Exchange Rate Volatility Forecasting: a Multivariate Realized-GARCH Approach

Janine Balter[*] Elena-Ivona Dumitrescu[†] Peter Reinhard Hansen[‡]

01/14/15

Abstract

We propose a model of exchange rates that jointly models associated realized measures of volatility and covariances within the Realized GARCH framework. The proposed model exploits identities arising from no arbitrage conditions, that facilitates a relatively parsimonious modeling of a panel of exchange rates. The model shares the simplicity of GARCH models while taking advantage of realized volatility measures that are computed from high-frequency (intraday) data. The latter leads to a better modeling of the variances and covariances, by providing a flexible modeling of their dynamic properties. The model easily produce forecasts at any horizon. The model is illustrated with an empirical application for exchange rates between the currencies: EUR, USD and JPY. An out-of-sample comparison shows that the proposed model dominates conventional benchmark models, in particular at shorter horizons.

Keywords: Exchange rates volatility, Forecasting, Realized GARCH

1 Introduction

This paper gauges whether the availability of high-frequency financial data enable the construction of more accurate forecasting models for the conditional covariance of daily exchange-rate returns.

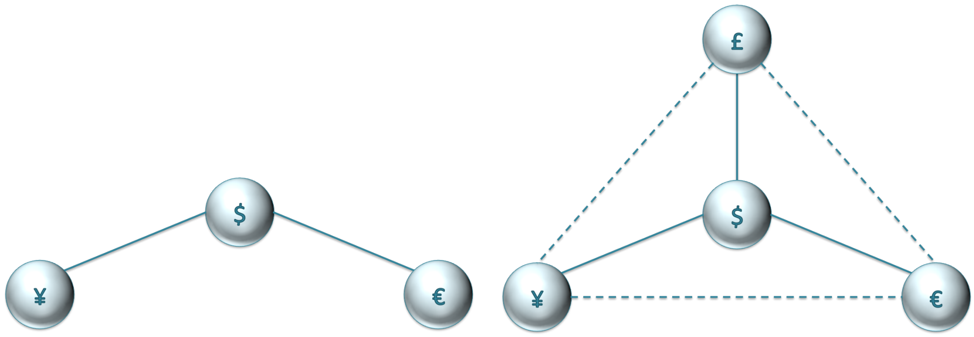
[*]Saarland University
[†]Paris West University Nanterre la Défense
[‡]European University Institute and CREATES

Relatively accurate estimators of ex-post volatility and covariances of daily returns, i.e. realized measures, can be computed from high-frequency financial data, including the realized covariance, realized kernels, Markov chain estimators, and other related measures (see Barndorff-Nielsen et al., 2011; Hansen and Horel, 2009 and references therein). According to Hansen and Lunde (2010), there are two main ways to exploit such realized measures in variance forecasting, labeled as reduced-form and model-based. If the former type of forecast assumes a formal time-series model for the series of realized measures, the latter specifies a parametric model for the return distribution. Model-based forecasting hence relies on a GARCH-type structure where the realized measures are included as exogenous variables. An appealing class of model-based approaches is the Realized GARCH framework of Hansen et al. (2012a) where the dynamics of realized measures and daily returns is jointly modeled. These models have been shown to lead to a significant improvement in the empirical fit over traditional GARCH models. Contrary to GARCH-X models, that are 'incomplete' in the sense that they are unable to provide information about volatility level beyond one period into the future, realized models are 'complete' and easy to use for multi-period forecasting.

In this paper we hence build upon the Realized GARCH framework and propose a multivariate model that improves the modeling and forecasting of exchange rates volatility and covariances. The model is specifically designed for the Forex market in the sense that it exploits the identities following from no-arbitrage conditions. Triangular arbitrage opportunities arise if the price quotes for currency pairs with correlating currencies do not match up. However, this type of arbitrage opportunity arises quite seldom and lasts only a few seconds. It follows that at a daily level the no-arbitrage hypothesis is valid and the cross-rates are equalized. To illustrate this, consider the triangular case of the EUR, USD, and JPY currencies. By absence of triangular arbitrage, the continuously compounded return on the €/¥ cross rate must be equal to the sum between the \$/¥ and €/\$ returns. Formally, the system in the left pane of figure 1 is characterized by the identity: $r_{EU} + r_{UJ} = r_{EJ}$, where $r_{(.)}$ denotes the return series for each exchange rate (EUR/USD; USD/JPY; EUR/JPY). This has two key consequences: i) the cross rate can be inferred from the two other sequences of returns and ii) we can infer the variance of the cross rate as $v_{EJ} = v_{EU} + v_{UJ} + 2cov(r_{EU}, r_{UJ})$, where $v(.)$ denotes the variance and $cov(.)$ the covariance of the exchange-rates (see Andersen et al., 2003a for a discussion on this topic). For a larger system, e.g. EUR, USD, JPY and GBP, the number of identities is given by the number of triangles the system can be split into (see the right pane in figure 1 for a visual identification of the three triangles).

Figure 1: Triangular arbitrage

(a) EUR / USD / JPY (b) EUR / USD / JPY / GBP



There are mainly two ways to account for these identities within the Realized GARCH framework. First, one can specify the full model, for all the countries in the panel, while respecting the no-arbitrage constraints. Second, one can estimate the minimum sufficient model, i.e. the model that includes the minimum number of exchanges rates to define the interconnections within the system, and infer the quantities associated with the other exchange rates. In our example, graphically this comes down to relating the 3 (4) currencies through any 2 (3) exchange-rates respectively. The more currencies, the more possibilities among which to select the exchange-rates to be modeled from the ones to be inferred. Generalizing to a panel of N countries, it is clear that the minimum model is relatively more parsimonious than the full one. It includes $N-1$ exchange rates instead of N(N-1)/2. Similarly, only $N(N-1)/2$ covariance elements ought to be estimated for the minimum model instead of roughly $N^4/8$ in the full model.[1] This is why we choose the minimum model as the multivariate Realized GARCH specification for exchange rates.

This multivariate realized model has appealing properties relatively to traditional multivariate GARCH models. It takes advantage of realized measures of variance, that have lower noise-to-signal ratio than the daily squared return traditionally used in GARCH models to form expectations about the future level of variance. The model is hence expected to react rapidly and perform well in situations where volatility and correlations are subject to abrupt changes. It has a dynamic covariance structure that is revised at every period through the measurement equations relating the conditional measures of volatility and correlation with the corresponding realized mea-

[1] $\frac{1}{2}(\frac{N(N-1)}{2}+1)\frac{N(N-1)}{2} \simeq \frac{N^4}{8}$

sures. It provides a flexible modeling of the dynamic properties of variances and covariances and it is simple to estimate by quasi-maximum likelihood. Most importantly, this model easily produces forecasts at any horizon.

The model is also related with the work of Hansen et al. (2012b), who propose a hierarchical bivariate model for stock returns' conditional beta in a Realized GARCH framework. It hence exploits the information from the correlation between the market and each individuals stock market in a parsimonious way. Another complete model for stock returns is the multivariate HEAVY by Noureldin et al. (2012) which operates with multiple latent volatility processes.

In an empirical application for the EUR, USD and JPY currencies we find that the multivariate realized model outperforms the traditional multivariate GARCH benchmarks, i.e. cDCC (Aielli, 2013), BEKK (Engle and Kroner, 1995), CCC (Engel and Hamilton, 1990) as well as the univariate Realized EGARCH model (Hansen et al., 2012a) up to 10 periods ahead. This is the case not only for the covariance-matrix forecasting but also for each of its components taken independently. What is more surprising is that the bivariate realized model dominates its competitors in the case of the inferred exchange-rate. This result emphasizes the proposed model's good forecasting abilities, since the bivariate realized model does not directly model the data for the cross rate, whereas the two competing models, i.e. GARCH and Realized EGARCH, use the cross rate historical series of returns and realized measure of volatility.

The paper proceeds as follows. Section 2 introduces the multivariate Realized GARCH model. We also discuss the estimation of the model and the multi-step forecasting procedure. Section 3 contains an empirical application on the EUR / USD / JPY currencies. Finally, section 4 concludes.

