The one-trading-day-ahead forecast errors of intra-day realized volatility

Degiannakis, Stavros

Department of Economic and Regional Development, Panteion University, 136 Syngrou Av., Athens, 176 71, Greece

January 2017

The one-trading-day-ahead forecast errors of intra-day realized volatility

Abstract

Two volatility forecasting evaluation measures are considered; the squared one-day-ahead forecast error and its standardized version. The mean squared forecast error is the widely accepted evaluation function for the realized volatility forecasting accuracy. Additionally, we explore the forecasting accuracy based on the squared distance of the forecast error standardized with its volatility. The statistical properties of the forecast errors point the standardized version as a more appropriate metric for evaluating volatility forecasts.

We highlight the importance of standardizing the forecast errors with their volatility. The predictive accuracy of the models is investigated for the FTSE100, DAX30 and CAC40 European stock indices and the exchange rates of Euro to British Pound, US Dollar and Japanese Yen. Additionally, a trading strategy defined by the standardized forecast errors provides higher returns compared to the strategy based on the simple forecast errors. The exploration of forecast errors is paving the way for rethinking the evaluation of ultra-high frequency realized volatility models.

Keywords: ARFIMA model, HAR model, intra-day data, predictive ability, realized volatility, ultra-high frequency modelling.

JEL Classifications: C14; C32; C50; G11; G15.

1. Introduction

A volatility forecasting evaluation framework that brings together a well defined measure with known statistical properties is applied for predicting one-trading-day-ahead realized volatility. Any evaluation (or loss) function is a measure of accuracy constructed upon the goals of its particular appliance. For example, the most widely applied evaluation function is the mean squared forecast error.

A number of model evaluation methods have been applied in financial literature. Most of them are based on measuring the ability of the models to fit in the data. The most cited in-sample fitting evaluation functions are the information criteria of Akaike (1973), Schwarz (1978) and Shibata (1980) that are based on the Kullback and Leibler (1951) measure. In the case that researchers focus on evaluating models' forecasting ability, they construct evaluation functions that take into consideration the characteristics of the predicting variable. For example, because of the non-linearity of volatility, Pagan and Schwert (1990), Heynen and Kat (1994) and Andersen *et al*. (1999), among others, have evaluated the predictive ability of volatility models with robust to heteroscedasticity evaluation functions. As Hendry and Clements (2001) noted *it seems natural that a stock broker measures the value of forecasts by their monetary return, not their mean squared error*. Thus, Engle *et al.* (1993), West *et al.* (1993) and Granger and Pesaran (2000), among others, have defined loss functions that evaluate the models according to the usage of the predictions.

Although evaluation functions are measures of accuracy that have been constructed upon the goals of their particular application, in the vast majority of the cases, their statistical properties are unknown. In volatility forecasting literature, the superiority of a loss function against others is not conducted according to a statistical based theoretical ground but it is based on empirical motivations. The present manuscript provides an empirical investigation of forecasting accuracy of ultra high frequency volatility models according to the one-step-ahead standardized forecast errors. These standardized forecast errors define an evaluation function that comprises a selection procedure whose distribution is explicitly derived.

The joint distribution of the half-sum of the squared one-step-ahead standardized forecast errors from a set of models is the multivariate gamma (Krishnamoorthy and Parthasarathy, 1951). The cumulative distribution function of the minimum half-sum of the squared one-step-ahead standardized forecast errors is the minimum multivariate gamma (Degiannakis and Xekalaki, 2005). Hence, a model selection algorithm is defined according to which the model with the lowest sum of squared standardized one-step-ahead prediction errors is considered as having a superior ability to predict the realized volatility.

The contribution of the paper is described concisely. First, we define an evaluation function of realized volatility forecasts that comprises a selection procedure whose distribution is explicitly derived. Second, we infer for the distributional properties of the standardized forecast errors of the most widely applied realized volatility models and third, we open new avenues for the exploitation of forecast errors in financial markets' literature.

The rest of the paper is structured as follows. Section 2 provides information for the stock indices (FTSE100, DAX30 and CAC40) and the exchange rates (Euro to British Pound, US Dollar and Japanese Yen) that comprise the dataset and the construction of the annualized realized volatility measures. Section 3 illustrates the forecasting models and the distributional assumptions under investigation. The most widely applied models for realized volatility forecasting, i.e. AFRIMA and HAR specifications, are estimated assuming that the standardized unpredictable components are normally, Student *t*, GED, and skewed Student *t* distributed. Section 4 describes the statistical properties of the forecast errors estimated from two generic model frameworks widely applied in financial literature (i.e. regression and ARFIMA models with heteroscedastic residuals), as well as the exploitation of the standardized forecast errors for defining an evaluation procedure of volatility models' predictability. Section 5 analyses the findings of the adopted forecasting evaluation method. In Section 6, we assume that the realized volatility measure is a tradeable asset and we define a trading framework under which we will investigate whether a trader based on the standardized forecast errors achieves higher returns compared to a trader whose trading is based on the simple forecast errors. Finally Section 7 concludes the study.

2. *Dataset – European Stock Indices and Euro Exchange Rates*

Figure 1 plots the daily prices, $\log P_{t_\tau}$, along with their logarithmic first differences, $y_t = \left(\log P_{t_\tau} - \log P_{t-1_\tau}\right)$, for the FTSE100 (20th August, 1998-12th January, 2011), DAX30 (3rd January, 2000-12th January, 2011), CAC40 (13th June, 2000-12th January, 2011) indices, as well as for the exchange rates of Euro to the Great Britain Pound (4th January, 1999-21st January, 2011), United States Dollar (20th April, 1998-24th January, 2011) and Japanese Yen (4th January, 1999-24th January, 2011). The dataset under investigation consists of 2686, 2784, 3106, 3308, 3091 and 3108 trading days for the CAC40, DAX30, FTSE100, EURUSD, EURGBP and EURJPY realized volatility series, respectively.

[Insert Figure 1 About here]

The *integrated variance* $\sigma^{2(IV)}_{[t_1,t_\tau]}$ is the actual, but unobservable, variance over the interval $[t_1, t_\tau]$. We are able to approximate the integrated variance of trading day t with the *realized volatility*:

$$RV_{[t_1,t_\tau]} = \sum\nolimits_{j=1}^{\tau} \left(\log P_{t_j} - \log P_{t_{j-1}}\right)^2, \tag{1}$$

where $j = 1,...,\tau$ denote the equidistant points in time at which the asset prices P_{t_j} are observed. The realized volatility converges in probability to the integrated volatility, or

$$p \lim_{\tau \to \infty} \left(RV_{[t_1,t_\tau]}\right) = \sigma^{2(IV)}_{[t_1,t_\tau]}. \tag{2}$$

The accuracy improves as the number of sub-intervals increases, but on the other hand, at a high sampling frequency the market frictions is a source of noise due to market microstructure features. As Andersen *et al*. (2006) noted, the realized volatility is constructed in the highest sampling frequency that the intra-day autocovariance minimizes:

$$\min\left(E\left(y_{t_i} y_{t_{i-j}}\right)\right), \tag{3}$$

where $y_{t_i} = \left(\log P_{t_i} - \log P_{t_{i-1}}\right)$.

Since the induction of ultra-high frequency based volatility estimation, the literature has proposed a number of variations of realized volatility measure in order to account for microstructure effects and various assumptions about efficient prices; see Aït-Sahalia *et al.* (2005, 2011), Barunik *et al.* (2016), Barndorff-Nielsen and Shephard (2005), among others. Some examples are the *bipower variation* of Barndorff-Nielsen and Shephard (2004), the *two time scale realized volatility* of Zhang *et al*., (2005), *realized range* of Christensen and Podolskij (2007), *semivariance* of Choobineh and Branting (1986) and Grootveld and Hallerbach (1999). As the scope of the study is not affected by the necessity to state a particular relation for the efficient prices and the market microstructure noise, we employ the most widely used measure the realized volatility adjusted for the overnight volatility.

Thus, we estimate the realized volatility at the optimal sampling frequency according to eq. (3) and adjust it with the overnight volatility according to the method of Hansen and Lunde (2005). Hence, the annualized realized volatility estimate for the whole day is:

$$\sqrt{252 RV_t^{(\tau)}} = \sqrt{252\left(\omega_1\left(\log P_{t_1} - \log P_{t-1_\tau}\right)^2 + \omega_2 \sum\nolimits_{j=1}^{\tau}\left(\log P_{t_j} - \log P_{t_{j-1}}\right)^2\right)}, \tag{4}$$

for the ω_1, ω_2 weights as defined by Hansen and Lunde (see also Degiannakis and Floros, 2015, 2016). Figure 2 illustrates the annualized adjusted realized volatilities, $\sqrt{252 RV_t^{(\tau)}}$, at the optimal sampling frequencies which are the 7 minutes for the CAC40 and FTSE100, the 13 minutes for the DAX30 and the 20 minutes for the three exchange rates.

