errors are constructed using the 'sandwich' formula that uses numerically estimated Hessian and Jacobian.

5.3 Estimation results

The diagnostics for all versions and specifications of the CAW and CTAW models are shown in Tables 3–8. All models except baseline ones with orders $(0, 1)$ have no problems with residual serial correlation: even minimal p-values of the LM test exceed 10%, while their median values exceed 40%.[10] The presence of the MIDAS component, however, removes the residual autocorrelation even when the order is (0,1). That is, the order (1,1) is sufficient to take care of serial correlation, but yet higher orders are generally needed to adequately describe the dynamics of realized variances and covariances. According to the BIC, in most specifications the orders (2,1) turn out to be optimal, while some diagonal versions (those that have much fewer parameters overall) require orders (3,2).

Inspection of Tables 4–5 and Tables 7–8 reveals that for the baseline versions of general CTAW models the A-flexible specifications fare better than the AB-flexible ones uniformly in model orders, but this is not the case for diagonal-switching and diagonal models and/or

[9]This proves to be crucial for relatively lower order models because some too parsimonious models like those of orders (0,1) or (1,1) are usually unable to capture all richness of the dynamics thus having their optimal solutions quite far from optimal solutions of higher order models. As a result, the simple bottom-up strategy may lead to suboptimal estimates of the whole category of models. Indeed, estimates that started from diagonal parameter matrices are chosen quite often.

[10]A visual inspection of LM test p-values in 'problematic' models indicates that errors in both variance and covariance equations may be equally subject to residual serial correlation, and that such autocorrelation tends to be present in the same equations for different 'problematic' models. At the same time, given the autoregressive orders, the CTAW structure by itself does not necessarily correct autocorrelation compared to CAW; it is an increase in orders that helps whiten the errors.

MIDAS versions. That is, except for most general (and highly parameterized) models, additional switches in the persistence matrices B_i may improve or may reduce the quality of in-sample fit.

Another interesting observation comes from comparing general and diagonal specifications in all models from each of Tables 3–8. For CAW models, both baseline (Table 3) and MIDAS (Table 6) versions, the diagonal specification is uniformly better than the general specification, which is more unambiguous than the evidence presented in Golosnoy, Gribisch, and Liesenfeld (2012). Somewhat analogously, in cases of both individual asset driven (Table 4) and market driven (Table 5) baseline CTAW models, the BIC favors diagonal-switching specifications over general ones, and diagonal specifications over diagonal-switching ones. That is, the reductions in the likelihood caused by restrictions on diagonals of the matrices A_i, B_i, G_i and H_i are sufficiently small to justify reductions in the degrees of freedom. However, such unambiguity vanishes when it comes to MIDAS versions: while the same tendency remains, some diagonal-switching specifications fare better than corresponding diagonal ones despite a big difference in degrees of parameterizations (see Table 7 in particular).

Table 9 summarizes information about the optimal orders and values of BIC across all model specifications. Among baseline versions, only the heavily parameterized (226 parameters) general AB-flexible CTAW is dominated by the original general CAW model (91 parameters). All baseline diagonal-switching and diagonal CTAW models (some containing more than 100 parameters) are better than the original diagonal CAW model (41 parameters). Among MIDAS versions, the same tendencies are observed, although the dominance is less sharp: not all diagonal MIDAS-CTAW models are better than the original diagonal MIDAS-CAW. Evidently, separating long-run and short-run movements by incorporating the MIDAS structure is more important for CAW than for CTAW in which the mechanism of regime switching itself is able to partially explain some long-run dynamics.[11]

[11]Similarly, one can explain strongly persistent behavior of stationary series using threshold

Interestingly, according to BIC, a best model in each CTAW category is necessarily diagonal with respect to regime switches, i.e. the parameter matrices A_i and B_i in different regimes differ only by their diagonals. That is, extending the degree of parameterization to switching the entire parameter matrices is not worth the improved quality of fit, but extending it to switching their diagonals is.

Tables 10 and 11 contain parameter estimates of the optimal (in-sample, according to BIC) models of both versions (baseline and MIDAS) – diagonal AB-flexible MCTAW(2,1) and diagonal-switching AB-flexible MIDAS-CTAW(2,1) (those that are highlighted in bold in Table 9). Note that both 'winners' are AB-flexible, so that eventually regime switching in the persistence matrices B_i turns out to be important.

In Tables 10 and 11 the statistically significant coefficients at the 1% level are highlighted in bold, those at the 5% level in italic. In the diagonal AB-flexible MCTAW model, the elements of the (diagonal) matrices A_i and B_i are big and highly statistically significant, while the elements in the matrices G_i and H_i that are added to the former are rather heterogenous though relatively small, and most are individually insignificant. In the diagonal-switching AB-flexible MIDAS-CTAW model, the matrices A_i and B_i are all filled by non-zero elements, and the pattern of statistical significance is quite blurry, while the situation with H_i and somewhat with G_i, on the contrary, is clearer. Incorporation of the MIDAS component seems to reduce the degree of persistence, which is to be expected. Interestingly, the estimates of the slope parameter θ and the decay parameter ω for these particular data, although statistically significant, are notably smaller than those for the data of Golosnoy, Gribisch, and Liesenfeld (2012).

One can see in Figures 2 and 3 how the in-sample predictions from the diagonal A-flexible CTAW(3,2) model fit the real realized variance for AXP and realized covariance for AXP-JPM.[12]

autoregressions (Lanne and Saikkonen, 2002).

[12]Recall that the earliest 240 observations are used only in MIDAS models for estimating long-run components, hence no predictions at the beginning of the sample.

5.4 Forecasting methodology

To compare forecasting ability of different models, we construct model confidence sets (MCS) of Hansen, Lunde, and Nason (2011). A model confidence set is a subcollection of models that contains the 'best' (according to a specified loss function or another performance criterion) model with an assigned coverage probability. The MCS procedure also delivers, for all models under consideration, their individual p-values which can be roughly interpreted as their chances to be 'best'.

We use, because of high computational duty, the fixed estimation scheme of generating forecasts when a model is estimated once on the in-sample part, and then one-step-ahead predictions are made using the estimated model and incoming information. After the predictions are made, we compute values of the loss function for the out-of-sample portion of the data. Then we apply the MCS methodology (Hansen, Lunde, and Nason, 2011) using the block bootstrap with length $B = 25$.[13] As performance criteria, we use several statistical loss functions listed below, as well as one economic criterion, the excess return from a mean-variance efficient portfolio. Denote the prediction of Y_t by $\hat{Y}_t$.

One criterion is the Stein loss (James and Stein, 1961):

$$L_S(Y_t, \hat{Y}_t) = \text{tr}[\hat{Y}_t^{-1}\, Y_t] - \log|\hat{Y}_t^{-1}\, Y_t| - n. \tag{12}$$

This is a scale-invariant loss function based on the standardized error. This loss function is asymmetric with respect to over/under-predictions: under-predictions are penalized more heavily than over-predictions. It is implied by the Wishart density, i.e. this prediction criterion is coherent with the estimation step. Moreover, it is robust to the unobservable nature of volatility in the sense of Patton (2011).

Another criterion is the Frobenius loss

$$L_F(Y_t, \hat{Y}_t) = \text{tr}[(Y_t - \hat{Y}_t)'\, (Y_t - \hat{Y}_t)]. \tag{13}$$

[13] We use Kevin Sheppard's MFE Toolbox for MATLAB, please find more details at http://www.kevinsheppard.com/MFE_Toolbox

This is a multivariate extension of the familiar quadratic loss in the scalar case. It is symmetric with respect to over- and under-predictions. It is also robust to the unobservable nature of volatility in the sense of Patton (2011).

In addition we make a comparison of economic significance of predictions of realized covariance matrices. We follow the procedure described in De Goeij and Marquering (2004) where the authors propose to use these predictions in order to construct a mean-variance efficient portfolio. An investor wants to construct a portfolio with the smallest possible variance but with some fixed required expected return. That is, he or she solves the following problem:

$$\begin{aligned} &\min_{w_{t+1}} w_{t+1}' Y_{t+1}^{-1} w_{t+1}, \\ \text{s.t.} \quad & w_{t+1}'\mu + (1 - w_{t+1}'\iota) r_{f,t+1} = \mu_p, \end{aligned}$$

where w_{t+1} are portfolio weights on n assets at moment $t+1$, μ is vector of expected returns for these n assets, ι is vector consisting of n ones, $r_{f,t+1}$ is risk free rate at the moment $t+1$, μ_p is required expected portfolio return. The solution to this problem is given by the following formula:

$$w_{t+1}^* = \frac{(\mu_p - r_{f,t+1}) Y_{t+1}^{-1} (\mu - r_{f,t+1}\iota)}{(\mu - r_{f,t+1}\iota) Y_{t+1}^{-1} (\mu - r_{f,t+1}\iota)}.$$

The authors use constant expected returns for all assets, since it is difficult to predict asset returns and they want to concentrate on volatility forecasts rather than on return forecasts. In case of a highly volatile out-of-sample period it seems that constant expected returns are unlikely as the environment and the market change quickly. Therefore we modify their approach. As predictions for day t we use an average ex-post realized daily return calculated for the next 3 days. On the one hand such approach can be seen as a "perfect foresight" by the investor and give good return predictions; on the other hand, 3 days is a long period, so only incomplete information is given to the investor, and large uncertainty is still present. This approach allows us to concentrate on evaluation of predictions of covariance matrices and not on predictions of returns.