2 Methodology

In this section we introduce the minimum multivariate realized model for exchange rates that encompasses information offered by the realized measures of volatility. We discuss first the general formulation of the model. Then we present the estimation method and subsequently we show how it can be used for forecasting.

2.1 The Minimum Realized Model

To formally introduce the model, let $r_t = (r_{0,t}, ..., r_{N-1,t})'$ denote the $N - 1 \times 1$ vector of low-frequency, typically daily, returns associated with the exchange rates directly included in the model. The series of corresponding high-frequency derived realized measures of variance are given by $x_t = (x_{0,t}, ..., x_{N-1,t})'$. Let us further denote by $y_t = y_{ij,t}$, for $i \neq j, i > j$, $i, j \in \{1, N-1\}$, the vector of realized measures of daily correlation. These observable variables lead to the natural filtration $\mathcal{F}_{t-1} = \sigma(r_{t-1}, x_{t-1}, y_{t-1}, r_{t-2}, x_{t-2}, y_{t-2}, ...)$ that includes both past realized measures and low-frequency returns.

The multivariate Realized GARCH model can be structured in three sets of equations: the ones for the returns, the multivariate GARCH system and the measurement equations. In vector form, the equation for the returns is given by

$$r_t = \mu + \sqrt{h_t} z_t \tag{1}$$

where μ is a $N-1 \times 1$ vector of intercepts, $h_t = (h_{1,t}, ..., h_{N-1,t})'$ represents the vector of conditional variances and $z_t = (z_{1,t}, ..., z_{N-1,t})$' is the vector of $iid\, \mathcal{N}_{N-1}(0,1)$-distributed studentized residuals.

Let us measure the dependence between the exchange-rate returns by the conditional correlation, ρ_t. From the equations in (1) it can easily be seen that each element $\rho_{ij,t}$ in the $(N-1)(N-2)/2 \times 1$ vector of correlations ρ_t corresponds also to the conditional covariance of the studentized returns, i.e. $\rho_{ij,t} = cov\,(z_{i,t}, z_{j,t}|\mathcal{F}_{t-1})$, for $i \neq j$, $i > j$ and $i, j \in \{1, N-1\}$. Let us further denote by $F : (-1, 1) \to \mathbb{R}$ the Fisher transformation $F(\rho) = \frac{1}{2}\log\left((1+\rho)/(1-\rho)\right)$, that maps the correlations $\rho_{ij,t}$ from the bounded interval $(-1, 1)$ to $\mathbb{R}$.

Next, we specify the multivariate GARCH system. We jointly model the dynamics of the log-conditional variances as well as the dynamics of the Fisher transformed conditional correlation as a VARMA(1,1) system:

$$V_t := \boldsymbol{\alpha} + \boldsymbol{\beta} V_{t-1} + \boldsymbol{\tau} z_t + \boldsymbol{\tau}^* z_t^* + \boldsymbol{\gamma} U_t \tag{2}$$

with

$$V_t = \begin{pmatrix} \log(h_t) \\ F(\rho_t) \end{pmatrix}, z_t = \begin{pmatrix} z_{1,t-1} \\ ... \\ z_{N-1,t-1} \end{pmatrix}, z_t^* = \begin{pmatrix} z_{1,t-1}^2 - 1 \\ ... \\ z_{N-1,t-1}^2 - 1 \end{pmatrix}, \mathrm{U}_t = \begin{pmatrix} u_{1,t-1} \\ ... \\ u_{N-1,t-1} \\ v_{t-1} \end{pmatrix}$$

and where α is a $N(N-1)/2 \times 1$ vector of intercepts, β and γ are squared $N(N-1)/2 \times N(N-1)/2$ matrices, while τ and τ^* are $N(N-1)/2 \times N-1$ matrices.

To complete the specification of our model, we further introduce the measurement equations. They define the dynamics of the realized measures of variance $x_t = (x_{1,t}, ..., x_{N-1,t})$ and correlation $y_t = \{y_{ij,t}\}$, $i \neq j$, $i > j$, $i, j \in \{1, N-1\}$ as:

$$\log(x_t) = \xi + \varphi \log(h_t) + \delta(z_t) + \delta^*(z_t^*) + u_t \tag{3}$$

$$F(y_t) = \zeta + \psi F(\rho_t) + v_t \tag{4}$$

where $u_t = (u_{1,t}, ..., u_{N-1,t})'$ ξ and ζ are $N-1 \times 1$ vectors of intercepts and φ, δ, δ^* and ψ are $N-1 \times N-1$ diagonal matrices.

The conditional variances are modeled via EGARCH equations augmented with information on the correlations between the exchange-rate returns. Hence the model preserves the ARMA structure of the traditional GARCH models while the logarithmic transformation ensures that the variance is always positive definite. Note also that the log-linear specification can be motivated by the well-known fact that the realized measures of variance are closely connected to the squared returns and hence to the conditional variance $\log(r_t - \mu)^2 = \log(h_t) + \log(z_t^2)$ (see Hansen et al., 2012a and Hansen et al., 2012b). The logarithmic measurement equations further make the logarithmic GARCH specification convenient.

The quadratic terms

$$\boldsymbol{\tau}(z) = \boldsymbol{\tau}(z_{t-1}) + \boldsymbol{\tau}^*(z_{t-1}^*) \text{ and } \boldsymbol{\delta}(z) = \boldsymbol{\delta}(z_{t-1}) + \boldsymbol{\delta}^*(z_{t-1}^*)$$

specify the leverage functions in the GARCH and measurement equations, respectively. Hansen et al. (2012a) find that a simple quadratic form appropriately models the dependence between returns and future volatility. Note that this term allows for asymmetric leverage effects, i.e. that volatility can react differently to positive and negative shocks to returns.

We assume that the vector of measurement errors $U_{t-1} = (u_{1,t-1}, ..., u_{N-1,t-1}, v_{t-1})' \sim \mathcal{N}_{N(N-1)/2}(0, \Omega)$, i.e. it follows a multivariate normal distribution with zero-mean and covariance matrix Ω:

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \cdot & \Omega_{22} \end{pmatrix}$$

where the sub-matrices Ω_{11}, Ω_{22}, Ω_{12} are defined as $\Omega_{11} = \begin{pmatrix} \sigma_{u_1}^2 & \dots & \sigma_{u_1}\sigma_{u_{N-1}} \\ \cdot & \sigma_{u_i}^2 & \sigma_{u_i}\sigma_{u_j} \\ \cdot & \cdot & \sigma_{u_{N-1}}^2 \end{pmatrix}$,

$$\Omega_{22} = \begin{pmatrix} \sigma_{v1}^2 & \dots & \sigma_{v_1}\sigma_{v_{(N-1)(N-2)/2}} \\ \cdot & \sigma_{v_i}^2 & \sigma_{v_i}\sigma_{v_j} \\ \cdot & \cdot & \sigma_{v_{(N-1)(N-2)/2}}^2 \end{pmatrix} \text{ and } \Omega_{12} = \begin{pmatrix} \sigma_{u1}\sigma_{v_1} & \dots & \sigma_{u_1}\sigma_{v_{(N-1)(N-2)/2}} \\ \cdot & \sigma_{u_i}\sigma_{v_j} & \sigma_{u_i}\sigma_{v_k} \\ \cdot & \cdot & \sigma_{u_{N-1}}\sigma_{v_{(N-1)(N-2)/2}} \end{pmatrix}.$$

Our choice relies on findings in Andersen et al. (2003b) inter alios who show that the realized variance is approximately log-normally distributed. Further, we assume that the error terms from the measurement equations, U_t, are independent from the studentized residuals z_t, what simplifies the estimation procedure. Due to the presence of the leverage functions $\boldsymbol{\tau}(z)$ and $\boldsymbol{\delta}(z)$, this independence is realistic in practice.

The coefficients on the main diagonal of the matrix $\boldsymbol{\beta}$ reveal the persistent, autoregressive behavior of the conditional variance and correlation, respectively. The system is stationary if they are less than one. The off-diagonal parameters may contain information on the direct variance spillover from one exchange rate to another.