[Insert Figure 2 About here]

3. *Set of competing models*

Literature has provided evidence in favour to the use of the ARFIMA and the HAR-RV models for predicting realised volatility[1]. Hence, these two model frameworks comprise our set of competing models. Additionally, in order to capture the statistical properties of integrated quarticity, we extend the ARFIMA and the HAR-RV model frameworks by incorporating a time varying framework for the conditional standard deviation of the dependent variable. In our case, the logarithmic annualized realized volatility is the dependent variable, and the integrated quarticity is expressed as the conditional standard deviation of logarithmic realized volatility.

3.1. ARFIMA-GARCH framework

The Autoregressive Fractionally Integrated Moving Average model with dynamic conditional volatility was initially proposed by Baillie *et al.* (1996). It captures the long memory property of the logarithmic annualized volatility $lRV_t^{(\tau)} \equiv \log\sqrt{252 RV_t^{(\tau)}}$ as well as its high persistence and the time-variation and clustering of the integrated quarticity via the GARCH approaching. The ARFIMA-GARCH model for $lRV_t^{(\tau)}$ is expressed as:

$$\left(1-\sum_{i=1}^{k} c_i L^i\right)(1-L)^d\left(lRV_t^{(\tau)}-\beta_0\right)=\left(1+\sum_{i=1}^{l} d_i L^i\right)\varepsilon_t, \tag{5}$$

where $\varepsilon_t = h_t z_t$, $h_t^2 = a_0 + \sum_{i=1}^{q} a_i L^i \varepsilon_t^2 + \sum_{i=1}^{p} b_i L^i h_t^2$, $z_t \sim f(0,1;\boldsymbol{\theta})$ and $f(.)$ is the density function of z_t. The h_t^2 defines an estimate of the volatility's volatility, $\sigma^{2(IQ)}_{[t_1,t_\tau]} = \int_{t_1}^{t_\tau} 2\sigma^4(t)dt$, which is termed integrated quarticity. Thus:

$$\sqrt{\tau}\left(\log\left(RV_{[t_1,t_\tau]}\right)-\int_{t_1}^{t_\tau}\sigma^2(t)dt\right)\Big/\sqrt{\int_{t_1}^{t_\tau}2\sigma^4(t)dt}\xrightarrow{d}N(0,1). \tag{6}$$

3.2. HAR-RV-GARCH framework

The Heterogeneous Autoregressive model with dynamic conditional volatility was was introduced by Corsi *et al.* (2008). The heterogeneous autoregressive framework accommodates the beliefs of heterogeneous traders, an idea proposed by Müller *et al.*, (1997). The HAR-RV-GARCH model is an autoregressive structure of the volatilities realized over different time intervals. Thus, current trading day's realized volatility is expressed by the daily, the weekly and the monthly volatilities:

[1] Indicatively, you are referred to Chiriac and Voev (2011), Prokopczuk *et al.* (2016) and Sevi (2014).

$$lRV_t^{(\tau)} = w_0 + w_1 lRV_{t-1}^{(\tau)} + w_2\left(5^{-1}\sum_{j=1}^{5} lRV_{t-j}^{(\tau)}\right) + w_3\left(22^{-1}\sum_{j=1}^{22} lRV_{t-j}^{(\tau)}\right) + \varepsilon_t, \quad (7)$$

where $\varepsilon_t = h_t z_t$, $h_t^2 = a_0 + \sum_{i=1}^{q} a_i L^i \varepsilon_t^2 + \sum_{i=1}^{p} b_i L^i h_t^2$, $z_t \sim f(0,1;\mathbf{\theta})$ and $f(.)$ is the density function of z_t .

The ARFIMA-GARCH and HAR-RV-GARCH model specifications are estimated for standardized residuals that are normally, $z_t \sim N(0,1)$, Student t , $z_t \sim t(0,1;v)$, GED of Box and Tiao (1973), $z_t \sim Ged(0,1;v)$, and skewed Student t of Fernandez and Steel (1998), $z_t \sim skT(0,1;v,g)$, distributed. The density functions of $z_t \sim t(0,1;v)$ is $f_{(t)}(z_t;v) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{\pi(v-2)}}\left(1+\frac{z_t^2}{v-2}\right)^{-\frac{v+1}{2}}$, for $\Gamma(.)$ denoting is the gamma function. The generalized error, or exponential power, distribution's density function is $f_{(GED)}(z_t;v) = \frac{v\exp\left(-0.5|z_t/\lambda|^v\right)}{\lambda 2^{(1+1/v)}\Gamma(v^{-1})}$ for $v > 0$ and $\lambda \equiv \sqrt{2^{-2/v}\Gamma(v^{-1})/\Gamma(3v^{-1})}$. For $z_t \sim skT(0,1;v,g)$ the density function is $f_{(skT)}(z_t;v,g) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{\pi(v-2)}}\left(\frac{2s}{g+g^{-1}}\right)\left(1+\frac{sz_t+m}{v-2}g\right)^{-\frac{v+1}{2}}$, for $z_t < -ms^{-1}$, and $f_{(skT)}(z_t;v,g) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{\pi(v-2)}}\left(\frac{2s}{g+g^{-1}}\right)\left(1+\frac{sz_t+m}{v-2}g^{-1}\right)^{-\frac{v+1}{2}}$, for $z_t \geq -ms^{-1}$, where $m = \Gamma((v-1)/2)\sqrt{(v-2)}\left(\Gamma(v/2)\sqrt{\pi}\right)^{-1}(g-g^{-1})$ and $s = \sqrt{g^2+g^{-2}-m^2-1}$. The g and v are the asymmetry and tail parameters of the distribution, respectively.

4. Evaluate models' predictability

We define a set of four competing models; two ARFIMA(k,d,l)-GARCH(p,q) and two HAR-RV-GARCH(p,q). Additionally, each one of the four model specifications is estimated under the four distributional assumptions. The lag orders k,d,l,p,q of the best performing models have been selected according to Schwarz Bayesian information criterion. The ARFIMA(0,d,1)-GARCH(1,1), ARFIMA(1,d,1)-GARCH(1,1), HAR-RV-GARCH(1,1) and HAR-RV-GARCH(0,1) comprise the set of the competing models. The construction of a set of well-performing models makes the models' evaluation even more challenging. Hence, in total, we estimate 16 models for each of the six assets under investigation. These 16 models are estimated at each trading day, for $\tilde{T} = T - \breve{T}$ days, where T denotes the whole available dataset and $\breve{T} = 1000$ is the rolling sample size.

The next trading day's log-realized volatility, $lRV_{t+1|t}^{(\tau)}$, and the $h_{t+1|t}$ for the ARFIMA(1,d,1)-GARCH(1,1) model are computed as

$$lRV_{t+1|t}^{(\tau)} = \beta_0^{(t)}\left(1 - c_1^{(t)}\right) + c_1^{(t)} lRV_t^{(\tau)} + \sum_{j=1}^{\infty} A(j-1)\varepsilon_{t|t} + \sum_{j=0}^{\infty} A(j) d_1^{(t)} \varepsilon_{t|t} \quad (8)$$

$$h_{t+1|t} = \sqrt{a_0^{(t)} + a_1^{(t)} \varepsilon_{t|t}^2 + b_1^{(t)} h_{t|t}^2} \quad (9)$$

where $A(j) \equiv \left(\frac{\Gamma\left(j + d^{(t)}\right)}{\Gamma\left(d^{(t)}\right)\Gamma(j+1)} L^j \right)$ and $A(j-1) \equiv \left(\frac{\Gamma\left(j + d^{(t)}\right)}{\Gamma\left(d^{(t)}\right)\Gamma(j+1)} L^{j-1} \right)$. Figure 3 presents, indicatively for the ARFIMA(1,d,1)-GARCH(1,1) model, the annualized realized volatility, the discrepancy between the actual realized volatility and the forecasted values as well as the standardized forecast errors. In eq.(8), for $c_1^{(t)} = 0$, we get the log-realized volatility forecast for the ARFIMA(0,d,1) specification.

For the HAR-RV-GARCH(1,1) model, the next trading day's $lRV_{t+1|t}^{(\tau)}$ is computed as

$$lRV_{t+1|t}^{(\tau)} = w_0^{(t)} + w_1^{(t)} lRV_t^{(\tau)} + w_2^{(t)} \left(5^{-1} \sum_{j=1}^{5} lRV_{t-j+1}^{(\tau)} \right) + w_3^{(t)} \left(22^{-1} \sum_{j=1}^{22} lRV_{t-j+1}^{(\tau)} \right) + \varepsilon_{t|t}, \quad (10)$$

In eq.(9), for $b_1^{(t)} = 0$, the forecast of integrated quarticity from the GARCH(0,1) specification is computed.