Having one-step-ahead predictions of Y_{t+1} in hand, we calculate the portfolio return as a value weighted return:

$$r_{p,t} = w_t^* r_t,$$

where $r_{p,t}$ is portfolio return at moment t, and r_t is ex-post realized return vector at moment t. Following De Goeij and Marquering (2004) we set the target return μ_p equal to 20%. As a performance measure that captures the trade-off between risk and return of the portfolio the authors use the following average utility function:

$$\hat{U}_p(\gamma) = \frac{1}{T}\sum_t \left[r_{p,t} - \frac{1}{2}\gamma r_{p,t}^2 \right],$$

where γ captures investor's risk aversion. In order to test economic significance of differences in forecasts for different models, we again use MCS, with the loss at moment t being equal to minus utility at moment t. As in De Goeij and Marquering (2004) we use the following values of γ: 3, 6, and 9, but report prediction results only for $\gamma = 6$; the experiments with the other values yield qualitatively similar results.

5.5 Forecasting results

Tables 12 and 13 contain, for the highly volatile and calm subperiods respectively, best performing models according to their predictive abilities. For each specification and loss function separately they show best autoregressive orders, MCS p-values and average loss values. Table 14 shows similar statistics, but for the autoregressive orders that are optimal in the BIC sense (see Table 9).

The conclusions that can be drawn from these tables are less clear-cut than when in-sample performances are compared, and almost every 'rule' has its exceptions. Nevertheless, one can see that for both subperiods and all loss functions, to each benchmark CAW model correspond some CTAW specifications, often several or many, that have equal or more chances to predict better. The same is also true with the CAW and CTAW models selected by BIC.

The optimal autoregressive orders are generally less stable than when BIC is used to select orders, they happen to be all allowed pairs, including (0,1) and (3,3). Comparing different specifications, one can see an expected tendency that more parsimonious models are likely to predict more successfully. However, for some loss functions prediction is better with even higher orders than BIC would prescribe. For instance, in the calm subperiod the orders are quite large with the Frobenius loss but are quite small with the economic loss. Except under the economic loss, the optimal orders tend to be smaller in the calm subperiod than in the highly volatile one.

The BIC-optimal specifications tend to be more successful in prediction the more diagonal their structure is. Recalling that the same tendency is observed in-sample, one should certainly prefer diagonal CTAW specifications in empirical analysis, though not necessarily of low autoregressive orders and not necessarily MIDAS versions. One can hardly say for sure whether A-flexible or AB-flexible specifications predict better (though the former seem a bit more attractive), or whether Market CTAW are better than asset-wise CTAW. Whether baseline or MIDAS versions are preferred highly depends on the loss function: for the Stein loss baseline versions have more chances to fare best, for the economic loss – MIDAS versions.

To see whether it is volatilities or covolatilities that get predicted better as a result of incorporating the leverage effects, we also split the Frobenius loss function into the realized variance and covariance parts separately: $L_V(Y_t, \hat{Y}_t) = (\mathrm{dg}(Y_t - \hat{Y}_t))' \, \mathrm{dg}(Y_t - \hat{Y}_t)$ and $L_C(Y_t, \hat{Y}_t) = \frac{1}{2}(L_F(Y_t, \hat{Y}_t) - L_V(Y_t, \hat{Y}_t))$, where $\mathrm{dg}(P)$ denotes a vector containing the main diagonal of square matrix P, and construct MCS (unreported) for both sublosses as well. It turns out that the models that are best in predicting variances are also among best in predicting covariances, and vice versa. For example, in the highly volatile period the two L_F-best predicting models (which are general CTAW(2,1) and CTAW(2,2)) with L_F-MCS p-values 100% and 81% respectively, are also L_V-best with p-values 100% and

87% and are L_C-best with p-values 100% and 90%. All models with L_F-MCS p-values exceeding 70% (namely, 10 various CTAW and MCTAW) have L_V-MCS p-values of at least 79% and L_C-MCS p-values of at least 63%.

Table 15 provides rankings of what type of models (whose total number is 224, viz. 16 CAW models, 48 CTAW models, 48 MCTAW models, and as many their MIDAS versions) the MCS contain. For each model type (CAW, CTAW, MCTAW), model version (baseline or MIDAS) and each of the three losses (Stein, Frobenius, economic) a corresponding cell counts best models whose p-values belong to one of equal intervals (from [0%,20%] to (80%,100%]). One can see immediately that for relatively small values of significance levels (like 20%) Stein MCS are generally narrower than economic MSC and much narrower than Frobenius MSC.[14] Naturally enough, for the two statistical losses the volume of MCS tends to be larger during the calm subperiod than during the turbulent subperiod, but for the economic criterion the opposite is true.

Again, one can see that the benchmark CAW members never rise above the 60% threshold, except for the MIDAS version with respect to the economic loss. At the same time, there are always a few CTAW or MCTAW models that rise above the 80% bar.

Interestingly, during the volatile subperiod the best predicting models tend to be baseline versions of asset-wise CTAW, while during the calm subperiod – the MIDAS version of Market CTAW. In aggregate terms, incorporating the MIDAS component does good for predictive abilities of the CAW and MCTAW classes but does no good or even worsens performance of the CTAW class. This can be explained by the aforementioned property of switching regimes (whose number for CTAW is 5 as opposed to 2 for MCTAW and 1 for CAW) to implicitly account for some long-run volatility persistence without explicit presence of the MIDAS component.

One can see in Figures 2 and 3 how the forecasts from the out-of-sample L_S-best diagonal A-flexible CTAW(3,2) model correspond to the realizations of the realized

[14] We conjecture that relative tightness of Stein MCS is due to the coherency of the Stein loss to the Wishart likelihood.

variance for AXP and realized covariance for AXP-JPM.

6 Conclusion

In this paper we extended the conditional autoregressive Wishart model from Golosnoy, Gribisch, and Liesenfeld (2012) to incorporate the leverage effect. We considered restricted specifications of the proposed CTAW model and studied the dynamics implied by these models. Next, we estimated the new models along with the benchmark – the original CAW model and its modifications – and evaluated their forecasting abilities using the framework of model confidence sets.

We find strong evidence that some restricted versions of the CTAW model have a better in-sample fit and tend to have a better out-of-sample predictive ability than the benchmark CAW models, and so do most of their MIDAS versions compared to MIDAS versions of the benchmarks. The results indicate that a possibility to account for leverage effect is important for forecasting realized covariance matrices, and that the class of CTAW models is flexible enough to capture this effect. There is hard to tell though which of the restricted models, in particular one with asset-wise regime switches or one with those driven by the market indicator, is universally better in balancing the degree of complexity and quality of in-sample and out-of-sample fit.

Possible venues of future research include a search of other functions of price movements or alternative variables driving regime switches, as well as accommodating smoother transitions from one regime to another. Introducing threshold effects in the long-run components in MIDAS versions may also constitute a useful research direction.

Acknowledgements

Our thanks go to the Editor and two anonymous referees for numerous useful suggestions that significantly improved the paper.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71(2), 579–625.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2011): "Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices with Noise and Non-Synchronous Trading," *Journal of Econometrics*, 162(2), 149–169.

Bauer, G. H., and K. Vorkink (2011): "Forecasting Multivariate Realized Stock Market Volatility," *Journal of Econometrics*, 160(1), 93–101.

Black, F. (1976): "Studies in Stock Price Volatility Changes," *Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economics Statistics Section*, 177–181.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988): "A Capital Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy*, 96(1), 116–131.

Campbell, J. and L. Hentschel (1992): "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," *Journal of Financial Economics*, 31, 281–318.

Cappiello, L., R. F. Engle, and N. Shephard (2006): "Asymmetric dynamics in the correlations of global equity and bond returns," *Journal of Financial Econometrics*, 4, 537–572.

Chiriac, R. and V. Voev (2011): "Modelling and Forecasting Multivariate Realized Volatility," *Journal of Applied Econometrics*, 26(6), 922–947.

Choi, J. and M. Richardson (2008): "The Volatility of a Firms Assets," *AFA 2010 Atlanta Meetings*.

Christie, A. (1982): "The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects," *Journal of Financial Economics*, 10, 407–432.

De Goeij, P. and W. Marquering (2004): "Modeling the Conditional Covariance Between Stock and Bond Returns: A Multivariate GARCH Approach," *Journal of Financial Econometrics*, 2(4), 531–564.

Engle, R. F. (2002): "Dynamic Conditional Correlation," *Journal of Business & Economic Statistics*, 20(3), 339–350.

Engle, R. F. and K. F. Kroner (1995): "Multivariate Simultaneous Generalized Arch," *Econometric Theory*, 11(1), 122–150.