Besides, the realized measures are connected with their conditional counterparts through the $\boldsymbol{\gamma}U_t$ term of errors defined by the measurement equations. The realized measures can hence drive the future dynamics of the conditional measures up and down. Furthermore, to make the specification more parsimonious, in the empirical application we assume that the coefficients φ in the measurement equations are equal to 1 (see Hansen and Huang, 2012 for the univariate case). This is a reasonable assumption, since the round-the clock realized measures of variance x_t are expected to be roughly proportional to the daily conditional volatilities h_t.

2.2 Estimation

The multivariate Realized GARCH model is easily estimated by Quasi Maximum-Likelihood. For this, we proceed in several steps.

Recall that our model consists of two sets of observable variables, the returns r_t and the realized measures x_t and y_t. Their joint density conditional on the existing past information $\mathcal{F}_{t-1}$ is given by

$$f(r_t, x_t, y_t|\mathcal{F}_{t-1}). \tag{5}$$

which can be decomposed into the product of the conditional and marginal densities as:

$$\underbrace{f(r_t|\mathcal{F}_{t-1})}_{I} \times \underbrace{f(x_t, y_t|r_t, \mathcal{F}_{t-1})}_{II}.$$

For the first term (I), we exploit the fact that the returns are assumed normally distributed with mean μ and covariance matrix H_t with elements $\sigma_{ij,t} = \begin{cases} h_{i,t}, & \text{for } i = j \\ \rho_{ij,t}\sqrt{h_{i,t}}\sqrt{h_{j,t}}, & for\, i \neq j \end{cases}$. The contribution of the first term to the quasi log-likelihood function is hence given by

$$lg_{I'} := -\frac{1}{2}\sum_{t=1}^{T}\Big((N-1)\log(2\pi) + log|H_t| + (r_t-\mu)^{'}H_t^{-1}(r_t-\mu)\Big).$$

For the second term (II), we can exploit the fact that the studentized residuals z_t are independent from the innovations in the measurement equations U_t. Besides, the vector of innovations U_t is normally distributed and thus are the associated realized measures (x_t, y_t). We can hence write

$$f(x_t, y_t|r_t, \mathcal{F}_{t-1}) = f(x_t, y_t|\mathcal{F}_{t-1}).$$

The second part of the quasi-log-likelihood function is

$$lg_{II} := -\frac{1}{2}\sum_{t=1}^{T}[(N-1)\log(2\pi) + \log(\det\Omega) + U_t'\Omega^{-1}U_t],$$

where Ω is the covariance matrix of the innovations.

The number of free parameters in the likelihood maximization process can be reduced by concentrating the likelihood function with respect to the variance-covariance matrix Ω. For this purpose, let us define the estimators of the covariance parameters:

$$\hat{\sigma}^2_{u_i} := \frac{1}{T}\sum_{t=1}^{T}\hat{u}^2_{i,t}, \quad \hat{\sigma}^2_{v_1} := \frac{1}{T}\sum_{t=1}^{T}\hat{v}^2_{1,t}, \quad \hat{\sigma}^2_{u_i,u_j} := \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{i,t}\hat{u}_{j,t},$$
$$\hat{\sigma}^2_{u_i,v_j} := \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{i,t}\hat{v}_{j,t}, \quad \hat{\sigma}^2_{v_i,v_j} := \frac{1}{T}\sum_{t=1}^{T}\hat{v}_{i,t}\hat{v}_{j,t},$$
$$\text{where } \hat{\Omega} := \frac{1}{T}\sum_{t=1}^{T}\hat{U}_t\hat{U}_t, \quad \hat{U}_1 := (\hat{u}_t, \hat{v}_t)'.$$

So finally,

$$lg_{II} := -\frac{1}{2}T[(N-1)\log(2\pi) + \log(\det\Omega) + \frac{N(N-1)}{2}].$$

The Quasi Log-Likelihood function is the sum of the two components presented above

$$l(\theta) = lg_I(\theta) + lg_{II}(\theta),$$

with $\theta = (vec(\alpha), diag(\beta), vec(\tau), vec(\tau^*)$,$vec(\gamma)$,$vec(\xi)$,$vec(\varphi)$,$vec(\delta)$, $vec(\delta^*)$, $vec(\zeta)$,$vec(\psi))$. Note that the initial values $h_{.,1}$ and $\rho_{.,1}$ are treated as unknown parameters.

2.3 Partial Log-Likelihood for Returns

To facilitate the comparison of the fitted returns outputted by the bivariate Realized GARCH model with those from a standard GARCH model or from a univariate EGARCH model, we consider as a measure of fit the partial log-likelihood associated with each series of returns

$$l(r_i, \theta) = -\frac{1}{2}\sum_{t=1}^{T}[log(2\pi) + log(h_{i,t}) + (r_{i,t} - \mu)^2/h_{i,t}],$$

where $r_{i,t}$ and $h_{i,t}$ denote the sequence of returns and conditional variance associated with each exchange rate for $i \in \{1, ..., N-1\}$.

2.4 Multi-period Ahead Forecasting

Our multivariate Realized GARCH model exploits the dynamics of the log-variances and the Fisher-transformed correlations for a panel of exchange-rates. The model is complete in the sense that it also relates the dynamics of the realized measures of variance and correlation to that of the conditional measures.

One advantage of this complete specification is that multi-period ahead forecasts are feasible. Indeed, log-variance forecasts are easy to compute based on the VARMAX(1,1) representation of this system. But we are most often interested in forecasting the variance itself and not the log-variance. A similar reasoning applies to the Fisher-transformed conditional correlation. In particular, forecasts of the full covariance matrix are of special interest for portfolio selection, risk management and asset pricing.

Let us partition $V_t = \begin{bmatrix} V_{1,t} \\ V_{2,t} \end{bmatrix}$ where $V_{1,t}$ is set to include the log-conditional variances and $V_{2,t}$ corresponds to the Fisher-transformed conditional correlations submatrix. On the one hand, the non-linear (logarithmic) transformation applied to the conditional variance sequence leads to the inequality between $E(\exp(V_{1,t}))$ and $\exp E(V_{1,t})$, where the first term is the one we are interested in. This feature is actually common to all realized models based on logarithmic specifications, including Hansen et al. (2012a)'s Realized GARCH and Hansen et al. (2012b)'s Realized Beta-GARCH. One way to tackle this problem was proposed by Hansen et al. (2012a) and it is based on bootstrapping or Monte-Carlo simulations. On the other hand, the presence of the Fisher-transformed conditional correlation is specific to multivariate realized models. Given the employment of the non-linear Fisher transform, the same forecasting issue arises as in the case of the conditional variance.

We hence propose joint bootstrap-based forecasts of the conditional variances and correlations. For this, we bootstrap from the residual set $\varepsilon_t = (z_t, U_t)'$, with $t \in \{1, T\}$, where T is the size of the in-sample dataset. We make use of the 'factor' structure of the studentized residuals to account for the time-varying dependence among exchange-rate returns.

k-step-ahead forecasts of the distribution of V_{T+k} conditional on the information set $\mathcal{F}_t$ can be obtained by iterating on equation (2) where z_t, z_t^* and U_t are replaced by the bootstrapped residuals. The variance forecasts $h_{i,T+k}$ are then given by the average of the exponential of all the possible log-variance forecasts obtained in S simulations $h_{i,T+k} = \frac{1}{S}\sum_{s=1}^{S} \exp(V_{1,T+k}(i))$. In the same vein, correlation forecasts are given by $\rho_{i,T+k} = \frac{1}{S}\sum_{s=1}^{S} F^{-1}(V_{2,T+k}(i))$.

3 Empirical Analysis

We implement the multivariate Realized GARCH model in the case of the EUR, USD, JYP currencies and discuss its implications in terms of covariance forecasts.

3.1 The Dataset

Intra-daily 1-minute data have been obtained from Dukascopy for the period August 7, 2003 to August 29, 2014 for all the exchange rates. We divide the sample into an in-sample period, August 7, 2003 - January 6, 2010, and an out-of-sample period spanning the period from January 7, 2010 to August 29, 2014.