[Insert Figure 3 About here]

4.1. One-trading-day-ahead forecast error

Going back to the grounds of financial theory, the actual volatility of the instantaneous log-returns during the continuous time interval $[t_1, t_\tau]$ is the integral of $\sigma(t)$, having assuming that the log-returns of the asset follow the diffusion process $d\log p(t) = \sigma(t) dW(t)$. The $\sigma(t)$ denotes the volatility of the instantaneous log-returns $\log p(t)$ and $W(t)$ is the Wiener process. Hence, the actual volatility over a trading day $[t_1, t_\tau]$ is denoted as $\sigma_{[t_1, t_\tau]}^{2(IV)} = \int_{t_1}^{t_\tau} \sigma^2(t) dt$; which is *integrated variance*. The $RV_{[t_1, t_\tau]}$ is a proxy measure of the unobservable $\sigma_{[t_1, t_\tau]}^{2(IV)}$.

Let us assume that a forecaster is interested in evaluating the ability of a set of models to forecast the one-day-ahead volatility. The forecaster is not able to observe the actual volatility; i.e. $\sigma^{2(IV)}_{[t_1,t_\tau]}$, so she works with the proxy measure $RV_{[t_1,t_\tau]}$. The evaluation of the models is conducted according the proxy measure and not based on the actual variable. Hence, a question arises concerning the ranking of the models according to their forecasting ability. Hansen and Lunde (2006) derived conditions which ensure that the ranking of variance forecasts by an evaluation function is the same (*consistent ranking*) whether the ranking is done via the true and unobserved variance, $\sigma^{2(IV)}_{[t_1,t_\tau]}$, or via an unbiased proxy as the $RV_{[t_1,t_\tau]}$.

Consider an evaluation function $L(,)$ that measures the distance between the proxy variable and its forecast. A sufficient condition for a consistent ranking is the quantity $\frac{\partial^2 L\left(RV_{[t_1,t_\tau]}, RV^{(\tau)}_{t|t-1}\right)}{\partial\left(RV_{[t_1,t_\tau]}\right)^2}$ not to depend on $RV^{(\tau)}_{t|t-1}$. For example, the mean squared forecast error (MSE):

$$\tilde{T}\sum_{t=1}^{\tilde{T}} \varepsilon^2_{t+1|t} \, , \tag{11}$$

ensures the equivalence of the ranking of volatility models, where

$$\varepsilon_{t+1|t} = lRV^{(\tau)}_{t+1} - lRV^{(\tau)}_{t+1|t} \tag{12}$$

denotes the one-trading-day-ahead forecast error. Hansen and Lunde (2006) and Patton (2011) investigate various evaluation functions robust for consistent ranking and study the effects of noisy volatility proxies in the evaluation of forecasting ability.

4.2. One-trading-day-ahead standardized forecast error

The MSE function is the most widely applied method of evaluating the forecasting accuracy. However, we are not aware of the explicit distribution of the squared forecast errors. On the other hand, we can infer for the distributional properties of the squared standardized forecast errors; i.e. the ratio of the forecast error to the conditional standard deviation.

Regression model with Arch errors

Let us define a generic form of a multiple regression model with dynamic conditional volatility of the error term (such a framework includes the HAR-RV-GARCH model where the explanatory variables express the lagged daily, weekly and monthly values of the dependent variable). For y_t and $\mathbf{x}_{t-1}$ denoting the dependent variable and the vector of explanatory variables, respectively, the regression model with the error term ε_t following an ARCH process is presented as:

$$\begin{aligned} y_t &= \mathbf{x}'_{t-1}\boldsymbol{\beta}^{(t)} + \varepsilon_t \\ \varepsilon_t &= z_t h_t \\ z_t &\sim f\left(0,1;\boldsymbol{\theta}^{(t)}\right) \\ h_t^2 &= g\left(\varepsilon_{t-j}, h_{t-j}, j \geq 1; \boldsymbol{\psi}^{(t)}\right), \end{aligned} \tag{13}$$

where $\boldsymbol{\varphi}^{(t)} = \left(\boldsymbol{\beta}^{(t)}, \boldsymbol{\theta}^{(t)}, \boldsymbol{\psi}^{(t)}\right)'$ is the vector of parameters to be estimated, $f(.)$ is the density function of z_t, and $g(.)$ is a measurable function of information set at time *t-1*, I_{t-1}, that represents the conditional variance of ε_t.

Degiannakis and Xekalaki (2005) noted that if the vector of parameters is constant over time, $\boldsymbol{\varphi}^{(t)} = \boldsymbol{\varphi}^{(t+1)} = ... = \boldsymbol{\varphi}$, and $z_t \sim N(0,1)$, then the estimated standardized one-step-ahead prediction errors are asymptotically standard normally distributed:

$$z_{t+1|t} = \left(y_{t+1} - y_{t+1|t}\right) h_{t+1|t}^{-1} \sim N(0,1), \tag{14}$$

where $y_{t+1|t} = \mathbf{x}'_t\boldsymbol{\beta}^{(t)}$ and $h_{t+1|t}$ is the one-step-ahead conditional standard deviation.

ARFIMA model with Arch errors

The model framework in eq.(13) is modified in order to include the ARFIMA-GARCH model. Let us assume that y_t is the log-realized volatility. Then an ARFIMA model with the error term ε_t following an ARCH process is presented as:

$$\begin{aligned} y_t &= C(L)^{-1}(1-L)^{-d} D(L) + \varepsilon_t \\ \varepsilon_t &= z_t h_t \\ z_t &\sim f\left(0,1;\boldsymbol{\theta}^{(t)}\right) \\ h_t^2 &= g\left(\varepsilon_{t-j}, h_{t-j}, j \geq 1; \boldsymbol{\psi}^{(t)}\right), \end{aligned} \tag{15}$$

where $C(L) = \left(1 - \sum_{j=1}^{k} c_j^{(t)} L^j\right)$, $D(L) = \left(1 + \sum_{j=1}^{l} d_j^{(t)} L^j\right)$, $(1-L)^{d^{(t)}} = \sum_{j=0}^{\infty} \pi_j L^j$, for $\pi_j = \frac{\Gamma\left(j + d^{(t)}\right)}{\Gamma(j+1)\Gamma\left(d^{(t)}\right)} = \prod_{k=0}^{j} \frac{k-1-d^{(t)}}{k}$, and $\Gamma(.)$ is the gamma function.

The vector of parameters to be estimated is defined as $\boldsymbol{\varphi}^{(t)} = \left(c_j^{(t)}, d_j^{(t)}, d^{(t)}, \boldsymbol{\theta}^{(t)}, \boldsymbol{\psi}^{(t)}\right)'$. Assuming constancy of parameters over time, $\boldsymbol{\varphi}^{(t)} = \boldsymbol{\varphi}^{(t+1)} = ... = \boldsymbol{\varphi}$, and $z_t \sim N(0,1)$, the statistic $z_{t+1|t} = \left(y_{t+1} - y_{t+1|t}\right) h_{t+1|t}^{-1}$ is asymptotically standard normally distributed as well.

Naturally, based on the mean squared standardized forecast error, we can compare the forecasting accuracy of a set of competing models. Degiannakis and Xekalaki (2005) showed that the ratio of the sum of the squared standardized forecast errors from two competing models follows the correlated gamma ratio distribution. Therefore, they proposed a model selection approach, named standardized prediction error criterion (SPEC), based on the evaluation of the predictability of the models in terms of standardized prediction errors.

Let us denote a set of M competing models, and $X_m = \frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^{2(m)}$ is the half-sum of the squared one-step-ahead standardized forecast errors from model m. The joint distribution of $(X_1, X_2, ..., X_M)$ is the multivariate gamma. For m=2, the interested reader is referred to Kibble (1941), whereas for m>2 you are referred to Krishnamoorthy and Parthasarathy (1951). For $X_{(1)} \equiv \min(X_1, X_2, ..., X_M)$ denoting the minimum half-sum of the $z_{t+1|t}^{2(m)}$, the cumulative distribution function of $X_{(1)}$ is the minimum multivariate gamma distribution (for details see Xekalaki and Degiannakis, 2010).

According to the SPEC criterion, the most appropriate forecasting model is the i^{th} model with the lowest half-sum of squared one-step-ahead standardized forecast errors:

$$\min_i \left(\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^{2} \right). \tag{16}$$

Moreover, Degiannakis and Livada (2016) provided theoretical and empirical evidence that the SPEC criterion can be applied to evaluate models with residuals that are leptokurtically (i.e. Student t and generalized error distributed), or even leptokurtically and asymmetrically distributed (i.e. skewed Student t distributed).