Engle, R. F., E. Ghysels, and B. Sohn (2013): "Stock Market Volatility and Macroeconomic Fundamentals," *Review of Economics and Statistics*, 95(3), 776–797.

French, K., G. Schwert, and R. Stambaugh (1987): "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3–30.

Golosnoy, V., B. Gribisch, and R. Liesenfeld (2012): "The Conditional Autoregressive Wishart Model for Multivariate Stock Market Volatility," *Journal of Econometrics*, 167(1), 211–223.

Gourieroux, C., J. Jasiak, and R. Sufana (2009): "The Wishart Autoregressive Process of Multivariate Stochastic Volatility," *Journal of Econometrics*, 150(2), 167–181.

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

Hautsch, N., L. M. Kyj, and R. C. A. Oomen (2012): "A Blocking and Regularization Approach to High-Dimensional Realized Covariance Estimation," *Journal of Applied Econometrics*, 27(4), 625–645.

HENTSCHEL, L. (1995): "All in the Family Nesting Symmetric and Asymmetric GARCH Models," *Journal of Financial Economics*, 39(1), 71–104.

JAMES, W., AND C. STEIN (1961): "Estimation with Quadratic Loss," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 361–379.

KRONER, K. F. AND V. NG (1998): "Modeling asymmetric comovements of assets returns," *Review of Financial Studies*, 11, 817–844.

LANNE, M. AND P. SAIKKONEN (2002): "Threshold autoregressions for strongly autocorrelated time series," *Journal of Business & Economic Statistics*, 20, 282–289.

LAURENT, S., J. V. K. ROMBOUTS, AND F. VIOLANTE (2012): "On the Forecasting Accuracy of Multivariate GARCH Models," *Journal of Applied Econometrics*, 27(6), 934–955.

LUNDE, A., N. SHEPHARD, AND K. SHEPPARD (2011): "Econometric Analysis of Vast Covariance Matrices Using Composite Realized Kernels," Manuscript, University of Aarhus.

NOURELDIN, D., N. SHEPHARD, AND K. SHEPPARD (2012): "Multivariate High-Frequency-Based Volatility (HEAVY) Models," *Journal of Applied Econometrics*, 27(6), 907–933.

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

VARNESKOV, R. T., AND V. VOEV (2013): "The Role of Realized Ex-post Covariance Measures and Dynamic Model Choice on the Quality of Covariance Forecasts," *Journal of Empirical Finance*, 20, 83-95.

WU, G. (2001): "The Determinants of Asymmetric Volatility," *Review of Financial Studies*, 14, 837–859.

Model	Dynamics for S_t	Specification	Parameter restrictions
CAW	$S_t = CC' + \sum_{i=1}^{p} B_i S_{t-i} B_i' + \sum_{i=1}^{q} A_i Y_{t-i} A_i'$	General	
		Diagonal	A_i and B_i are diagonal
AB-flexible CTAW	$S_t = CC' + \sum_{i=1}^{p} \left(B_i + \sum_{j=1}^{n} H_{i,j} I_{j,t-i}\right) S_{t-i} \left(B_i + \sum_{j=1}^{n} H_{i,j} I_{j,t-i}\right)' + \sum_{i=1}^{q} \left(A_i + \sum_{j=1}^{n} G_{i,j} I_{j,t-i}\right) Y_{t-i} \left(A_i + \sum_{j=1}^{n} G_{i,j} I_{j,t-i}\right)'$	General	
		Diagonal-switching	$H_{i,j}$ and $G_{i,j}$ are zero except $(j,j)^{th}$ entries
		Diagonal	$H_{i,j}$ and $G_{i,j}$ are zero except $(j,j)^{th}$ entries A_i and B_i are diagonal
A-flexible CTAW	$S_t = CC' + \sum_{i=1}^{p} B_i S_{t-i} B_i' + \sum_{i=1}^{q} \left(A_i + \sum_{j=1}^{n} G_{i,j} I_{j,t-i}\right) Y_{t-i} \left(A_i + \sum_{j=1}^{n} G_{i,j} I_{j,t-i}\right)'$	General	
		Diagonal-switching	$G_{i,j}$ are zero except$(j,j)^{th}$ entries
		Diagonal	$G_{i,j}$ are zero except$(j,j)^{th}$ entries A_i and B_i are diagonal
AB-flexible MCTAW	$S_t = CC' + \sum_{i=1}^{p} \left(B_i + H_{i,m} I_{m,t-i}\right) S_{t-i} \left(B_i + H_{i,m} I_{m,t-i}\right)' + \sum_{i=1}^{q} \left(A_i + G_{i,m} I_{m,t-i}\right) Y_{t-i} \left(A_i + G_{i,m} I_{m,t-i}\right)'$	General	
		Diagonal-switching	$H_{i,m}$ and $G_{i,m}$ are diagonal
		Diagonal	A_i, B_i, $H_{i,m}$ and $G_{i,m}$ are diagonal
A-flexible MCTAW	$S_t = CC' + \sum_{i=1}^{p} B_i S_{t-i} B_i' + \sum_{i=1}^{q} \left(A_i + G_{i,m} I_{m,t-i}\right) Y_{t-i} \left(A_i + G_{i,m} I_{m,t-i}\right)'$	General	
		Diagonal-switching	$G_{i,m}$ are diagonal
		Diagonal	A_i, B_i and $G_{i,m}$ are diagonal

Table 1: Classification of the baseline CAW, CTAW and MCTAW models. The MIDAS versions have similar classifications for the short-run volatility dynamics S_t^*.

Summary Statistics						
Stock	Mean	Max.	Min.	Std. dev.	Skewness	Kurtosis
Realized Variances						
AXP	4.42	201.88	0.08	9.15	8.54	130.99
JPM	5.06	176.48	0.11	11.09	7.52	84.69
GE	3.20	114.26	0.10	7.11	7.24	72.65
DP	2,53	63.87	0.16	3.72	6.44	68.05
IBM	1.93	57.54	0.08	3.36	7.32	82.19
Realized Covariances						
AXP-JPM	2.46	64.4	−4.45	5.49	5.60	45.70
AXP-GE	1.76	72.84	−0.41	4.29	7.51	85.11
AXP-DP	1.45	56.52	−0.50	3.28	7.37	85.80
AXP-IBM	1.26	51.82	−0.76	2.98	8.27	97.84
JPM-GE	1.86	70.03	−5.22	4.54	7.56	85.33
JPM-DP	1.52	57.56	−0.88	3.34	7.17	80.12
JPM-IBM	1.33	56.58	−2.21	30.09	8.76	116.58
GE-DP	1.29	49.40	−0.63	2.83	7.55	88.42
GE-IBM	1.13	49.27	−0.78	2.57	8.45	109.61
DP-IBM	0.98	39.88	−0.53	2.22	8.44	102.03
Daily Returns						
AXP	0.00	18.32	−17.02	2.72	0.11	7.90
JPM	0.00	22.77	−22.98	2.96	0.27	14.88
GE	0.00	17.36	−13.89	2.20	0.00	7.78
DP	0.00	10.26	−12.14	1.91	−0.22	8.49
IBM	0.00	10.97	−11.18	1.72	0.13	12.73
SPY	0.00	10.76	−9.60	1.38	−0.07	10.76

Table 2: Summary statistics for realized variances and covariances and daily stock returns. The sample period is from February 2001 to December 2009.

p	q	N_p	BIC	20_{min}	20_{med}
General CAW					
0	1	41	4311	0.00	0.10
1	1	66	431	0.19	0.65
1	2	91	306	0.17	0.59
2	1	91	**232**	0.29	0.62
2	2	116	285	0.30	0.64
2	3	141	398	0.27	0.62
3	2	141	378	0.30	0.68
3	3	166	509	0.30	0.62
Diagonal CAW					
0	1	21	4214	0.00	0.06
1	1	26	228	0.16	0.59
1	2	31	212	0.19	0.57
2	1	31	162	0.17	0.57
2	2	36	181	0.20	0.57
2	3	41	161	0.22	0.57
3	2	41	**123**	0.22	0.62
3	3	46	154	0.23	0.62

Table 3: CAW models. The sample period is from February 2001 to December 2005. Orders of autoregression are in columns p and q, number of parameters is in column N_p, value of BIC is in column BIC, and 20_{min} and 20_{med} denote minimum and median (across 5 variances and 10 covariances) p-values of heteroskedasticity-robust LM test with 20 lags for residual serial correlation. Minimal values of BIC for each specification are highlighted in bold.