Since trading is slowish over weekends, we follow the usual practice in this literature and remove the weekend quotes from Friday 21:00 GMT to Sunday 21:00 GMT. Many public holidays (e.g. days around Christmas, New Year, President's Day, Good Friday, Easter Monday, Memorial Day, US National Day, Labor Day, Thanksgiving, Japan's Emperor Day) are removed for the same reasons.

The daily 9PM to 9PM returns are given by $r_{i,t} = \ln p_{i,t} - \ln p_{i,t-1}$, where $i \in \{1, 2\}$ designates the exchange rate (EUR/USD and USD/JPY respectively), and $p_{i,t}$ represents the last price observed on day t before 9PM. We also make use of realized measures of volatility and correlation computed using 1-minute data. We adopt the multivariate realized kernel by Barndorff-Nielsen et al. (2011), which is robust to noise and insures a positive definite realized covariance matrix.

3.2 Preliminary Analysis

This analysis sheds light on a simplified specification of the multivariate Realized GARCH model. We then show that the multivariate Realized GARCH framework is superior to standard GARCH models in terms of partial log-likelihood both in-sample and out-of-sample. A summary of the estimation results and residual tests is further provided.

3.2.1 Simplifications in the Specification of the Bivariate Realized GARCH Model

Table 1 presents full and partial likelihood results for three specifications of our model: the unconstrained bivariate model introduced in section 2.1 and two restricted versions. The first constrained model assumes that β is a diagonal matrix, i.e. there are no direct spillovers between the elements of the exchange-rate covariance matrix. The second restricted model gauges the stylized fact that exchange rates do not exhibit an asymmetry effect. This specification hence assumes that the covariance matrix does not react differently to positive / negative innovations, i.e. τ is set to 0. The out-of-sample likelihood is obtained simply by plugging the in-sample estimates into the out-of-sample log-likelihood function, whereas the partial likelihood is computed as described in section 2.3. The inference we draw from this table is that the diagonal-β model is generally a good model.

Note that we report the likelihood values and not likelihood ratio statistics for two reasons. First, in the QMLE framework, the limit distribution of the likelihood-ratio statistic is generally a weighted sum of χ^2-distributed random variables and the usual χ^2 critical values are only indicative of significance. Second, the asymptotic distribution of the out-of-sample likelihood-ratio statistic is non-standard (see Hansen et al., 2012a for a more thorough discussion on this topic). For example, comparing the likelihood-ratio statistic $LR = -2(l_{bivRG,diag\beta} - l_{bivRG}) = 11.06$ with a $\chi^2(6)$ suggests that leaving the off-diagonal β parameters unrestricted does not improve the in-sample fit substantially and hence the restricted model is preferable.

In view of the out-of-sample full likelihood results, the model supports a simplification to the diagonal-β specification. Besides, the out-of-sample partial log-likelihoods suggest that asymmetry could play a certain role in the case of the USD/JPY exchange rate. Although the partial likelihood is not the object of maximization in the realized framework, this empirical evidence may justify that the bivariate model does not support a further simplification.

3.2.2 A Partial Log-likelihood Model Comparison

Table 2 reports in-sample and out-of-sample partial log-likelihoods that allow us to compare the empirical fit of the diagonal-β bivariate realized model and that of univariate EGARCH and traditional GARCH models. We find that the realized models dominate the GARCH one not only in-sample but also out-of-sample. This result is very impressive because the realized models do not seek to maximize the partial likelihood, which is the aim of the GARCH model. Another very impressive finding is that the bivariate model exhibits the largest partial likelihood for the inferred (not directly modeled) EUR/JPY exchange-rate, although the other models use the actual observed series of returns and realized-variance in the estimation. [2]These results emphasize the good fit of the bivariate Realized GARCH model relatively to its competitors.

3.2.3 Estimation Results

The results for the diagonal-β Realized GARCH model with EUR/USD and USD/JPY returns are

$$\begin{aligned} r_{0,t} &= \sqrt{h_{0,t}} z_{0,t} \\ r_{1,t} &= \sqrt{h_{1,t}} z_{1,t} \end{aligned}$$

$$\begin{pmatrix} \log(h_{0,t}) \\ \log(h_{1,t}) \\ F(\rho_{01,t}) \end{pmatrix} := \hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\beta}} \begin{pmatrix} \log(h_{0,t-1}) \\ \log(h_{1,t-1}) \\ F(\rho_{01,t-1}) \end{pmatrix} + \hat{\boldsymbol{\tau}} \begin{pmatrix} z_{0,t-1} \\ z_{1,t-1} \end{pmatrix} + \hat{\boldsymbol{\tau}}^* \begin{pmatrix} z_{0,t-1}^2 - 1 \\ z_{1,t-1}^2 - 1 \end{pmatrix} + \hat{\boldsymbol{\gamma}} \begin{pmatrix} u_{0,t-1} \\ u_{1,t-1} \\ v_{1,t-1} \end{pmatrix}$$

with

$$\hat{\boldsymbol{\alpha}} := (\underset{(0.013)}{-0.012}, \underset{(0.017)}{-0.034}, \underset{(0.004)}{-0.019})', \hat{\boldsymbol{\beta}} := \begin{pmatrix} \underset{(0.027)}{0.989} & 0 & 0 \\ 0 & \underset{(0.062)}{0.963} & 0 \\ 0 & 0 & \underset{(0.028)}{0.959} \end{pmatrix}, \hat{\boldsymbol{\tau}} := \begin{pmatrix} \underset{(0.014)}{-0.005} & \underset{(0.008)}{-0.009} \\ \underset{(0.014)}{-0.013} & \underset{(0.014)}{-0.071} \\ \underset{(0.003)}{-0.032} & \underset{(0.002)}{-0.033} \end{pmatrix},$$

[2]The no-arbitrage condition implies that the returns (variance) for the EUR/YEN series are a linear combination of the returns (variances, respectively) of the two exchange rates directly modeled: $r_{2,t} = r_{1,t} + r_{0,t}$ and $h_{2,t} = h_{1,t} + h_{0,t} + 2\rho_{01}\sqrt{h_{0,t}h_{1,t}}$.

$$\hat{\boldsymbol{\tau}}^* := \begin{pmatrix} \underset{(0.011)}{0.028} & \underset{(0.008)}{0.007} \\ \underset{(0.010)}{0.012} & \underset{(0.006)}{0.043} \\ \underset{(0.002)}{-0.002} & \underset{(0.007)}{0.005} \end{pmatrix} \text{ and } \hat{\boldsymbol{\gamma}} := \begin{pmatrix} \underset{(0.010)}{0.126} & \underset{(0.001)}{0.001} & \underset{(0.016)}{0.071} \\ \underset{(0.023)}{-0.014} & \underset{(0.012)}{0.197} & \underset{0.018}{0.109} \\ \underset{(0.012)}{0.046} & \underset{(0.018)}{0.027} & \underset{(0.011)}{0.326} \end{pmatrix}$$

and measurement equations

$$\begin{aligned} \log(x_{0,t}) &= \underset{(0.019)}{-0.195} + \log(h_{0,t}) - \underset{(0.022)}{0.016}(z_{0,t}) + \underset{(0.003)}{0.092}(z_{0,t}^2 - 1) + u_{0,t} \\ \log(x_{1,t}) &= \underset{(0.022)}{-0.163} + \log(h_{1,t}) - \underset{(0.011)}{0.032}(z_{1,t}) + \underset{(0.004)}{0.103}(z_{1,t}^2 - 1) + u_{1,t} \\ F(y_{01,t}) &= \underset{(0.022)}{0.019} + \underset{(0.056)}{1.031}F(\rho_{01,t}) + v_{0,t}, \end{aligned}$$

where the numbers in parentheses represent robust standard errors based on the sandwich method.

The log-volatilities and Fisher-transformed correlation are highly persistent given the very close to 1 estimates on the diagonal of the $\hat{\beta}$ matrix. The diagonal $\hat{\gamma}$ estimates are generally significant and particularly large relatively to the 0.05 traditional value of the ARCH parameter in a GARCH model. This emphasizes the good quality of the information contained in the realized measures of volatility about future volatility. The spillovers between the exchange rates are hence indirect and occur through the residuals of the measurement equations. The significant off-diagonal estimates of the $\hat{\gamma}$ matrix emphasize the role of the realized correlation dynamics for future volatility and that of the realized measures of variance for future correlation.