Let us consider the frameworks (13) or (15) and assume that $\boldsymbol{\varphi}^{(\breve{T})}$ is a consistent estimator, for $\breve{T}$ denoting the sample size that has been used to estimate $\boldsymbol{\varphi}^{(t)}$.[2] Slutsky's theorem states that $p\lim g(x_T) = g(p\lim x_T)$, for any continuous function $g(x_T)$ that is not a function of T. Henceforth, $p\lim(z_{t+1|t}) = z_t$ and as convergence in probability implies convergence in distribution, we get that $z_{t+1|t} \xrightarrow{p} z_t \Rightarrow z_{t+1|t} \xrightarrow{d} z_t \overset{i.i.d.}{\sim} f(0,1;\boldsymbol{\theta}^{(t)})$, for $\boldsymbol{\theta}^{(t)} \subseteq \boldsymbol{\varphi}^{(t)}$.[3]

[2] If $\boldsymbol{\varphi}^{(\breve{T})}$ is a strongly consistent estimator and asymptotically normally distributed, then $p\lim(\boldsymbol{\varphi}^{(\breve{T})}) = \boldsymbol{\varphi}^{(t)}$.

[3] Hence, $z_{t+1|t}$ are asymptotically $f(0,1;\boldsymbol{\theta}^{(t)})$ distributed since, from the definition of convergence in probability, component wise convergence in probability implies convergence of vectors.

Also, Xekalaki and Degiannakis (2005) provided evidence that traders who base their trading preferences on SPEC algorithm (for selecting volatility forecasts extracted by a set of ARCH models) achieve higher profits than those who use only a single ARCH model.

5. *Model selection evaluation framework*

According to the aforementioned analysis of forecast errors' statistical properties, the predictive ability evaluation is based on the standardized version of the one-step-ahead forecast errors.

Forecast errors with symmetric distribution

The values of the $\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^2$ evaluation function are presented in the 2nd column of Table 1. We observe that the minimum value of $\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^2$ is not achieved by the same model for all the realized volatility series. Additionally, Table 2 illustrates the sum of the squared forecast errors, $\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2$. In the case of the unstandardized forecast errors, for all the realized volatility series under investigation (CAC40, DAX30, FTSE100, EURUSD, EURJPY), except for the EURGBP[4], the ARFIMA(1,d,1)-GARCH(1,1) model has the lowest value of $\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2$.

[Insert Table 1 about here]
[Insert Table 2 about here]

Table 3 provides the descriptive statistics of the standardized one-step-ahead forecast errors. The kurtosis for all the models is greater than the normal value of three, rejecting the hypothesis that the standardized one-step-ahead forecast errors, $z_{t+1\backslash t}$, could be normally distributed. Naturally, the unstandardized forecast errors, $\varepsilon_{t+1\backslash t}$, are characterized by even higher values of kurtosis[5] due to the fact that the unconditional distribution of a stochastic process with heteroscedastic formation is more platykurtic compared to the conditional distribution. Let us consider the model frameworks of equations (13) and (15). The conditional distribution is derived as $\varepsilon_t \backslash I_{t-1} \sim f\left(0, h_t^2; \boldsymbol{\theta}^{(t)}\right)$, whereas the unconditional

[4] In the case of the Euro/Pound rate, the HAR-RV-GARCH(1,1) model is being selected.

[5] The relevant table with the descriptive statistics of the unconditional forecast errors is not presented due to space limitations, but it is available to the interested reader upon request.

distribution $\varepsilon_t \sim \tilde{f}(.;.)$, which is not explicitly related with $f\left(0,h_t^2;\boldsymbol{\theta}^{(t)}\right)$, has fatter tails as $E\left(\varepsilon_t^4\right)/E\left(\varepsilon_t^2\right)^2 \geq 3E\left(\sigma_t^2\right)^2/E\left(\sigma_t^2\right)^2$, for $f\left(0,h_t^2;\boldsymbol{\theta}^{(t)}\right) \equiv N\left(0,h_t^2\right)$.[6]

[Insert Table 3 about here]

Forecast errors with platykurtic distribution

According to the 3rd column of Table 1, which presents the values of $\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^2$ for Student *t* distributed z_t, we reach to qualitatively similar conclusion: the model *i* with $\min_i\left(\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^2\right)$ differs for the various volatility series. Concerning the unstandardized forecast errors (in Table 2), for all the realized volatility series under investigation except for the DAX30[7], the ARFIMA(1,d,1)-GARCH(1,1) model has the lowest value of $\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2$.

The 4th column of Table 1 illustrates the $\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^2$ values for $z_t \sim Ged(0,1;v)$. We observe that for different distributional assumptions, the minimum value of $\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^2$ is achieved from different models. For the unstandardized forecast errors (4th column of Table 2), the ARFIMA(1,d,1)-GARCH(1,1) model has the $\min\left(\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2\right)$ for all the realized volatility series.

Table 4 provides the descriptive statistics of the standardized one-step-ahead forecast errors from the models with conditionally Student *t* distributed innovations. Concerning the estimated kurtosis (for the 4 models and the 6 realized volatility series), there are no significant differences in the estimated values of kurtosis under the two assumed distributions $z_t \sim N(0,1)$ and $z_t \sim t(0,1;v)$. The skewness of $z_{t+1\setminus t}$ is negative in all the cases except for the EURUSD exchange rate. In the case of the simulated data of Degiannakis and Livada (2016), the skewness was much higher than in the case of the models with normally distributed innovations. However, in the case of real data we do not observe any significant differences in the estimated skewness for normally and Student *t* distributed innovations.

---

[6] According to Jensen's inequality.

[7] In the case of the DAX30 stock index, the ARFIMA(0,d,1)-GARCH(1,1) model is selected.

[Insert Table 4 about here]

Table 5 provides the descriptive statistics of the standardized one-step-ahead forecast errors from the models with conditionally GED distributed innovations. The descriptive statistics provide similar information with the case of the Student *t* distributional assumption. For all the models and all the time series, there are no significant differences in the estimated values of the descriptive measures compared to the models with normally distributed innovations.

[Insert Table 5 about here]

Forecast errors with platykurtic and asymmetric distribution

According to the 5th column of both Tables 1 and 2, we reach to similar conclusions: the minimum value of $\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^2$ is achieved from different models, whereas the $\min\left(\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2\right)$ holds for the same model, the ARFIMA(1,d,1)-GARCH(1,1), across all realized volatility series.

Table 6 provides the descriptive statistics of the standardized one-step-ahead forecast errors from the models with conditionally skewed Student *t* distributed innovations. In the case of the six actual realized volatility series, we do not observe any significant differences in the estimated skewness and kurtosis for the various assumed distributions of innovations.

[Insert Table 6 about here]

6. Financial Application

Let us define a framework under which we will investigate whether predictive information is being extracted by the forecast errors. We assume that the realized volatility measure is a tradeable asset (i.e. an ETF or a future contract on the implied volatility index). We outline an agent who defines her trading strategy according to the half-sum of the squared one-step-ahead standardized forecast errors. At each trading day *t*, the trader takes into consideration the *s*-trading days ahead forecast, $RV_{t+s|t}^{(\tau)}$, estimated by the model *m* with $\min\left(\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^{2(m)}\right)$. Hence, the trader is able to buy (long position) or sell (short position) the realized volatility measure. Thus, at each trading day, the trader proceeds to a long position when $RV_{t+s|t}^{(\tau)} > RV_t^{(\tau)}$ and to a short position when $RV_{t+s|t}^{(\tau)} < RV_t^{(\tau)}$. The daily trading returns are computed as:

$$r_t^{(s)} = \begin{cases} \dfrac{RV_{t+s}^{(\tau)} - RV_t^{(\tau)}}{RV_t^{(\tau)}} & \text{if} \quad RV_{t+s|t}^{(\tau)} > RV_t^{(\tau)} \\ \\ \dfrac{-RV_{t+s}^{(\tau)} + RV_t^{(\tau)}}{RV_t^{(\tau)}} & \text{if} \quad RV_{t+s|t}^{(\tau)} < RV_t^{(\tau)} \end{cases} \qquad (17)$$

So, the average return of the trading strategy is:

$$\bar{r}^{(s)} = \tilde{T}\sum_{t=1}^{\tilde{T}} r_t^{(s)}\,, \qquad (18)$$

where $\tilde{T}$ is the number of out-of-sample forecasted values. The trading strategy is being replicated for s=1,5,10 and 15 trading days ahead and for $\tilde{T}$ =10 and 20. The average returns, $\bar{r}^{(s)}$, are presented in Table 7. In all the cases, the returns are positive and statistically significant.