p	q	N_p	BIC	20_{min}	20_{med}	N_p	BIC	20_{min}	20_{med}	N_p	BIC	20_{min}	20_{med}
AB-flexible CTAW													
		General				Diagonal-switching				Diagonal			
0	1	86	4467	0.00	0.08	46	4287	0.00	0.05	26	4190	0.00	0.04
1	1	156	429	0.14	0.44	76	171	0.13	0.57	36	−32	0.14	0.53
1	2	226	529	0.16	0.46	106	92	0.15	0.57	46	−13	0.15	0.57
2	1	226	**422**	0.30	0.62	106	**3**	0.23	0.66	46	**−62**	0.14	0.53
2	2	296	728	0.31	0.65	136	83	0.21	0.64	56	−12	0.17	0.57
2	3	366	1108	0.25	0.60	166	229	0.18	0.59	66	−1	0.17	0.56
3	2	366	1041	0.23	0.51	166	174	0.22	0.58	66	−51	0.21	0.52
3	3	436	1427	0.19	0.55	196	334	0.32	0.71	76	2	0.23	0.56
A-flexible CTAW													
		General				Diagonal-switching				Diagonal			
0	1	86	4467	0.00	0.08	46	4287	0.00	0.05	26	4190	0.00	0.04
1	1	111	322	0.24	0.63	71	180	0.19	0.54	31	−24	0.18	0.55
1	2	181	399	0.25	0.56	101	104	0.18	0.61	41	−3	0.17	0.57
2	1	136	**136**	0.28	0.64	96	**−2**	0.26	0.62	36	−79	0.17	0.57
2	2	206	379	0.26	0.62	126	78	0.28	0.63	46	−36	0.23	0.59
2	3	276	739	0.22	0.66	156	219	0.24	0.60	56	−22	0.24	0.58
3	2	231	504	0.30	0.61	151	174	0.30	0.59	51	**−80**	0.25	0.60
3	3	301	865	0.16	0.63	181	317	0.35	0.67	61	−34	0.27	0.57

Table 4: CTAW models. The sample period is from February 2001 to December 2005. Orders of autoregression are in columns p and q, number of parameters is in column N_p, value of BIC is in column BIC, and 20_{min} and 20_{med} denote minimum and median (across 5 variances and 10 covariances) p-values of heteroskedasticity-robust LM test with 20 lags for residual serial correlation. Minimal values of BIC for each specification are highlighted in bold.

p	q	N_p	BIC	20_{min}	20_{med}	N_p	BIC	20_{min}	20_{med}	N_p	BIC	20_{min}	20_{med}
						AB-flexible MCTAW							
			General				Diagonal-switching				Diagonal		
0	1	66	4217	0.00	0.06	46	4152	0.00	0.07	26	4057	0.00	0.01
1	1	116	252	0.23	0.49	76	58	0.20	0.53	36	−157	0.18	0.49
1	2	166	288	0.23	0.57	106	−23	0.16	0.59	46	−140	0.23	0.51
2	1	166	**229**	0.32	0.63	106	**−76**	0.30	0.61	46	**−183**	0.21	0.53
2	2	216	422	0.30	0.71	136	26	0.28	0.67	56	−134	0.22	0.51
2	3	266	670	0.32	0.77	166	153	0.23	0.66	66	−126	0.24	0.56
3	2	266	672	0.33	0.69	166	129	0.32	0.71	66	−178	0.25	0.58
3	3	316	973	0.36	0.71	196	300	0.29	0.71	76	−117	0.26	0.58
						A-flexible MCTAW							
			General				Diagonal-switching				Diagonal		
0	1	66	4217	0.00	0.06	46	4152	0.00	0.07	26	4057	0.00	0.01
1	1	91	250	0.21	0.60	71	132	0.19	0.58	31	−87	0.17	0.53
1	2	141	261	0.17	0.58	101	54	0.17	0.54	41	−68	0.17	0.52
2	1	116	**74**	0.26	0.61	96	**−36**	0.26	0.59	36	−153	0.20	0.53
2	2	166	261	0.23	0.59	126	44	0.27	0.58	46	−106	0.21	0.55
2	3	216	489	0.29	0.63	156	176	0.18	0.62	56	−102	0.21	0.56
3	2	191	330	0.33	0.63	151	130	0.26	0.58	51	**−168**	0.25	0.60
3	3	241	593	0.40	0.57	181	287	0.24	0.66	61	−118	0.24	0.60

Table 5: Market CTAW models. The sample period is from February 2001 to December 2005. Orders of autoregression are in columns p and q, number of parameters is in column N_p, value of BIC is in column BIC, and 20_{min} and 20_{med} denote minimum and median (across 5 variances and 10 covariances) p-values of heteroskedasticity-robust LM test with 20 lags for residual serial correlation. Minimal values of BIC for each specification are highlighted in bold.

p	q	N_{p}	BIC	20_{min}	20_{med}
General MIDAS-CAW					
0	1	43	736	0.10	0.41
1	1	68	267	0.35	0.59
1	2	93	163	0.23	0.65
2	1	93	**10**	0.22	0.58
2	2	118	24	0.27	0.69
2	3	143	89	0.26	0.70
3	2	143	70	0.31	0.74
3	3	168	205	0.28	0.72
Diagonal MIDAS-CAW					
0	1	23	668	0.06	0.36
1	1	28	−42	0.27	0.67
1	2	33	−99	0.24	0.62
2	1	33	−106	0.28	0.67
2	2	38	−125	0.32	0.67
2	3	43	−152	0.31	0.68
3	2	43	**−171**	0.35	0.75
3	3	48	−150	0.36	0.75

Table 6: MIDAS-CAW models. The sample period is from February 2001 to December 2005. Orders of autoregression are in columns p and q, number of parameters is in column N_{p}, value of BIC is in column BIC, and 20_{min} and 20_{med} denote minimum and median (across 5 variances and 10 covariances) p-values of heteroskedasticity-robust LM test with 20 lags for residual serial correlation. Minimal values of BIC for each specification are highlighted in bold.

p	q	N_p	BIC	20_{min}	20_{med}	N_p	BIC	20_{min}	20_{med}	N_p	BIC	20_{min}	20_{med}
AB-flexible MIDAS-CTAW													
		General				Diagonal-switching				Diagonal			
0	1	88	950	0.17	0.35	48	754	0.15	0.41	28	686	0.09	0.38
1	1	158	333	0.18	0.60	78	108	0.23	0.65	38	−17	0.28	0.72
1	2	228	338	0.14	0.70	108	−77	0.17	0.72	48	−60	0.26	0.71
2	1	228	**190**	0.01	0.27	108	**−230**	0.18	0.51	48	−75	0.26	0.71
2	2	298	393	0.06	0.44	138	−209	0.18	0.46	58	−56	0.31	0.70
2	3	368	657	0.10	0.42	168	−114	0.07	0.43	68	−55	0.31	0.71
3	2	368	649	0.05	0.55	168	−122	0.11	0.56	68	**−103**	0.31	0.70
3	3	438	996	0.14	0.40	198	12	0.16	0.57	78	−51	0.33	0.68
A-flexible MIDAS-CTAW													
		General				Diagonal-switching				Diagonal			
0	1	88	950	0.17	0.35	48	754	0.15	0.41	28	686	0.09	0.38
1	1	113	311	0.21	0.51	73	137	0.25	0.67	33	−23	0.27	0.70
1	2	183	353	0.11	0.57	103	19	0.24	0.73	43	−63	0.26	0.66
2	1	138	**147**	0.17	0.50	98	**−20**	0.18	0.54	38	−90	0.28	0.72
2	2	208	341	0.18	0.41	128	−1	0.24	0.59	48	−88	0.28	0.67
2	3	278	616	0.13	0.49	158	79	0.23	0.57	58	−86	0.35	0.72
3	2	233	399	0.15	0.45	153	56	0.25	0.66	53	**−136**	0.31	0.73
3	3	303	734	0.16	0.44	183	217	0.29	0.59	63	−86	0.29	0.72

Table 7: MIDAS-CTAW models. The sample period is from February 2001 to December 2005. Orders of autoregression are in columns p and q, number of parameters is in column N_p, value of BIC is in column BIC, and 20_{min} and 20_{med} denote minimum and median (across 5 variances and 10 covariances) p-values of heteroskedasticity-robust LM test with 20 lags for residual serial correlation. Minimal values of BIC for each specification are highlighted in bold.

p	q	N_p	BIC	20_{min}	20_{med}	N_p	BIC	20_{min}	20_{med}	N_p	BIC	20_{min}	20_{med}
AB-flexible MIDAS-MCTAW													
		General				Diagonal-switching				Diagonal			
0	1	68	857	0.19	0.38	48	766	0.09	0.40	28	662	0.06	0.35
1	1	118	214	0.07	0.34	78	156	0.22	0.61	38	−11	0.24	0.60
1	2	168	189	0.21	0.44	108	54	0.29	0.62	48	−41	0.29	0.63
2	1	168	**149**	0.17	0.49	108	**−94**	0.21	0.48	48	−112	0.17	0.45
2	2	218	327	0.23	0.52	138	−46	0.22	0.49	58	−94	0.18	0.51
2	3	268	525	0.24	0.56	168	−4	0.18	0.50	68	−120	0.17	0.47
3	2	268	515	0.19	0.47	168	−24	0.20	0.54	68	**−187**	0.19	0.48
3	3	318	793	0.27	0.55	198	139	0.19	0.56	78	−141	0.20	0.50
A-flexible MIDAS-MCTAW													
		General				Diagonal-switching				Diagonal			
0	1	68	857	0.19	0.38	48	766	0.09	0.41	28	662	0.06	0.35
1	1	93	217	0.15	0.52	73	151	0.24	0.59	33	−25	0.24	0.63
1	2	143	227	0.11	0.47	103	49	0.32	0.61	43	−56	0.24	0.61
2	1	118	100	0.13	0.43	98	**−70**	0.25	0.67	38	−91	0.26	0.68
2	2	168	**82**	0.22	0.61	128	−57	0.25	0.60	48	−93	0.26	0.61
2	3	218	383	0.20	0.59	158	127	0.25	0.59	58	−101	0.23	0.61
3	2	193	211	0.24	0.60	153	82	0.25	0.59	53	**−144**	0.33	0.63
3	3	243	551	0.24	0.60	183	289	0.25	0.59	63	−97	0.32	0.63

Table 8: MIDAS-MCTAW models. The sample period is from February 2001 to December 2005. Orders of autoregression are in columns p and q, number of parameters is in column N_p, value of BIC is in column BIC, and 20_{min} and 20_{med} denote minimum and median (across 5 variances and 10 covariances) p-values of heteroskedasticity-robust LM test with 20 lags for residual serial correlation. Minimal values of BIC for each specification are highlighted in bold.