Traditionally, exchange rates do not exhibit leverage effects. However, in the case of the USD/JPY variance the leverage function seems to matter ($\hat{\tau}$ is negative, whereas $\hat{\tau}^*$ is positive). This goes along the lines of the previous result on out-of-sample partial likelihoods. The Fisher transformed conditional correlation also reacts differently to positive and negative news about returns. Furthermore, leverage plays an important role in the measurement equations in the sense that it insures independence between studentized innovations and the innovations in the measurement equations.

3.2.4 Residuals Checks

To check the statistical adequacy of the model, we compute standard model diagnostics. The normality assumption for the five standardized residuals $\hat{z}_0$, $\hat{z}_1$, $\hat{u}_0$, $\hat{u}_1$ and $\hat{v}$ is assessed in the framework of QQ-plots. Figure 2 compares the empirical distribution of these residuals to a normal distribution. The rejection of the normality hypothesis for $\hat{z}_0$ and $\hat{z}_1$ supports the use of bootstrap in the forecasting exercise.

We further implement the Ljung-Box test of the null hypothesis of no autocorrelation in the sequences of standardized residuals $\hat{z}_t$ and squared standardized residuals $\hat{z}_t^2$ up to a specified lag. The results are presented in table 3 for 1, 2, 5, 10, 15 and 20 lags for the three estimation periods considered. The tests indicate the absence of autocorrelation and heteroskedasticity in the standardized residuals. In this sense, the model fits the daily exchange-rate returns successfully.

Besides, figure 3 presents empirical evidence that the leverage function succeeds in obtaining independence between z_t and u_t. The limited amount of unmodeled residual dependence is in agreement with the underlying hypotheses of the model.

3.3 Forecasting Analysis

Forecasting the covariance matrix is essential in finance, but the main characteristic of volatility and correlations is that they are unobserved. A model's predictive abilities are hence traditionally evaluated against an ex-post estimator generally called 'proxy'. The multivariate realized kernel measure is used, but we also check the robustness of the results to the choice of the proxy by considering the Realized Covariance (Andersen et al., 2003b) based on 5-minute data as well as the Markov Chain covariance measure (Hansen and Horel, 2009) which makes use of ultra-high-frequency information contained in tick-by-tick data. For the latter, we use data obtained from Oanda (Olsen) that is cleaned according to the methods indicated in Barndorff-Nielsen et al. (2009).

The out-of-sample analysis is performed for the period January 7, 2010 to August 29, 2014. Our forecasting strategy consists in a rolling window with parameters re-estimated on a daily basis. The fitted model is then used to construct k-periods ahead forecasts of the daily realized variance. Four forecast horizons $k \in \{1, 5, 10, 20\}$ are considered to emphasize the forecasting abilities of the models at longer horizons of up to one month.

The forecasting abilities of the bivariate realized model are compared to those of traditional volatility models, generally considered in forecasting exercises: the corrected Dynamic Conditional Correlation, cDCC (Aielli, 2013), the Constant Conditional Correlation model of Bollerslev (1990) and the diagonal-BEKK model of Engle and Kroner (1995), based on the univariate GARCH(1,1) specification. Additionally, the univariate Realized EGARCH(1,1) model is included in the set of alternatives. To this aim, a multiple comparison-based test, the Model Confidence Set (MCS) approach by Hansen et al. (2011) is implemented. This test allows to identify the subset of models that are equivalent in terms of forecasting ability, and which outperform all the other models at a confidence level α. We set the significance level for the MCS to $\alpha = 10\%$ and use 10,000 bootstrap resamples (with block length of 12 daily observations) to obtain the distribution under the null of equal predictive accuracy.

One essential issue here is that the object of interest, i.e. the covariance matrix, is unobserved even ex-post and that the use of a proxy may distort the comparison of the losses. In this context, Patton and Sheppard (2009) introduce loss functions robust in the sense that they asymptotically generate the same ranking of the models regardless of the proxy considered. Following Patton and Sheppard (2009), the quasi-likelihood loss function (*QLike*) is employed and in the spirit of a robustness check, the squared Frobenius distance is also considered.

$$
\begin{aligned}
QLike(\Sigma_t, H_t) &= \operatorname{tr}(H_t^{-1}\hat{\Sigma}_t) - \log|H_t^{-1}\hat{\Sigma}_t| - K \text{ and} \\
L_F(\Sigma_t, H_t) &= \operatorname{tr}[(\hat{\Sigma}_\mathrm{t} - \mathrm{H_t})^2] = \sum_N \lambda_i,
\end{aligned}
$$

where $\hat{\Sigma}_t$, is a proxy of the true but latent conditional variance matrix, H_t is a candidate model for the conditional covariance matrix and λ_i are the positive eigenvalues of the matrix $(\hat{\Sigma}_t - H_t)^2$. *QLike* is generally privileged as it presents the advantage to lead to iid loss series under the null hypothesis that the forecasting model is correctly specified. This comes from the fact that *QLike* depends on the multiplicative forecast error $H_t^{-1}\Sigma_t$ and its bias is independent of the volatility level (Brownlees et al., 2012), making easier the comparison of losses across volatility regimes.

The analysis is performed in two steps: we first scrutinize the ability of the competing models to reasonably forecast the covariance matrix and then we investigate their behavior for each element of the covariance matrix and for the variance of the inferred EUR/JPY exchange rate.

Table 4 reports the MCS results from the comparison of covariance matrix forecasts by presenting the p-value of the MCS test as well as the average loss for each model at 1, 5, 10 and 20 periods ahead. One and two asterisks indicate the models selected by the MCS procedure as performing better, i.e. included in $\hat{\mathcal{M}}_{90\%}$ and $\hat{\mathcal{M}}_{75\%}$, respectively. The test results indicates that the bivariate realized model has superior forecasting abilities than its competitors up to 10 periods (two weeks) ahead. This result holds regardless of the loss function considered.

Tables 5 and 6 summarize the results of the element by element analyses for the conditional variances and conditional correlation respectively. The univariate version of the *QLike* loss function is employed for the conditional variances. But since it is not appropriate for the evaluation of non-positive quantities such as correlation forecasts, the robust MSE loss function is considered for the results in table 6.

Our main finding is that the bivariate Realized GARCH generally does better than the GARCH and the Realized EGARCH models, in particular at shorter horizons. Most importantly, the bivariate model is clearly performing better than its competitors for the inferred EUR/JPY exchange-rate. This is impressive, because the EUR/JPY conditional volatility is obtained by making use of the no-arbitrage condition, whereas the two competing models rely on the observed EUR/JPY returns and realized variance. Against all odds, the bivariate model beats its competitors and appears as the only model retained in the model confidence set whatever the forecast horizon. Table 6 shows that the bivariate model does good not only for variance forecasting but also in the case of correlation forecasting. MCS p-values equal to 1 indicate that the bivariate Realized-GARCH model dominates all its competitors at all forecast horizons.

Along the lines of Patton (2011), the test of equality of predictive abilities of Diebold-Mariano (DM) is used as a robustness check to assess the differences in forecasting abilities among the models considered. Tables 7 and 8 report the results of the Diebold-Mariano test of equal forecasting abilities. If the test-statistic is significant and positive, the bivariate Realized GARCH model outperforms the model considered in the left column of the table. The DM test results support our previous findings by showing that the bivariate Realized GARCH model outperforms its main competitors, especially up to two weeks ahead. These results are a clear indication of the important role of the realized measures of volatility in forecasting the covariance matrix and its elements.

A robustness check with respect to the use of alternative proxies, i.e. the realized covariance and the realized Markov chain estimators, has also been considered and our main findings hold (results are available upon request).