In order to compare the trading strategy based on the standardized forecast errors with an alternative one based on the unstandardized forecast errors, we assume another trader who defines her trading strategy according to the sum of the squared forecast errors, $\min\left(\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2\right)$. Table 8 illustrates the average return, $\bar{r}^{(s)}$, of the trading strategy based on the model m with $\min\left(\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2\right)$. Comparing the Tables 7 and 8, we conclude that the trader who takes into consideration the s-trading days ahead forecast, $RV_{t+s|t}^{(\tau)}$, estimated by the model m with $\min\left(\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^{2(m)}\right)$ achieves higher average returns, $\bar{r}^{(s)}$, compared to the trader whose forecasts are based on the $\min\left(\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2\right)$ criterion. More specifically, in 40 out of 48 cases the strategy based on the standardized forecast errors provides higher returns compared to the strategy based on the simple forecast errors.

[Insert Table 7 about here]
[Insert Table 8 about here]

7. Conclusion

We have estimated a set of ARFIMA-GARCH and HAR-RV models with i) normally, ii) Student t, iii) GED and iv) skewed Student t distributed standardized innovations. The models were estimated for the major European Union's stock market indices (FTSE100, DAX30, CAC40) and for the exchange rates of Euro to the Great Britain Pound, United States Dollar and Japanese Yen. The models were re-estimated for $\tilde{T}$ days, where $\tilde{T}=1686$, 1784, 2106, 2308, 2091, 2108 for the CAC40, DAX30, FTSE100, EURUSD, EURGBP and EURJPY realized volatility series, respectively, based on a rolling sample of constant size of $\breve{T}=1000$ days.

Concerning the evaluation according to the standardized forecast errors, the most accurate realized volatility predictions are not produced by the same model for all the realized volatility series. On the contrary, in almost all the cases the ARFIMA(1,d,1)-GARCH(1,1) model has the lowest value of the sum of the squared forecast errors. For the 6 realized volatility series and the 4 distributional assumptions, in total 24 cases, there are only two exceptions; the HAR-RV-GARCH(1,1) model for the Euro/Pound rate under the normal distribution and the ARFIMA(0,d,1)-GARCH(1,1) model for the DAX30 index under the Student *t* distribution.

Finally, under the assumption that the realized volatility measure could be a tradeable asset, we measure via a trading framework the forecasting performance of a trader whose strategy is based on the standardized forecast errors. We conclude that the trader who predicts the *s*-trading days ahead volatility forecast from the models with $\min\left(\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^{2(m)}\right)$ achieves higher average returns compared to the trader whose predictions are based on the models with $\min\left(\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^{2}\right)$.

Therefore, the definition of the evaluation criterion is highly important for the predictability assessment of the realized volatility models. The distribution of the squared standardized forecast errors is explicitly defined compared to the squared unstandardized forecast errors. The one-trading-day-ahead forecast errors' investigation provides avenues of further research in the utilization of ultra-high frequency based volatility modeling. In future research, the forecast errors' properties may be utilized for portfolio evaluation taking into consideration multivariate frameworks and the dynamics of correlations across assets.

References

Aït-Sahalia, Y., Mykland, P.A. and Zhang, L. (2005). How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise. *Review of Financial Studies*, 18, 351–416.

Aït-Sahalia, Y., Mykland, P.A. and Zhang, L. (2011). Ultra-High Frequency Volatility Estimation with Dependent Microstructure Noise, *Journal of Econometrics*, 160, 190-203.

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. *Proceedings of the second international symposium on information theory.* B.N. Petrov and F. Csaki (eds.), Budapest, 267-281.

Andersen, T., T. Bollerslev and S. Lange (1999). Forecasting Financial Market Volatility: Sample Frequency vis-à-vis Forecast Horizon. *Journal of Empirical Finance,* 6, 457-477.

Andersen, T., Bollerslev, T., Christoffersen, P. and Diebold, F.X. (2006). Volatility and Correlation Forecasting. In (eds.) Elliott, G. Granger, C.W.J. and Timmermann, A. *Handbook of Economic Forecasting*, North Holland Press, Amsterdam.

Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O. (1996). Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 74, 3-30.

Barndorff-Nielsen O.E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2, 1-48.

Barndorff-Nielsen, O.E. and Shephard, N. (2005). How Accurate is the Asymptotic Approximation to the Distribution of Realised Volatility? In (eds.) Andrews, D., Powell, J., Ruud, P., and Stock, J., *Identification and Inference for Econometric Models*. Cambridge University Press, Cambridge.

Barunik, J., Krehlik, T. and Vacha, L. (2016). Modeling and forecasting exchange rate volatility in time-frequency domain. *European Journal of Operational Research*, 251, 329–340.

Box, G.E.P. and Tiao, G.C. (1973). Bayesian Inference in Statistical Analysis. *Addison-Wesley*, Reading, Mass.

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

Choobineh, F. and Branting, D. (1986). A Simple Approximation for Semivariance. *European Journal of Operational Research*, 27, 364–370.

Christensen K. and Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141, 323-349.

Corsi, F., Mittnik, S., Pigorsch, C. and Pigorsch, U. (2008). The Volatility of Realised Volatility. *Econometric Reviews*, 27(1-3), 46-78.

Degiannakis, S. and Floros, C. (2015). Modelling and Forecasting High Frequency Financial Data, Palgrave - MacMillan Ltd., Hampshire.

Degiannakis, S. and Floros, C. (2016). Intra-Day Realized Volatility for European and USA Stock Indices, *Global Finance Journal*, 29, 24-41.

Degiannakis, S. and Livada, A. (2016). Evaluation of Realized Volatility Predictions from Models with Leptokurtically and Asymmetrically Distributed Forecast Errors, *Journal of Applied Statistics*, 43(5), 871-892.

Degiannakis, S. and Xekalaki, E. (2005). Predictability and Model Selection in the Context of ARCH Models. *Journal of Applied Stochastic Models in Business and Industry*, 21, 55-82.

Engle, R.F., C.H. Hong, A. Kane and J. Noh (1993). Arbitrage valuation of variance forecasts with simulated options. *Advances in Futures and Options Research*, 6, 393-415.

Fernandez, C. and Steel, M. (1998). On Bayesian Modeling of Fat Tails and Skewness. *Journal of the American Statistical Association*, 93, 359-371.

Granger C.W.J. and M.H. Pesaran (2000). Economic and Statistical Measures of Forecasting Accuracy. *Journal of Forecasting*, 19, 537-560.

Grootveld H. and Hallerbach, W. (1999). Variance vs downside risk: Is there really that much difference? *European Journal of Operational Research*, 114, 304–319.

Hansen, P.R. and Lunde, A. (2005). A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data. *Journal of Financial Econometrics*, 3(4), 525-554.

Hansen, P.R. and Lunde, A. (2006). Realized Variance and Market Microstructure Noise. *Journal of Business and Economic Statistics*, 24(2), 127-161.

Hendry D.F. and M.P. Clements (2001). Economic Forecasting: Some Lessons from Recent Research. *European Central Bank*, Working Paper Series, 82.

Heynen, R. and H. Kat (1994). Volatility Prediction: A Comparison of the Stochastic Volatility, Garch(1,1), and Egarch(1,1) Models. *Journal of Derivatives*, 94, 50-65.

Kibble, W.F. (1941). A Two Variate Gamma Type Distribution. *Sankhya*, 5, 137-150.

Krishnamoorthy, A.S. and Parthasarathy, M. (1951). A Multivariate Gamma - Type Distribution. *Annals of Mathematical Statistics*, 22, 4, 549-557.

Kullback, S. and Leibler, R.A. (1951). On Information and Sufficiency. *Annals of Mathematical Statistics*, 22, 79-86.

Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V. and VonWeizsäcker, J.E. (1997). Volatilities of Different Time Resolutions – Analyzing the Dynamics of Market Components. *Journal of Empirical Finance*, 4, 213-239.

Pagan, A.R. and G.W. Schwert (1990). Alternative Models for Conditional Stock Volatility. *Journal of Econometrics,* 45, 267-290.

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

Prokopczuk, M., Symeonidis, L., and Wese Simen, C. (2016). Do Jumps Matter for Volatility Forecasting? Evidence from Energy Markets. *Journal of Futures Markets*, forthcoming.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics,* 6, 461-464.

Sevi, B. (2014). Forecasting the Volatility of Crude Oil Futures using Intraday Data. *European Journal of Operational Research*, 235, 643-659.

Shibata, R. (1980). Asymptotically Efficient Selection of the Order of the Model for Estimating Parameters of a Linear Process. *Annals of Statistics*, 8, 147-164.