Model specification	p	q	N_{p}	BIC
		Baseline versions		
General CAW	2	1	91	232
Diagonal CAW	3	2	41	123
General AB-flexible CTAW	2	1	226	422
Diagonal-switching AB-flexible CTAW	2	1	106	3
Diagonal AB-flexible CTAW	2	1	46	−62
General A-flexible CTAW	2	1	136	136
Diagonal-switching A-flexible CTAW	2	1	96	−2
Diagonal A-flexible CTAW	3	2	51	−80
General AB-flexible MCTAW	2	1	166	229
Diagonal-switching AB-flexible MCTAW	2	1	106	−76
Diagonal AB-flexible MCTAW	2	1	46	**−183**
General A-flexible MCTAW	2	1	116	74
Diagonal-switching A-flexible MCTAW	2	1	96	−36
Diagonal A-flexible MCTAW	3	2	51	−168
		MIDAS versions		
General CAW	2	1	93	10
Diagonal CAW	3	2	43	−171
General AB-flexible CTAW	2	1	228	190
Diagonal-switching AB-flexible CTAW	2	1	108	**−230**
Diagonal AB-flexible CTAW	3	2	68	−103
General A-flexible CTAW	2	1	138	147
Diagonal-switching A-flexible CTAW	2	1	98	−20
Diagonal A-flexible CTAW	3	2	53	−136
General AB-flexible MCTAW	2	1	168	149
Diagonal-switching AB-flexible MCTAW	2	1	108	−94
Diagonal AB-flexible MCTAW	3	2	68	−188
General A-flexible MCTAW	2	2	168	82
Diagonal-switching A-flexible MCTAW	2	1	98	−70
Diagonal A-flexible MCTAW	3	2	53	−144

Table 9: Best CAW, CTAW and MCTAW models, according to BIC. The sample period is from February 2001 to December 2005. Orders of autoregression are in columns p and q, number of parameters is in column N_{p}, value of BIC is in column BIC. Minimal values of BIC for baseline and MIDAS versions are highlighted in bold.

	Diagonal AB-flexible MCTAW(2,1)				
C	**0.159**	0	0	0	0
	0.085	**0.187**	0	0	0
	0.065	**0.054**	**0.132**	0	0
	0.080	**0.052**	*0.041*	**0.185**	0
	0.069	**0.064**	**0.048**	**0.032**	**0.155**
B_1	**0.674**	**0.603**	**0.706**	**0.664**	**0.690**
B_2	*0.437*	**0.478**	**0.464**	**0.502**	**0.411**
A_1	**0.520**	**0.556**	**0.438**	**0.448**	**0.497**
$H_{1,m}$	*0.117*	0.095	0.054	*0.074*	*0.103*
$H_{2,m}$	−0.010	0.041	0.029	0.021	0.002
$G_{1,m}$	−0.017	−0.003	0.053	0.030	0.006
K	**20.8**				

Table 10: Estimation results for diagonal AB-flexible MCTAW(2,1) model. Rows B_j, A_j, $H_{j,m}$ and $G_{j,m}$ present diagonal elements of matrices B_j, A_j, $H_{j,m}$ and $G_{j,m}$, respectively. Sample period is from February 2001 to December 2005. Statistically significant at 1% level coefficients are highlighted in bold, at 5% in italic, standard errors are constructed with 'sandwich' formula using numerically estimated Hessian and Jacobian.

Diagonal-switching AB-flexible MIDAS-CTAW(2,1)											
C	**0.372**	0	0	0	0	B_1	**0.694**	*−0.244*	0.170	0.105	0.083
	0.106	**0.360**	0	0	0		0.033	*0.396*	−0.092	**0.392**	0.084
	0.086	*0.089*	**0.291**	0	0		−0.061	0.109	**0.603**	**−0.289**	−0.110
	0.105	**0.093**	0.077	**0.420**	0		−0.028	−0.072	0.073	*0.258*	*0.315*
	0.105	**0.121**	**0.194**	0.075	*0.212*		−0.028	0.076	*0.187*	*−0.265*	**0.394**
A_1	**0.443**	*0.085*	0.016	0.012	−0.011	B_2	0.000	−0.064	0.079	−0.031	0.058
	0.012	**0.484**	−0.059	−0.026	−0.030		0.220	−0.179	**0.418**	−0.031	0.203
	−0.013	0.009	**0.333**	0.066	0.074		0.066	0.023	*0.321*	*0.188*	−0.145
	−0.049	0.004	0.021	**0.440**	−0.016		−0.025	0.057	−0.065	**0.474**	**−0.220**
	−0.034	0.014	−0.047	0.006	**0.466**		0.091	−0.059	**−0.330**	0.163	**0.471**
H_1	**0.233**	*0.314*	*0.134*	**0.370**	0.062	G_1	0.006	−0.029	*0.129*	−0.051	*0.129*
H_2	0.112	−0.066	*0.194*	*0.224*	*0.135*						
K	**21.3**										
$\sqrt{\theta}$	**0.151**										
ω	*4.58*										

Table 11: Estimation results for diagonal-switching AB-flexible MIDAS-CTAW(2,1) model. Rows H_j and G_j present diagonal elements $H_{j,i}$ and $G_{j,i}$, $i = 1, \ldots, 5$, respectively, of corresponding matrices. Sample period is from February 2001 to December 2005. Statistically significant at 1% level coefficients are highlighted in bold, at 5% in italic, standard errors are constructed with 'sandwich' formula using numerically estimated Hessian and Jacobian.

Model specification	$(p,q)_S$	p_S	$\bar{L}_S$	$(p,q)_F$	p_F	$\bar{L}_F$	$(p,q)_e$	p_e	$\bar{L}_e$
Baseline versions									
General CAW	(1,2)	24.8%	1.290	(3,3)	44.0%	2332	(1,1)	54.8%	0.0291
Diagonal CAW	(3,3)	24.8%	1.260	(3,3)	49.0%	2306	(1,1)	54.8%	0.0307
General AB-flexible CTAW	(1,2)	24.8%	1.291	(2,1)	100.0%	2160	(0,1)	96.0%	0.0293
Diagonal-switching AB-flexible CTAW	(1,1)	24.8%	1.280	(3,2)	59.9%	2228	(1,1)	96.0%	0.0290
Diagonal AB-flexible CTAW	(3,2)	89.4%	1.248	(3,2)	59.9%	2265	(1,1)	54.8%	0.0307
General A-flexible CTAW	(1,1)	24.8%	1.292	(3,2)	71.0%	2204	(1,1)	100.0%	0.0277
Diagonal-switching A-flexible CTAW	(1,1)	24.8%	1.279	(3,3)	59.9%	2235	(1,1)	54.8%	0.0292
Diagonal A-flexible CTAW	(3,2)	100.0%	1.247	(3,2)	49.0%	2275	(1,1)	54.8%	0.0310
General AB-flexible MCTAW	(1,1)	1.6%	1.331	(3,2)	71.0%	2191	(1,1)	54.8%	0.0310
Diagonal-switching AB-flexible MCTAW	(2,3)	10.5%	1.300	(3,3)	59.9%	2216	(0,1)	54.8%	0.0294
Diagonal AB-flexible MCTAW	(3,2)	24.8%	1.295	(3,2)	59.9%	2258	(0,1)	54.8%	0.0308
General A-flexible MCTAW	(1,1)	5.4%	1.313	(3,2)	59.9%	2239	(1,1)	54.8%	0.0290
Diagonal-switching A-flexible MCTAW	(1,2)	24.8%	1.298	(3,2)	59.9%	2243	(1,1)	96.0%	0.0287
Diagonal A-flexible MCTAW	(3,3)	24.8%	1.287	(3,2)	49.0%	2279	(1,1)	54.8%	0.0305
MIDAS versions									
General CAW	(3,2)	40.4%	1.258	(1,1)	44.0%	2359	(0,1)	54.8%	0.0291
Diagonal CAW	(3,2)	40.4%	1.257	(3,2)	44.0%	2371	(0,1)	96.0%	0.0289
General AB-flexible CTAW	(3,2)	1.5%	1.423	(0,1)	49.0%	2348	(1,1)	54.8%	0.0296
Diagonal-switching AB-flexible CTAW	(3,3)	24.8%	1.329	(1,1)	44.0%	2425	(0,1)	54.8%	0.0293
Diagonal AB-flexible CTAW	(1,2)	24.8%	1.263	(3,3)	44.0%	2338	(0,1)	54.8%	0.0290
General A-flexible CTAW	(1,2)	24.8%	1.284	(0,1)	54.9%	2348	(1,1)	54.8%	0.0308
Diagonal-switching A-flexible CTAW	(2,2)	24.8%	1.265	(1,1)	44.0%	2373	(0,1)	54.8%	0.0293
Diagonal A-flexible CTAW	(3,2)	24.8%	1.260	(3,2)	44.0%	2332	(0,1)	96.0%	0.0290
General AB-flexible MCTAW	(3,2)	1.6%	1.335	(0,1)	44.0%	2399	(3,3)	54.8%	0.0318
Diagonal-switching AB-flexible MCTAW	(1,2)	10.5%	1.297	(1,1)	44.0%	2385	(3,3)	96.0%	0.0284
Diagonal AB-flexible MCTAW	(1,2)	24.8%	1.282	(1,2)	49.0%	2315	(0,1)	54.8%	0.0293
General A-flexible MCTAW	(1,2)	24.8%	1.270	(3,3)	44.0%	2337	(2,2)	54.8%	0.0317
Diagonal-switching A-flexible MCTAW	(1,2)	24.8%	1.287	(2,1)	44.0%	2331	(2,1)	96.0%	0.0291
Diagonal A-flexible MCTAW	(2,2)	24.8%	1.268	(3,2)	44.0%	2391	(0,1)	54.8%	0.0293