3.3.1 Mitigate the Impact of Outliers

In this section we study the sensitivity of the forecasting results to the presence of outliers in the data. Since the realized models are known to adapt quickly to sudden large changes in volatility, this analysis is expected to shed light on the forecasting abilities of the multivariate realized model in a setting where it is *a priori* on a more equal footing with its competitors. Figure 4 displays the log-realized kernel estimator for the EUR/USD and USD/JPY exchange rates over the period 2003 - 2014. Positive (negative) outliers are identified (in red and black, respectively) and the causes of extremely high (low) unexpected volatility are labeled: news in the financial markets, some public holidays, etc.. The experiment consists in using for forecasting purposes this 'smoother' sample obtained by ignoring all the extreme periods, i.e. about 4.4% of the initial dataset. Additionally, four forecast evaluation scenarios are investigated. Each evaluation scenario is defined by ignoring the days with the 10 largest covariance forecast losses registered in a competing model.

Table 9 reports the average loss and p-value of the MCS test based on the $QLike$ loss function for each of the scenarios. The results indicate that the bivariate model dominates its competitors in all cases. This is quite impressive, since in 3 out of 4 scenarios, the 10 days removed from the out-of-sample are identified with respect to the largest losses registered by the competitor models and not by the bivariate model. The result holds up to two-weeks ahead, supporting the use of the bivariate Realized GARCH specification at shorter horizons.

4 Conclusion

This paper proposes a complete multivariate model for exchange rate returns and realized measures of volatility and correlation. The model directly links the realized measures with their conditional counterparts and entails a flexible modeling of their dynamic properties. It exploits the identities arising from the no-arbitrage conditions characterizing the forex market. The model is simple to estimate and offers substantial improvement in covariance forecasting.

The model is illustrated in an empirical analysis for the EUR, USD and JPY currencies. We find that multivariate Realized GARCH model outperforms the traditional CCC, cDCC and diagonal BEKK models up to 10 periods-ahead. It is particularly the case of the inferred EUR/JPY exchange rate volatility and that of the conditional correlation, where it is the only model that belongs to the set of superior forecasting models regardless of the forecast horizon.

5 Tables and Figures

Figure 2: QQ plot for the studentized residuals

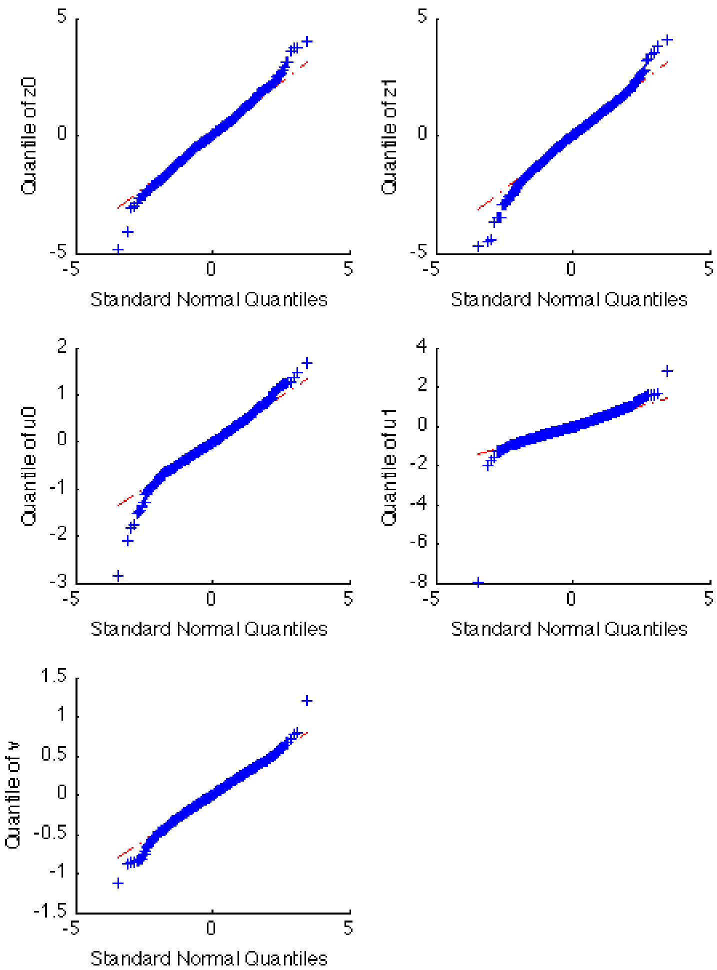


Table 1: Log-Likelihood comparison

	INS	OOS	Partial Log-Like OOS		
			EUR/USD	USD/YEN	EUR/YEN
Unrestricted model	-1360.73	-1454.48	-956.44	-974.52	-1222.59
Diagonal-β	-1366.26	**-1444.22**	-956.23	-975.02	-1220.37
D-β, no leverage	-1429.14	-1462.05	-956.65	-980.76	-1200.81

Table 2: Partial Log-Likelihood

		EUR/USD	USD/YEN	EUR/YEN (implied)
INS	GARCH	-1458.52	-1578.32	-1613.89
	R-EGARCH	**-1456.11**	**-1529.70**	-1588.16
	Biv-REG	-1457.62	-1576.75	**-1497.47**
OOS	GARCH	-966.60	-1026.70	-1301.60
	R-EGARCH	-960.80	-1056.56	-1284.15
	Biv-REG	**-956.23**	**-975.02**	**-1220.37**

Figure 3: Residuals independence

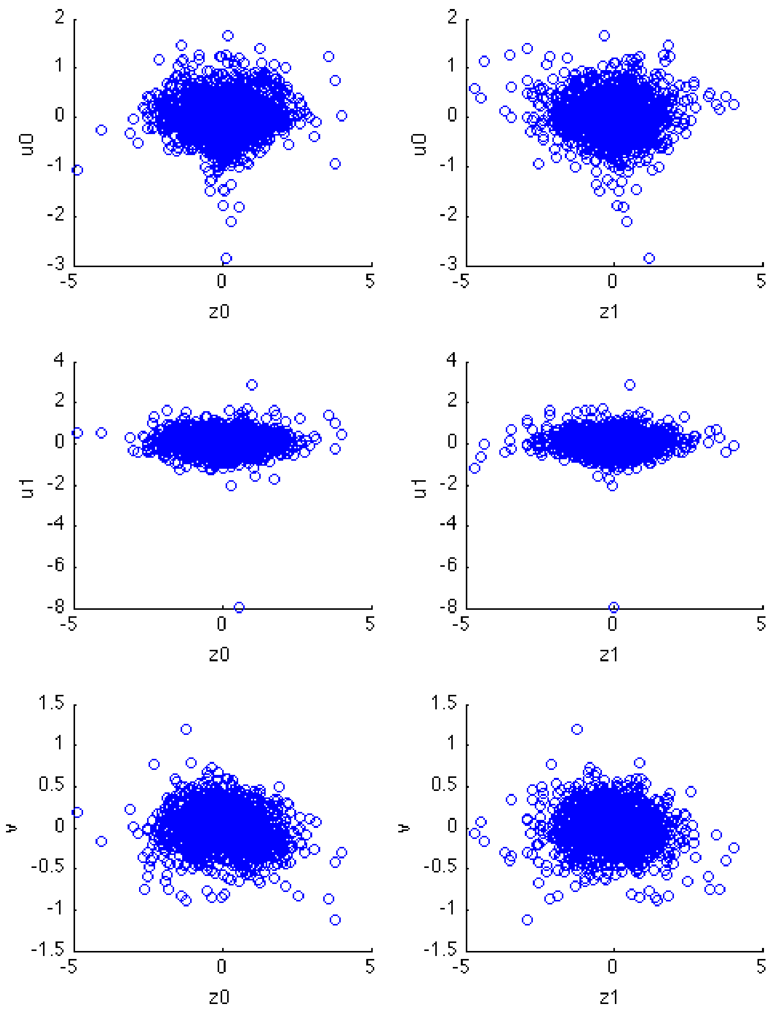
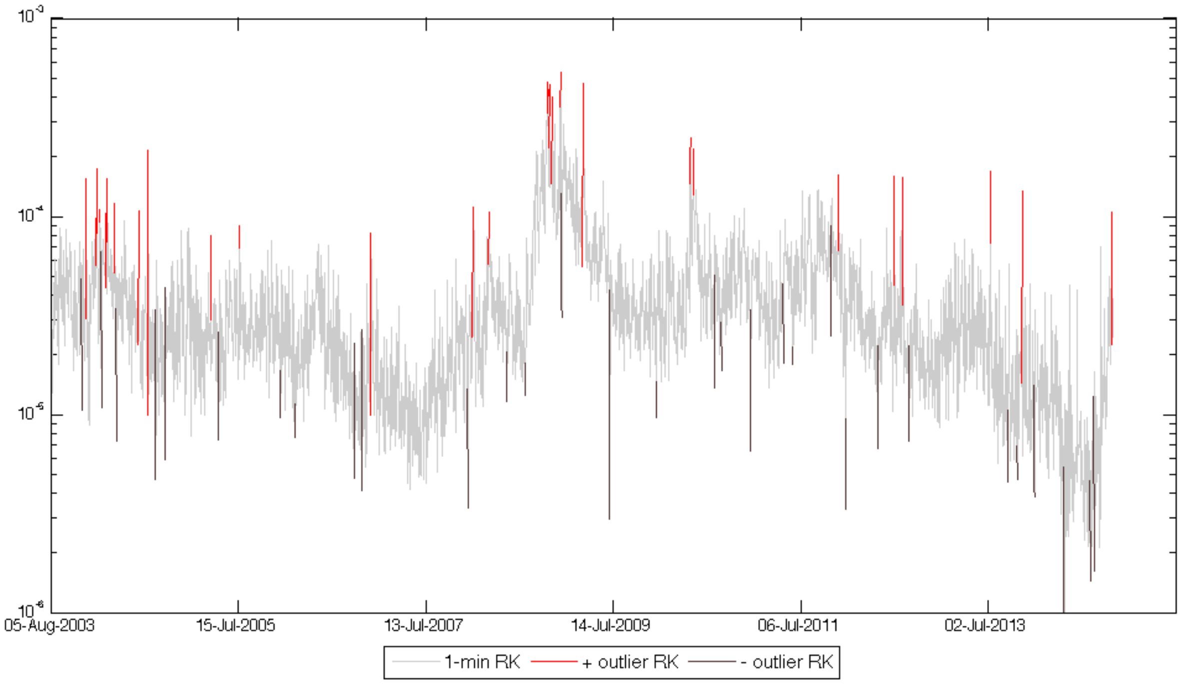