West, K. D., H. J. Edison, and D. Cho, (1993). A Utility Based Comparison of Some Models for Exchange Rate Volatility. *Journal of International Economics*, 35, 23-45.

Xekalaki, E. and Degiannakis, S. (2005). Evaluating Volatility Forecasts in Option Pricing in the Context of a Simulated Options Market. *Computational Statistics and Data Analysis,* 49(2), 611-629.

Xekalaki, E. and Degiannakis, S. (2010). *ARCH Models for Financial Applications*. John Wiley & Sons Ltd., New York.

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

Tables

Table 1. The half-sum of the squared standardized forecast errors $\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^2$ of the four models for i) normally, ii) Student *t*, ii) GED, and iv) skewed Student *t* distributed standardized innovations.

Model	$\sum_{t=1}^{\tilde{T}} z_{t+1\backslash t}^2$ for $z_t \sim N(0,1)$	$\sum_{t=1}^{\tilde{T}} z_{t+1\backslash t}^2$ for $z_t \sim t(0,1;v)$	$\sum_{t=1}^{\tilde{T}} z_{t+1\backslash t}^2$ for $z_t \sim Ged(0,1;v)$	$\sum_{t=1}^{\tilde{T}} z_{t+1\backslash t}^2$ for $z_t \sim skT(0,1;v,g)$
		CAC 40		
1	909.7	900.4	911.9	894.8
2	914.7	897.5	912.7	893.4
3	909.2	902.8	902.7	903.8
4	**885.6**	**882.6**	**902.6**	**883.3**
		DAX 30		
1	**992.1**	**991.9**	995.3	991.1
2	993.3	1017.7	996.4	**987.6**
3	993.9	998.4	**994.4**	999.8
4	1014.5	1008.7	994.5	1013.0
		FTSE 100		
1	**1075.1**	**1081.2**	**1079.1**	**1079.6**
2	1076.8	1082.8	1082.6	1086.3
3	1086.3	1091.0	1091.7	1096.2
4	1087.6	1092.1	1091.7	1097.5
		EURUSD		
1	1132.7	1139.1	1136.9	1134.0
2	1136.6	1138.3	1141.7	1130.6
3	1140.2	1128.8	1135.9	1127.9
4	**1129.8**	**1126.7**	**1135.8**	**1126.0**
		EURGBP		
1	**1084**.1	1075.9	**1084.3**	**1068.4**
2	1091.6	**1075.6**	1086.7	1071.9
3	1084.4	1085.7	1088.9	1087.5
4	1101.1	1102.0	1088.7	1104.3
		EURJPY		
1	1126.2	**1094.7**	1128.6	1098.5
2	1134.6	1114.7	1142.3	1112.2
3	**1094.4**	1116.4	**1119.6**	**1095.6**
4	1155.8	1154.9	1119.8	1157.9

Model 1: ARFIMA(0,d,1)-GARCH(1,1), Model 2: ARFIMA(1,d,1)-GARCH(1,1), Model 3: HAR-RV-GARCH(1,1), Model 4: HAR-RV-GARCH(0,1).

Table 2. The sum of the squared forecast errors $\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^2$ of the four models for conditionally i) normally, ii) Student t, ii) GED, and iv) skewed Student t distributed innovations.

Model	$\sum_{t=1}^{\tilde{T}} z_{t+1\backslash t}^2$ for $z_t \sim N(0,1)$	$\sum_{t=1}^{\tilde{T}} z_{t+1\backslash t}^2$ for $z_t \sim t(0,1;v)$	$\sum_{t=1}^{\tilde{T}} z_{t+1\backslash t}^2$ for $z_t \sim Ged(0,1;v)$	$\sum_{t=1}^{\tilde{T}} z_{t+1\backslash t}^2$ for $z_t \sim skT(0,1;v,g)$
		CAC 40		
1	1178.2	1185.6	1183.4	1185.8
2	**1174.8**	**1178.8**	**1180.6**	**1178.0**
3	1184.2	1189.8	1190.8	1187.8
4	1184.4	1187.8	1191.0	1187.0
		DAX 30		
1	1672.6	**1672.8**	1674.6	1675.0
2	**1665.0**	1705.4	**1663.6**	**1664.6**
3	1683.4	1681.0	1680.0	1681.0
4	1682.2	1680.8	1680.2	1680.8
		FTSE 100		
1	1639.8	1642.2	1641.2	1637.4
2	**1634.8**	**1637.8**	**1638.6**	**1637.1**
3	1657.6	1658.8	1660	1658.4
4	1656.2	1658.4	1660	1658.6
		EURUSD		
1	1780.4	1773.8	1776.6	1774.8
2	**1773.6**	**1772.4**	**1772.8**	**1771.4**
3	1776.4	1773.1	1774.2	1774
4	1776.4	1773.0	1774.5	1774
		EURGBP		
1	1124.6	1117.8	1124.0	1121.2
2	1117.8	**1099.2**	**1105.4**	**1102.0**
3	**1104.6**	1105.6	1106.2	1105.8
4	1105.2	1105.6	1106.2	1106.6
		EURJPY		
1	1693.0	1696	1692.2	1706.2
2	**1685.0**	**1689.4**	**1686.0**	**1690.6**
3	1697.2	1698.8	1695.4	1698.4
4	1700.0	1700.2	1696.1	1700.2

Model 1: ARFIMA(0,d,1)-GARCH(1,1), Model 2: ARFIMA(1,d,1)-GARCH(1,1), Model 3: HAR-RV-GARCH(1,1), Model 4: HAR-RV-GARCH(0,1).

Table 3. Descriptive statistics of the standardized one-step-ahead prediction errors, $z_{t+1\backslash t}$, from the four models under normally distributed innovations.

Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
CAC 40							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.007	0.002	1.080	-1.559	0.265	-0.393	5.464
2	-0.006	0.005	1.097	-1.560	0.265	-0.389	5.502
3	-0.010	-0.004	1.083	-1.530	0.265	-0.374	5.344
4	-0.009	-0.004	1.083	-1.541	0.265	-0.376	5.347
DAX 30							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.007	-0.002	1.119	-1.340	0.307	-0.249	3.958
2	-0.006	0.001	1.112	-1.348	0.307	-0.248	3.957
3	-0.009	-0.007	1.109	-1.348	0.308	-0.240	3.865
4	-0.008	-0.006	1.106	-1.347	0.308	-0.236	3.861
FTSE 100							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.006	0.002	1.166	-1.754	0.280	-0.452	5.015
2	-0.005	0.005	1.188	-1.743	0.279	-0.455	5.056
3	-0.010	0.000	1.227	-1.696	0.281	-0.436	4.918
4	-0.007	0.005	1.238	-1.696	0.281	-0.427	4.926
EURUSD							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	0.009	0.008	1.533	-1.246	0.278	0.092	4.184
2	0.008	0.010	1.539	-1.242	0.277	0.094	4.215
3	0.012	0.015	1.533	-1.250	0.277	0.093	4.187
4	0.012	0.015	1.533	-1.252	0.277	0.094	4.186
EURGBP							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	0.005	0.021	0.721	-1.368	0.233	-0.542	4.741
2	-0.002	0.013	0.732	-1.411	0.232	-0.541	4.791
3	0.000	0.017	0.731	-1.380	0.231	-0.532	4.775
4	0.001	0.018	0.739	-1.385	0.231	-0.528	4.779
EURJPY							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	0.000	0.005	1.149	-1.305	0.284	-0.237	4.119
2	0.001	0.004	1.143	-1.297	0.283	-0.221	4.167
3	-0.002	0.002	1.167	-1.246	0.284	-0.216	4.097
4	-0.003	0.001	1.166	-1.233	0.284	-0.213	4.073

Model 1: ARFIMA(0,d,1)-GARCH(1,1), Model 2: ARFIMA(1,d,1)-GARCH(1,1), Model 3: HAR-RV-GARCH(1,1), Model 4: HAR-RV-GARCH(0,1).

Table 4. Descriptive statistics of the standardized one-step-ahead prediction errors, $z_{t+1\backslash t}$, from the four models under Student *t* distributed innovations.

Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
CAC 40							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.008	0.002	1.077	-1.548	0.266	-0.400	5.466
2	-0.006	0.005	1.095	-1.540	0.265	-0.400	5.512
3	-0.013	-0.006	1.075	-1.530	0.266	-0.391	5.367
4	-0.012	-0.005	1.080	-1.527	0.266	-0.381	5.343
DAX 30							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.006	-0.002	1.119	-1.347	0.307	-0.243	3.961
2	-0.002	0.002	1.873	-1.359	0.310	-0.124	4.549
3	-0.011	-0.007	1.110	-1.359	0.308	-0.241	3.868
4	-0.011	-0.009	1.110	-1.359	0.308	-0.238	3.867
FTSE 100							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.007	-0.001	1.168	-1.753	0.280	-0.451	5.017
2	-0.005	0.006	1.191	-1.740	0.280	-0.461	5.072
3	-0.012	-0.001	1.226	-1.697	0.281	-0.444	4.945
4	-0.011	0.001	1.228	-1.697	0.281	-0.442	4.947
EURUSD							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	0.008	0.008	1.533	-1.241	0.277	0.101	4.195
2	0.009	0.008	1.537	-1.236	0.277	0.095	4.209
3	0.010	0.012	1.536	-1.249	0.277	0.095	4.199
4	0.010	0.013	1.536	-1.251	0.277	0.095	4.199
EURGBP							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	0.005	0.020	0.720	-1.369	0.232	-0.531	4.729
2	0.001	0.014	0.733	-1.407	0.230	-0.527	4.806
3	-0.004	0.013	0.727	-1.389	0.231	-0.534	4.779
4	-0.004	0.012	0.730	-1.392	0.231	-0.532	4.780
EURJPY							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.003	0.003	1.146	-1.318	0.284	-0.251	4.145
2	0.000	0.004	1.141	-1.296	0.284	-0.226	4.184
3	-0.005	-0.001	1.167	-1.249	0.284	-0.215	4.102
4	-0.006	-0.002	1.167	-1.236	0.284	-0.213	4.086

Model 1: ARFIMA(0,d,1)-GARCH(1,1), Model 2: ARFIMA(1,d,1)-GARCH(1,1), Model 3: HAR-RV-GARCH(1,1), Model 4: HAR-RV-GARCH(0,1).

Table 5. Descriptive statistics of the standardized one-step-ahead prediction errors, $z_{t+1|t}$, from the four models under GED distributed innovations.

Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
CAC 40							
1	-0.008	0.002	1.076	-1.554	0.265	-0.397	5.483
2	-0.007	0.004	1.097	-1.548	0.265	-0.394	5.518
3	-0.013	-0.006	1.075	-1.535	0.266	-0.389	5.373
4	-0.013	-0.006	1.075	-1.535	0.266	-0.389	5.373
DAX 30							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.007	-0.002	1.121	-1.346	0.307	-0.249	3.951
2	-0.005	0.002	1.113	-1.327	0.306	-0.246	3.947
3	-0.010	-0.007	1.110	-1.358	0.308	-0.241	3.870
4	-0.010	-0.007	1.110	-1.358	0.308	-0.241	3.870
FTSE 100							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.006	0.000	1.169	-1.746	0.280	-0.459	5.017
2	-0.005	0.006	1.192	-1.740	0.280	-0.460	5.068
3	-0.011	-0.001	1.225	-1.695	0.281	-0.445	4.939
4	-0.011	-0.001	1.225	-1.695	0.281	-0.445	4.939
EURUSD							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	0.008	0.009	1.532	-1.242	0.278	0.098	4.198
2	0.008	0.009	1.537	-1.236	0.277	0.094	4.208
3	0.010	0.014	1.535	-1.251	0.277	0.094	4.196
4	0.010	0.014	1.535	-1.251	0.277	0.094	4.196
EURGBP							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	0.003	0.018	0.720	-1.370	0.233	-0.539	4.725
2	0.000	0.014	0.732	-1.408	0.231	-0.526	4.788
3	-0.005	0.012	0.726	-1.391	0.231	-0.534	4.778
4	-0.005	0.012	0.726	-1.391	0.231	-0.534	4.778
EURJPY							
Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
1	-0.001	0.004	1.145	-1.305	0.284	-0.239	4.119
2	0.001	0.004	1.139	-1.295	0.283	-0.222	4.164
3	-0.003	0.000	1.167	-1.231	0.284	-0.212	4.085
4	-0.003	0.000	1.167	-1.231	0.284	-0.212	4.085

Model 1: ARFIMA(0,d,1)-GARCH(1,1), Model 2: ARFIMA(1,d,1)-GARCH(1,1), Model 3: HAR-RV-GARCH(1,1), Model 4: HAR-RV-GARCH(0,1).

Table 6. Descriptive statistics of the standardized one-step-ahead prediction errors, $z_{t+1\backslash t}$, from the four models under skewed Student *t* distributed innovations.

Model	Mean	Median	Maximum	Minimum	Std.Dev	Skewness	Kurtosis
CAC 40							
1	-0.006	0.001	1.148	-1.546	0.266	-0.392	5.493
2	-0.005	0.006	1.095	-1.540	0.265	-0.404	5.510
3	-0.009	-0.003	1.080	-1.527	0.266	-0.391	5.364
4	-0.008	-0.002	1.083	-1.523	0.266	-0.383	5.343
DAX 30							
1	-0.005	-0.001	1.118	-1.345	0.307	-0.252	3.938
2	-0.004	0.002	1.111	-1.357	0.307	-0.248	3.959
3	-0.008	-0.004	1.110	-1.353	0.308	-0.240	3.862
4	-0.007	-0.005	1.109	-1.352	0.308	-0.239	3.862
FTSE 100							
1	-0.007	0.000	1.167	-1.753	0.279	-0.453	5.032
2	-0.004	0.006	1.192	-1.738	0.280	-0.461	5.074
3	-0.008	0.004	1.229	-1.694	0.281	-0.444	4.945
4	-0.007	0.006	1.232	-1.694	0.281	-0.444	4.954
EURUSD							
1	0.010	0.009	1.533	-1.241	0.277	0.096	4.192
2	0.009	0.010	1.537	-1.236	0.277	0.095	4.210
3	0.012	0.014	1.536	-1.248	0.277	0.096	4.196
4	0.012	0.014	1.536	-1.249	0.277	0.096	4.196
EURGBP							
1	0.006	0.021	0.720	-1.366	0.232	-0.543	4.783
2	0.000	0.014	0.732	-1.407	0.230	-0.532	4.813
3	-0.001	0.015	0.728	-1.386	0.231	-0.539	4.797
4	-0.001	0.016	0.731	-1.389	0.231	-0.537	4.795
EURJPY							
1	-0.002	0.005	1.144	-1.315	0.285	-0.253	4.119
2	0.001	0.005	1.139	-1.294	0.284	-0.226	4.183
3	-0.001	0.003	1.168	-1.245	0.284	-0.215	4.106
4	-0.002	0.003	1.168	-1.232	0.285	-0.214	4.089

Model 1: ARFIMA(0,d,1)-GARCH(1,1), Model 2: ARFIMA(1,d,1)-GARCH(1,1), Model 3: HAR-RV-GARCH(1,1), Model 4: HAR-RV-GARCH(0,1).

Table 7. The average return, $\bar{r}^{(s)}$, of the trading strategy based on the model m with $\min\left(\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^{2(m)}\right)$.

	$\tilde{T}=10$ $s=1$	$\tilde{T}=10$ $s=5$	$\tilde{T}=10$ $s=10$	$\tilde{T}=10$ $s=15$	$\tilde{T}=22$ $s=1$	$\tilde{T}=22$ $s=5$	$\tilde{T}=22$ $s=10$	$\tilde{T}=22$ $s=15$
CAC 40	11.84%	13.37%	14.86%	15.95%	11.87%	13.36%	14.79%	16.27%
DAX 30	16.98%	17.80%	20.38%	19.84%	17.01%	17.90%	20.77%	20.45%
FTSE 100	14.60%	15.66%	17.21%	18.85%	14.57%	15.45%	17.18%	18.64%
EURUSD	18.96%	18.52%	18.86%	18.62%	18.97%	18.59%	18.72%	18.70%
EURGBP	13.94%	14.84%	14.60%	15.25%	13.95%	14.86%	14.75%	15.64%
EURJPY	15.35%	16.56%	17.75%	18.79%	15.35%	16.78%	17.89%	18.41%

At each trading day t, the trader takes into consideration the s-trading days ahead forecast, $RV_{t+s|t}^{(\tau)}$, estimated by the model m with $\min\left(\frac{1}{2}\sum_{t=1}^{\tilde{T}} z_{t+1|t}^{2(m)}\right)$. The trader proceeds to a long position when $RV_{t+s|t}^{(\tau)} > RV_{t}^{(\tau)}$ and to a short position when $RV_{t+s|t}^{(\tau)} < RV_{t}^{(\tau)}$.

Table 8. The average return, $\bar{r}^{(s)}$, of the trading strategy based on the model m with $\min\left(\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^{2(m)}\right)$.