Table 12: Best CAW, CTAW and MCTAW models, according to out-of-sample criteria, accompanied by MCS p-values and mean loss values during highly volatile period from August, 29 2008 to June 9, 2009. Optimal autoregressive orders for Stein, Frobenius and economic losses are in columns (p,q); corresponding MCS p-values are in columns p_S, p_F and p_e; corresponding average losses are in columns $\bar{L}_S$, $\bar{L}_F$ and $\bar{L}_e$.

Model specification	$(p,q)_S$	p_S	$\bar{L}_S$	$(p,q)_F$	p_F	$\bar{L}_F$	$(p,q)_e$	p_e	$\bar{L}_e$
Baseline versions									
General CAW	(1,2)	2.1%	0.972	(1,2)	51.6%	37.05	(1,2)	44.9%	−0.0188
Diagonal CAW	(3,3)	2.1%	0.967	(3,2)	51.6%	36.99	(2,1)	44.9%	−0.0188
General AB-flexible CTAW	(1,1)	0.0%	0.990	(1,1)	51.6%	37.14	(1,1)	44.9%	−0.0186
Diagonal-switching AB-flexible CTAW	(1,2)	0.6%	0.975	(1,1)	70.2%	36.02	(1,1)	44.9%	−0.0186
Diagonal AB-flexible CTAW	(2,2)	2.1%	0.969	(3,2)	51.6%	36.20	(2,1)	44.9%	−0.0184
General A-flexible CTAW	(1,1)	0.6%	0.977	(1,1)	81.2%	35.22	(2,1)	44.9%	−0.0187
Diagonal-switching A-flexible CTAW	(1,2)	1.7%	0.973	(1,2)	51.6%	36.42	(2,2)	44.9%	−0.0190
Diagonal A-flexible CTAW	(3,2)	2.1%	0.964	(3,2)	70.2%	36.02	(2,1)	44.9%	−0.0187
General AB-flexible MCTAW	(1,1)	2.1%	0.966	(1,2)	51.6%	37.40	(1,1)	44.9%	−0.0188
Diagonal-switching AB-flexible MCTAW	(1,2)	45.7%	0.960	(1,1)	70.2%	35.91	(1,1)	44.9%	−0.0188
Diagonal AB-flexible MCTAW	(3,2)	53.6%	0.951	(3,2)	70.2%	35.79	(1,1)	44.9%	−0.0186
General A-flexible MCTAW	(1,1)	2.1%	0.962	(2,1)	51.6%	36.71	(2,1)	44.9%	−0.0191
Diagonal-switching A-flexible MCTAW	(1,2)	45.7%	0.958	(1,2)	51.6%	36.20	(2,1)	44.9%	−0.0192
Diagonal A-flexible MCTAW	(3,2)	53.6%	0.948	(3,2)	70.2%	35.77	(1,1)	44.9%	−0.0189
MIDAS versions									
General CAW	(2,3)	45.7%	0.956	(3,2)	51.6%	37.24	(2,3)	99.8%	−0.0200
Diagonal CAW	(2,3)	45.7%	0.954	(3,2)	51.6%	37.11	(2,3)	99.8%	−0.0200
General AB-flexible CTAW	(0,1)	0.0%	1.027	(2,2)	51.6%	38.98	(2,2)	44.9%	−0.0186
Diagonal-switching AB-flexible CTAW	(1,1)	0.6%	0.987	(1,2)	51.6%	36.34	(2,2)	44.9%	−0.0191
Diagonal AB-flexible CTAW	(1,2)	2.1%	0.962	(2,1)	51.6%	36.35	(1,2)	44.9%	−0.0196
General A-flexible CTAW	(2,1)	0.1%	0.986	(1,1)	37.0%	40.33	(1,2)	44.9%	−0.0194
Diagonal-switching A-flexible CTAW	(1,2)	2.1%	0.968	(0,1)	51.6%	39.07	(3,2)	44.9%	−0.0195
Diagonal A-flexible CTAW	(1,2)	45.7%	0.959	(2,1)	51.6%	37.20	(3,2)	99.8%	−0.0200
General AB-flexible MCTAW	(2,1)	1.7%	0.977	(2,2)	81.2%	35.10	(2,1)	44.9%	−0.0190
Diagonal-switching AB-flexible MCTAW	(3,2)	45.7%	0.962	(3,2)	100.0%	34.66	(2,1)	44.9%	−0.0193
Diagonal AB-flexible MCTAW	(1,2)	2.1%	0.967	(3,2)	81.2%	35.35	(2,2)	97.9%	−0.0198
General A-flexible MCTAW	(3,2)	45.7%	0.951	(3,3)	51.6%	36.93	(1,1)	98.0%	−0.0198
Diagonal-switching A-flexible MCTAW	(2,2)	100.0%	0.942	(2,2)	51.6%	37.47	(3,2)	44.9%	−0.0193
Diagonal A-flexible MCTAW	(1,2)	45.7%	0.960	(2,2)	51.6%	37.78	(2,3)	100.0%	−0.0201

Table 13: Best CAW, CTAW and MCTAW models, according to out-of-sample criteria, accompanied by MCS p-values and mean loss values during calm period from January 4, 2006 to August 28, 2009. Optimal autoregressive orders for Stein, Frobenius and economic losses are in columns (p,q); corresponding MCS p-values are in columns p_S, p_F and p_e; corresponding average losses are in columns $\bar{L}_S$, $\bar{L}_F$ and $\bar{L}_e$.