Figure 4: Outliers identification

(a) EUR/USD



(b) USD/JPY

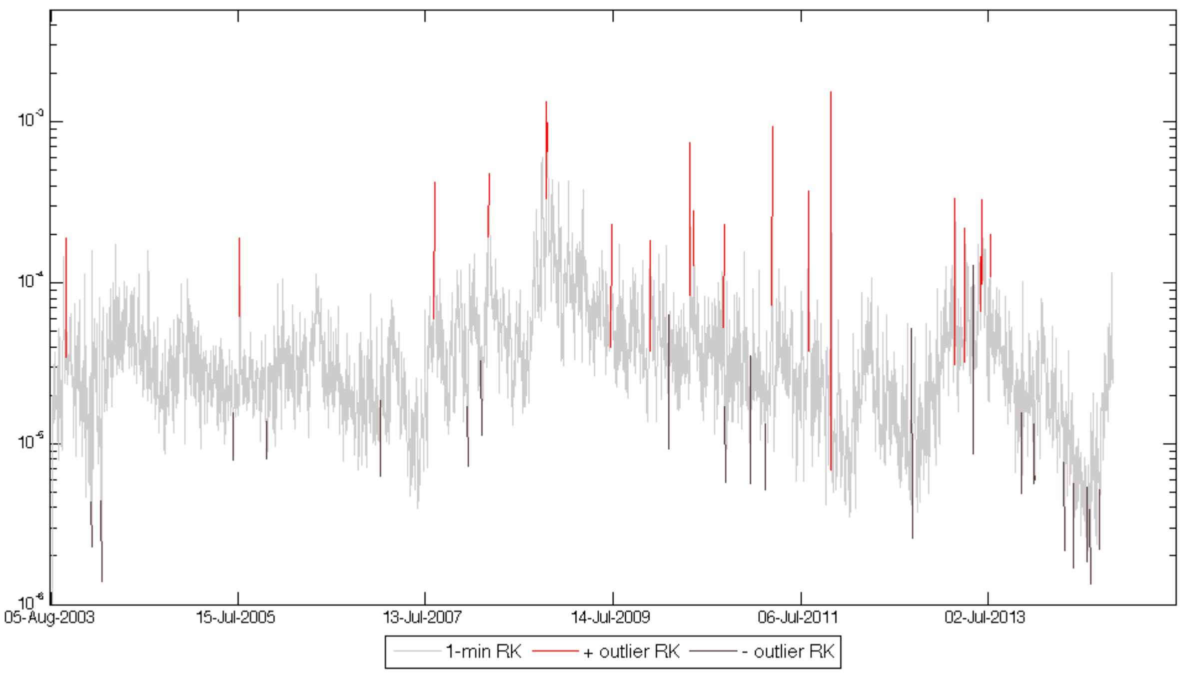


Table 3: Studentized residuals: autocorrelation and heteroskedasticity tests

		z_0						z_1					
		Q(1)	Q(2)	Q(5)	Q(10)	Q(15)	Q(20)	Q(1)	Q(2)	Q(5)	Q(10)	Q(15)	Q(20)
$\hat{z}_t$	statistic	0.016	0.073	1.225	8.489	13.40	13.05	0.570	0.685	3.851	7.448	13.94	16.20
	p-value	0.900	0.964	0.943	0.581	0.571	0.875	0.450	0.710	0.571	0.683	0.530	0.704
$\hat{z}_t^2$	statistic	3.469	4.958	11.99	13.85	18.99	24.15	0.922	2.242	5.275	9.164	12.63	16.51
	p-value	0.063	0.084	0.035*	0.180	0.214	0.236	0.337	0.326	0.383	0.517	0.631	0.685

Note: This Table reports the test-statistics and p-values of the Ljung-Box test of no autocorrelation up to the lags specified for the series of standardized residuals z_0 and z_1 and the series of squared standardized residuals. An asterisk indicates the rejection of the null hypothesis at 5%.

Table 4: Conditional covariance forecast evaluation: MCS test (Realized Kernel proxy)

Qlike loss								
	$k=1$		$k=5$		$k=10$		$k=20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv REG	0.468	1.000**	0.575	1.000**	0.612	1.000**	0.745	1.000**
cDCC	0.575	0.000	0.647	0.001	0.682	0.006	0.760	0.444**
BEKK	0.579	0.000	0.654	0.001	0.692	0.003	0.777	0.110*
CCC	0.646	0.000	0.707	0.001	0.731	0.003	0.779	0.177*
Squared Frobenius distance								
	$k=1$		$k=5$		$k=10$		$k=20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv REG	0.442	1.000**	0.495	1.000**	0.495	1.000**	0.519	0.593**
cDCC	0.479	0.019	0.501	0.333**	0.507	0.125*	0.517	0.593**
BEKK	0.481	0.019	0.508	0.001	0.516	0.001	0.530	0.000
CCC	0.479	0.019	0.501	0.333**	0.507	0.125*	0.517	1.000**

Note: This Table displays the average loss over the evaluation sample as well as the Model Confidence Set p-values. The conditional covariance forecasts in the $\hat{\mathcal{M}}_{90\%}$ and $\hat{\mathcal{M}}_{75\%}$ are identified by one and two asterisks, respectively.

Table 5: Forecast evaluation: MCS evaluation test

Conditional variance forecast								
EUR/USD	$k = 1$		$k = 5$		$k = 10$		$k = 20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv-REG	0.125	1.000**	0.144	0.700**	0.160	1.000**	0.193	0.157*
GARCH	0.135	0.015	0.151	0.199*	0.163	0.772**	0.183	1.000**
R-EGARCH	0.127	0.081	0.144	1.000**	0.160	0.966**	0.191	0.157*
USD/JPY	$k = 1$		$k = 5$		$k = 10$		$k = 20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv-REG	0.274	1.000**	0.334	0.832**	0.347	1.000**	0.411	0.628**
GARCH	0.347	0.001	0.382	0.003	0.389	0.030	0.411	0.858**
R-EGARCH	0.280	0.040	0.334	1.000**	0.349	0.530**	0.408	1.000**
EUR/JPY	$k = 1$		$k = 5$		$k = 10$		$k = 20$	
(implied)	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv-REG	0.150	1.000**	0.228	1.000**	0.247	1.000**	0.323	1.000**
GARCH	0.207	0.001	0.273	0.001	0.296	0.001	0.356	0.007
R-EGARCH	0.155	0.075	0.234	0.202*	0.265	0.006	0.352	0.016

Note: This Table displays the average *Qlike* loss over the evaluation sample as well as the Model Confidence Set p-values. The conditional variance forecasts in the $\hat{\mathcal{M}}_{90\%}$ and $\hat{\mathcal{M}}_{75\%}$ are identified by one and two asterisks, respectively.

Table 6: Forecast evaluation: MCS evaluation test

Conditional correlation forecast								
	$k = 1$		$k = 5$		$k = 10$		$k = 20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv-REG	0.150	1.000**	0.228	1.000**	0.247	1.000**	0.323	1.000**
cDCC	0.207	0.001	0.273	0.001	0.296	0.001	0.356	0.007
d-BEKK	0.155	0.075	0.234	0.202*	0.265	0.006	0.352	0.016
CCC	55.17	0.000	77.04	0.000	111.1	0.000	174.0	0.000

Note: This Table displays the average MSE loss over the evaluation sample as well as the Model Confidence Set p-values. The conditional variance forecasts in the $\hat{\mathcal{M}}_{90\%}$ and $\hat{\mathcal{M}}_{75\%}$ are identified by one and two asterisks, respectively.

Table 7: Forecast evaluation: DMW test

Conditional variance forecasts					
EUR/USD		$k = 1$	$k = 5$	$k = 10$	$k = 20$
GARCH	t-statistic	2.651***	1.654**	0.659	-1.879**
	p-value	0.008	0.098	0.510	0.060
R-EGARCH	t-statistic	1.733**	-0.406	0.050	-0.925
	p-value	0.083	0.684	0.960	0.355
USD/JPY		$k = 1$	$k = 5$	$k = 10$	$k = 20$
GARCH	t-statistic	4.334***	3.441***	2.281***	-0.043
	p-value	0.000	0.001	0.023	0.966
R-EGARCH	t-statistic	2.047***	-0.220	0.631	-1.112
	p-value	0.041	0.826	0.528	0.266
EUR/JPY (implied)		$k = 1$	$k = 5$	$k = 10$	$k = 20$
GARCH	t-statistic	4.146***	3.670***	4.206***	3.377***
	p-value	0.000	0.000	0.000	0.001
R-EGARCH	t-statistic	1.839**	1.243	3.193***	2.874***
	p-value	0.066	0.214	0.001	0.004

Note: This Table displays the t-statistics and p-value for the Diebold-Mariano test of equal predictive accuracy based on the $QLike$ loss function. *,** and *** indicate a significant difference between the forecasting abilities of the bivariate Realized GARCH model and those of the competitor (GARCH or R-EGARCH) at the 1%, 5% and 10% level. If the t-statistic is positive, the bivariate Realized GARCH performs better (it has a smaller average forecast loss).

Table 8: Forecast evaluation: DMW test

Conditional correlation forecasts					
EUR/USD - USD/JPY		$k = 1$	$k = 5$	$k = 10$	$k = 20$
cDCC	t-statistic	-7.299***	-4.892***	-6.672***	-8.345***
	p-value	0.000	0.000	0.000	0.000
d-BEKK	t-statistic	-6.895***	-3.956***	-4.905***	-5.783***
	p-value	0.000	0.000	0.000	0.000
CCC	t-statistic	-14.97***	-12.24***	-12.33***	-11.98***
	p-value	0.000	0.000	0.000	0.000

Note: This Table displays the t-statistics of the Diebold-Mariano test of equal predictive accuracy based on the MSE loss function. The asterisks indicate a significant difference between the forecasting abilities of the bivariate Realized GARCH model and those of the competitor (cDCC) at the 1% significance level. If the t-statistic is positive, the bivariate Realized GARCH performs better (it has a smaller average forecast loss).

Table 9: Conditional covariance forecast evaluation: MCS test (Realized Kernel proxy)

Biv REG forecast error reference								
	$k = 1$		$k = 5$		$k = 10$		$k = 20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv REG	0.321	1.000**	0.398	1.000**	0.448	1.000**	0.555	1.000**
cDCC	0.423	0.000	0.462	0.000	0.503	0.000	0.572	0.341**
BEKK	0.420	0.000	0.459	0.000	0.500	0.000	0.571	0.341**
CCC	0.491	0.000	0.514	0.000	0.541	0.000	0.582	0.238**
cDCC forecast error reference								
	$k = 1$		$k = 5$		$k = 10$		$k = 20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv REG	0.325	1.000**	0.401	1.000**	0.452	1.000**	0.557	1.000**
cDCC	0.417	0.000	0.457	0.000	0.500	0.001	0.571	0.468**
BEKK	0.415	0.000	0.454	0.000	0.498	0.001	0.570	0.468**
CCC	0.485	0.000	0.509	0.000	0.536	0.000	0.582	0.310**
BEKK forecast error reference								
	$k = 1$		$k = 5$		$k = 10$		$k = 20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv REG	0.325	1.000**	0.403	1.000**	0.453	1.000**	0.557	1.000**
cDCC	0.417	0.000	0.458	0.000	0.502	0.001	0.573	0.414**
BEKK	0.415	0.000	0.454	0.000	0.497	0.002	0.570	0.414**
CCC	0.485	0.000	0.510	0.000	0.539	0.000	0.584	0.237**
CCC forecast error reference								
	$k = 1$		$k = 5$		$k = 10$		$k = 20$	
	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}	Loss	P_{MCS}
Biv REG	0.330	1.000**	0.406	1.000**	0.454	1.000**	0.557	1.000**
cDCC	0.420	0.000	0.459	0.000	0.501	0.000	0.572	0.437**
BEKK	0.418	0.000	0.455	0.001	0.499	0.001	0.571	0.437**
CCC	0.481	0.000	0.506	0.000	0.534	0.000	0.580	0.343**

Note: This Table displays the average loss over the evaluation sample as well as the Model Confidence Set p-values for the scenarios where outliers and important forecast losses are removed. The conditional covariance forecasts in the $\hat{\mathcal{M}}_{90\%}$ and $\hat{\mathcal{M}}_{75\%}$ are identified by one and two asterisks, respectively.

References

Aielli, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003a). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003b). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal*, 12(3):C1–C32.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (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.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. *The Review of Economics and Statistics*, 72(3):498–505.

Brownlees, C., Engle, R., and Kelly, B. (2012). A practical guide to volatility forecasting through calm and storm. *Journal of Risk*, 14(2):3.

Engel, C. and Hamilton, J. (1990). Long swings in the dollar: Are they in the data and do markets know it? *American Economic Review*, 80:689–713.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric theory*, 11(01):122–150.

Hansen, P. and Horel, G. (2009). Quadratic variation by markov chains. *Univ. of Aarhus Dept. of Economics Research Paper*, 2009-13.

Hansen, P. and Lunde, A. (2010). Forecasting volatility using high frequency data. *Oxford Handbook of Economic Forecasting?, Oxford University Press. forthcoming.*

Hansen, P. R. and Huang, Z. (2012). Exponential garch modeling with realized measures of volatility. CADMUS working paper.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012a). Realized garch: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Hansen, P. R., Lunde, A., and Voev, V. (2012b). Realized beta garch: Multivariate garch model with realized measures of volatility and covolatility. CADMUS working paper.

Noureldin, D., Shephard, N., and Sheppard, K. (2012). Multivariate high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics*, 27(6):907–933.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 150(1):246–256.

Patton, A. J. and Sheppard, K. (2009). Evaluating volatility and correlation forecasts. In *Handbook of financial time series*, pages 801–838. Springer.