Model specification	$(p,q)_{BIC}$	Volatile subperiod						Calm subperiod					
		p_S	$\bar{L}_S$	p_F	$\bar{L}_F$	p_e	$\bar{L}_e$	p_S	$\bar{L}_S$	p_F	$\bar{L}_F$	p_e	$\bar{L}_e$
Baseline versions													
General CAW	(2,1)	0.2%	1.363	44.0%	2315	15.8%	0.0345	0.1%	0.986	51.6%	37.67	44.9%	−0.0188
Diagonal CAW	(3,2)	24.8%	1.260	44.0%	2333	54.8%	0.0313	2.1%	0.967	51.6%	36.99	44.9%	−0.0184
General AB-flexible CTAW	(2,1)	0.2%	1.362	100.0%	2160	7.1%	0.0385	0.0%	1.037	33.4%	41.09	3.8%	−0.0166
Diagonal-switching AB-flexible CTAW	(2,1)	1.6%	1.337	54.9%	2251	15.8%	0.0347	0.0%	0.988	51.7%	37.22	28.0%	−0.0181
Diagonal AB-flexible CTAW	(2,1)	24.8%	1.261	49.0%	2296	54.8%	0.0308	2.1%	0.970	51.6%	36.27	44.9%	−0.0184
General A-flexible CTAW	(2,1)	10.5%	1.327	71.0%	2218	15.8%	0.0345	0.0%	0.995	51.6%	36.76	44.9%	−0.0187
Diagonal-switching A-flexible CTAW	(2,1)	10.5%	1.318	59.9%	2258	7.1%	0.0345	0.6%	0.979	51.6%	36.82	44.9%	−0.0188
Diagonal A-flexible CTAW	(3,2)	100.0%	1.247	49.0%	2275	15.8%	0.0321	2.1%	0.964	70.2%	36.02	44.9%	−0.0183
General AB-flexible MCTAW	(2,1)	0.2%	1.413	59.9%	2232	7.1%	0.0367	0.6%	0.987	51.6%	38.48	28.0%	−0.0181
Diagonal-switching AB-flexible MCTAW	(2,1)	0.2%	1.373	59.9%	2231	7.1%	0.0340	2.1%	0.974	51.6%	37.17	28.0%	−0.0185
Diagonal AB-flexible MCTAW	(2,1)	24.8%	1.305	49.0%	2268	15.8%	0.0319	45.7%	0.952	59.6%	35.84	44.9%	−0.0185
General A-flexible MCTAW	(2,1)	0.2%	1.375	59.9%	2269	7.1%	0.0350	2.1%	0.969	51.6%	36.71	44.9%	−0.0191
Diagonal-switching A-flexible MCTAW	(2,1)	0.2%	1.370	49.0%	2279	15.8%	0.0340	2.1%	0.962	51.6%	36.31	44.9%	−0.0192
Diagonal A-flexible MCTAW	(3,2)	24.8%	1.288	49.0%	2279	54.8%	0.0316	53.6%	0.948	70.2%	35.78	44.9%	−0.0183
MIDAS versions													
General CAW	(2,1)	24.8%	1.285	44.0%	2434	54.8%	0.0303	2.1%	0.966	51.6%	38.37	44.9%	−0.0191
Diagonal CAW	(3,2)	40.4%	1.257	44.0%	2371	15.8%	0.0319	45.7%	0.955	51.6%	37.11	97.9%	−0.0199
General AB-flexible CTAW	(2,1)	0.2%	1.496	44.0%	5346	15.8%	0.0363	0.0%	1.059	33.4%	43.39	44.9%	−0.0179
Diagonal-switching AB-flexible CTAW	(2,1)	10.5%	1.346	44.0%	3191	15.8%	0.0334	0.0%	0.998	51.6%	37.76	44.9%	−0.0186
Diagonal AB-flexible CTAW	(3,2)	24.8%	1.265	44.0%	2339	54.8%	0.0306	2.1%	0.969	51.6%	36.59	44.9%	−0.0193
General A-flexible CTAW	(2,1)	1.6%	1.329	44.0%	2511	15.8%	0.0334	0.1%	0.986	33.4%	40.23	44.9%	−0.0190
Diagonal-switching A-flexible CTAW	(2,1)	24.8%	1.296	44.0%	2450	54.8%	0.0301	2.1%	0.974	37.0%	38.47	44.9%	−0.0189
Diagonal A-flexible CTAW	(3,2)	24.8%	1.260	44.0%	2332	15.8%	0.0309	45.7%	0.960	51.6%	37.59	99.8%	−0.0200
General AB-flexible MCTAW	(2,1)	0.2%	1.413	44.0%	2662	54.8%	0.0319	1.7%	0.977	51.6%	35.89	44.9%	−0.0190
Diagonal-switching AB-flexible MCTAW	(2,1)	0.2%	1.383	44.0%	2790	54.8%	0.0319	2.1%	0.972	74.5%	35.58	44.9%	−0.0193
Diagonal AB-flexible MCTAW	(3,2)	1.6%	1.360	49.0%	2379	54.8%	0.0310	0.6%	0.979	81.2%	35.35	44.9%	−0.0195
General A-flexible MCTAW	(2,2)	10.5%	1.300	44.0%	2339	54.8%	0.0317	45.7%	0.951	51.6%	37.02	44.9%	−0.0191
Diagonal-switching A-flexible MCTAW	(2,1)	10.5%	1.296	44.0%	2331	96.0%	0.0291	45.7%	0.952	51.6%	37.68	44.9%	−0.0192
Diagonal A-flexible MCTAW	(3,2)	24.8%	1.270	44.0%	2391	15.8%	0.0329	2.1%	0.965	51.6%	37.87	99.8%	−0.0201

Table 14: Out-of-sample MCS p-values and median loss values for in-sample best model orders of all specifications during highly volatile period. Out-of-sample period is from January 2006 to December 2009. Autoregressive orders are in column $(p,q)_{BIC}$; MCS p-values for Stein, Frobenius and economic losses are in columns p_S, p_F and p_e; corresponding average losses are in columns $\bar{L}_S$, $\bar{L}_F$ and $\bar{L}_e$.

Loss	MCS interval		Volatile subperiod						Calm subperiod					
	from	to	CAW	CTAW	MCTAW	MIDAS -CAW	MIDAS -CTAW	MIDAS -MCTAW	CAW	CTAW	MCTAW	MIDAS -CAW	MIDAS -CTAW	MIDAS -MCTAW
Stein	80%	100%	0	3	0	0	0	0	0	0	0	0	0	1
	60%	80%	0	0	0	0	0	0	0	0	0	0	0	0
	40%	60%	0	3	0	3	0	0	0	0	16	11	3	10
	20%	40%	8	17	14	10	24	12	0	0	1	0	0	0
	0%	20%	8	25	34	3	24	36	16	48	31	5	45	37
Frobenius	80%	100%	0	2	0	0	0	0	0	2	0	0	0	5
	60%	80%	0	7	1	0	0	0	0	3	5	0	0	4
	40%	60%	16	39	47	16	48	48	12	31	38	15	22	32
	20%	40%	0	0	0	0	0	0	4	12	5	1	26	7
	0%	20%	0	0	0	0	0	0	0	0	0	0	0	0
Economic	80%	100%	0	4	1	1	1	2	0	0	0	7	3	8
	60%	80%	0	0	0	0	0	0	0	0	0	0	0	0
	40%	60%	7	16	19	12	32	27	13	28	34	9	44	40
	20%	40%	0	0	0	0	0	0	2	9	11	0	1	0
	0%	20%	9	28	28	3	15	19	1	11	3	0	0	0

Table 15: Summary of contents of model confidence sets. Out-of-sample period is from January 2006 to December 2009. Figures signify number of models of each category (with varying orders and specifications) whose p-values fall into designates intervals. There are 16 CAW models, 48 CTAW models, 48 MCTAW models, and as many their MIDAS versions, all totalling to 224 models.

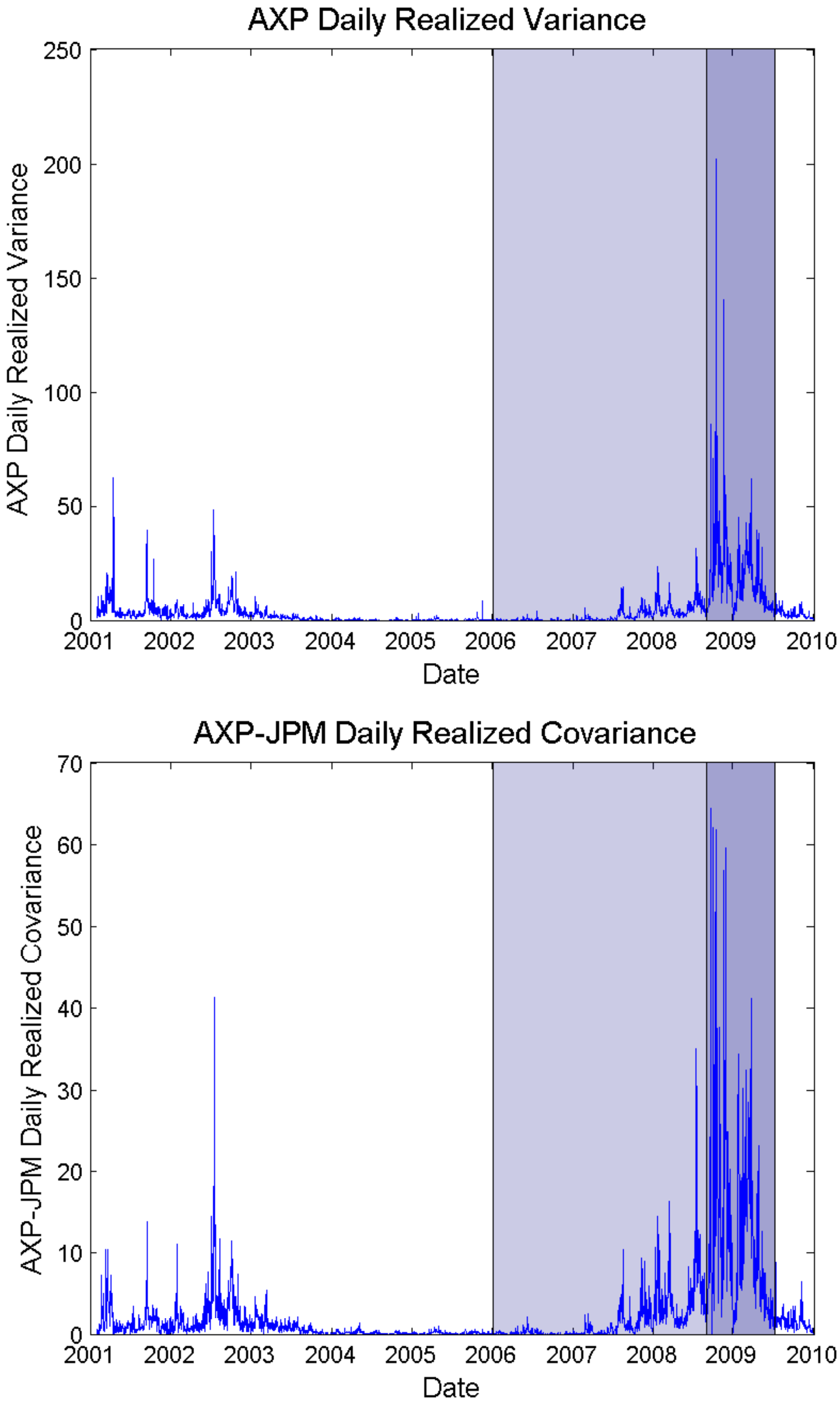

Figure 1: Realized variance for AXP and realized covariance for AXP-JPM. Sample period is from February 2001 to December 2009. The out-of-sample part of the sample is shaded.

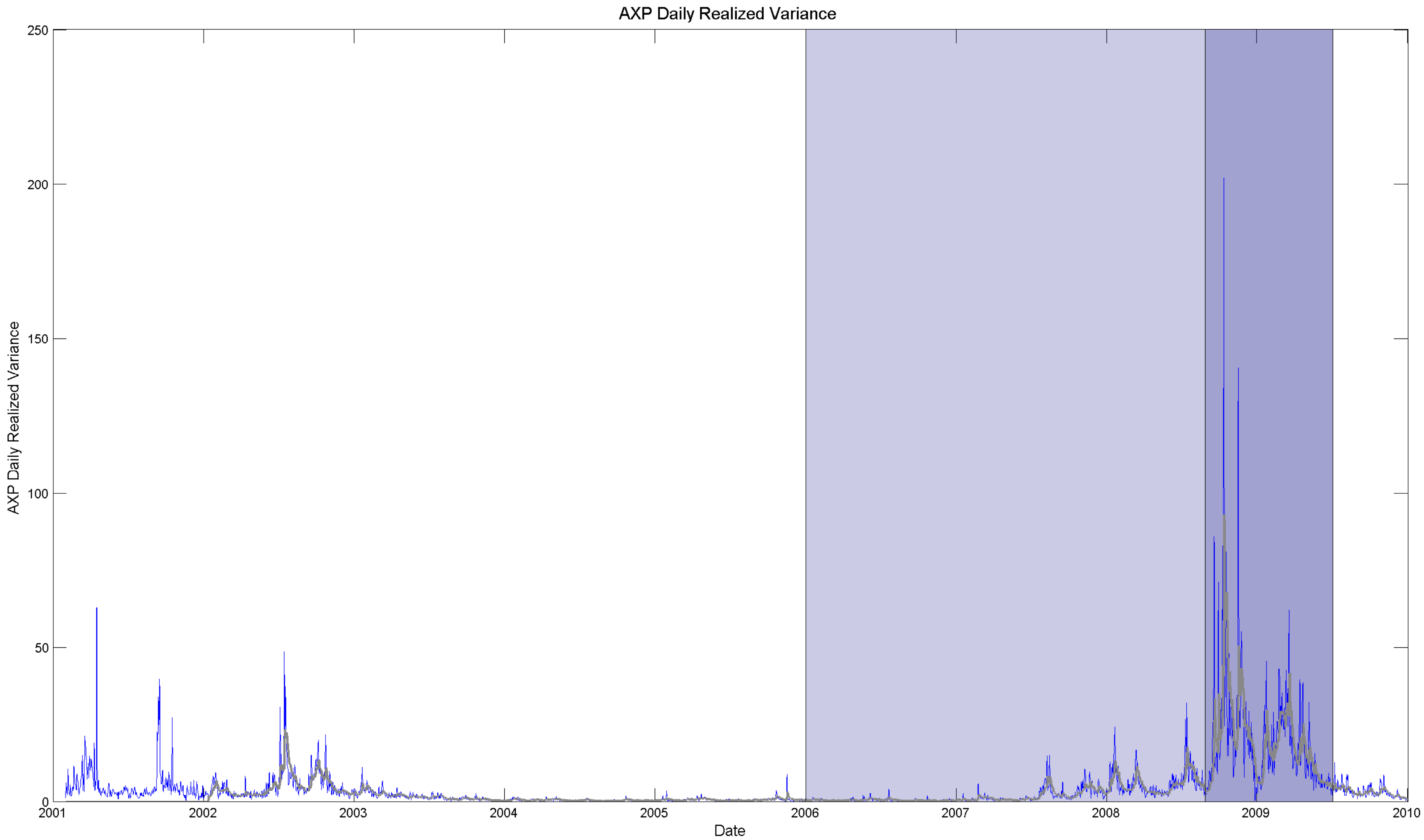

Figure 2: Realized variance for AXP and its in-sample and out-of-sample predictions from the diagonal A-flexible CTAW(3,2) model. Sample period is from February 2001 to December 2009. The out-of-sample part of the sample is shaded.

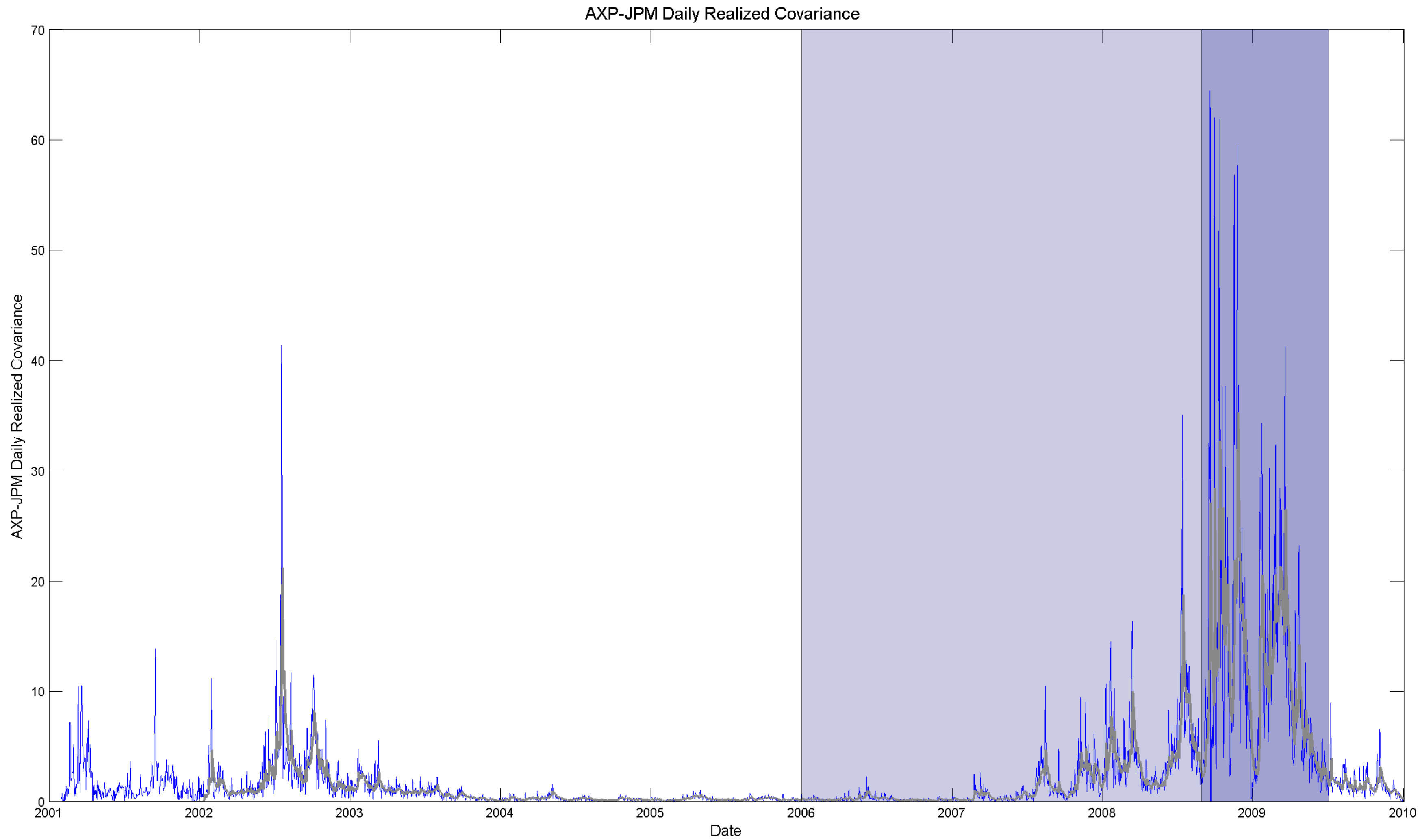

Figure 3: Realized covariance for AXP-JPM and its in-sample and out-of-sample predictions from the diagonal A-flexible CTAW(3,2) model. Sample period is from February 2001 to December 2009. The out-of-sample part of the sample is shaded.