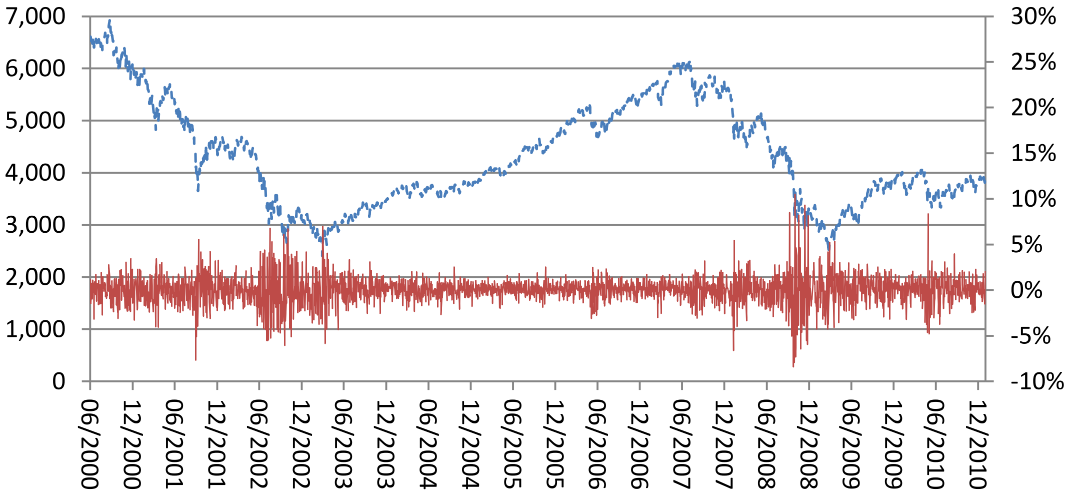
	$\tilde{T}=10$ $s=1$	$\tilde{T}=10$ $s=5$	$\tilde{T}=10$ $s=10$	$\tilde{T}=10$ $s=15$	$\tilde{T}=22$ $s=1$	$\tilde{T}=22$ $s=5$	$\tilde{T}=22$ $s=10$	$\tilde{T}=22$ $s=15$
CAC 40	11.68%	13.27%	14.72%	15.81%	11.75%	13.17%	14.72%	15.79%
DAX 30	16.78%	17.54%	20.09%	19.67%	17.00%	17.64%	20.66%	20.35%
FTSE 100	14.55%	15.63%	17.06%	18.79%	14.45%	15.44%	17.11%	18.52%
EURUSD	18.86%	18.45%	18.78%	18.61%	18.87%	18.49%	18.76%	18.78%
EURGBP	13.86%	14.56%	14.63%	15.37%	14.03%	15.11%	14.56%	15.62%
EURJPY	15.14%	16.59%	17.74%	18.72%	15.35%	16.70%	17.88%	18.47%

At each trading day t, the trader takes into consideration the s-trading days ahead forecast, $RV_{t+s|t}^{(\tau)}$, estimated by the model m with $\min\left(\sum_{t=1}^{\tilde{T}} \varepsilon_{t+1|t}^{2(m)}\right)$. The trader proceeds to a long position when $RV_{t+s|t}^{(\tau)} > RV_{t}^{(\tau)}$ and to a short position when $RV_{t+s|t}^{(\tau)} < RV_{t}^{(\tau)}$.

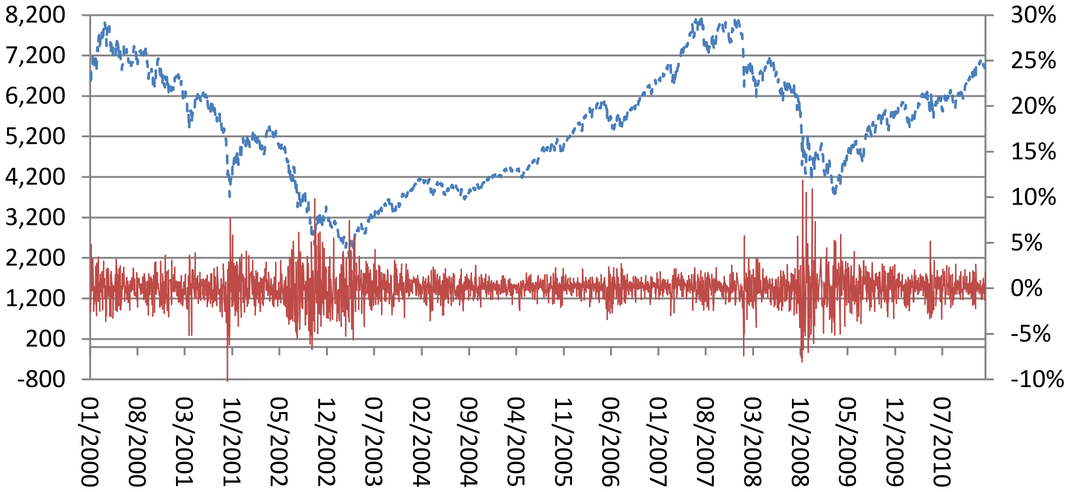
Figures

Figure 1. The daily prices, P_{t_τ} , along with the log-returns, $y_t = \left(\log P_{t_\tau} - \log P_{t-1_\tau}\right)$.

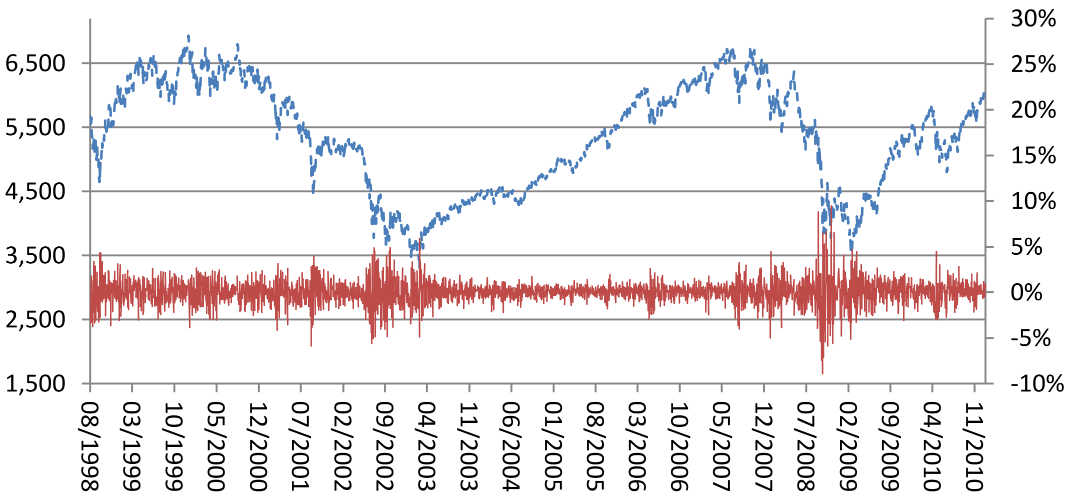
CAC 40 (13th June 2000 to 12th January 2011)



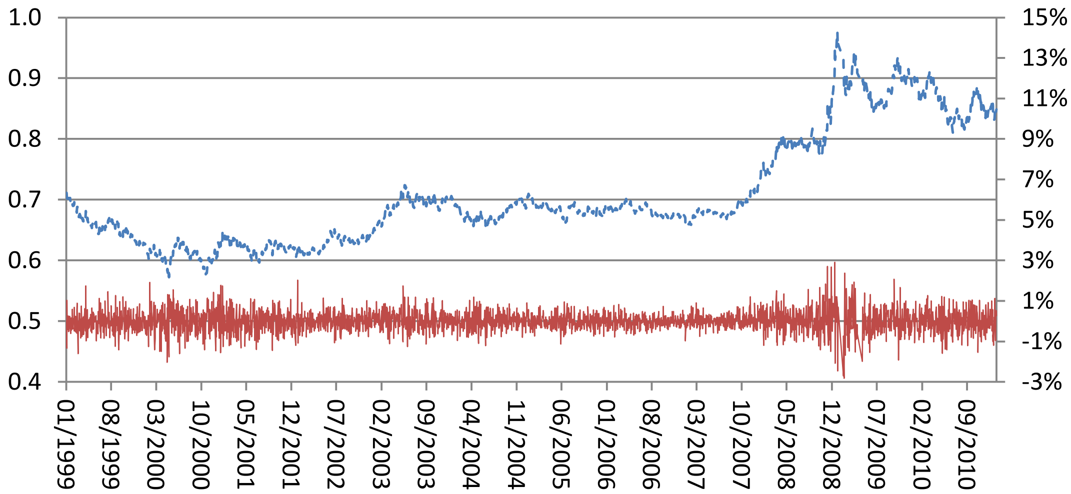
DAX 30 (3rd January 2000 to 12th January 2011)



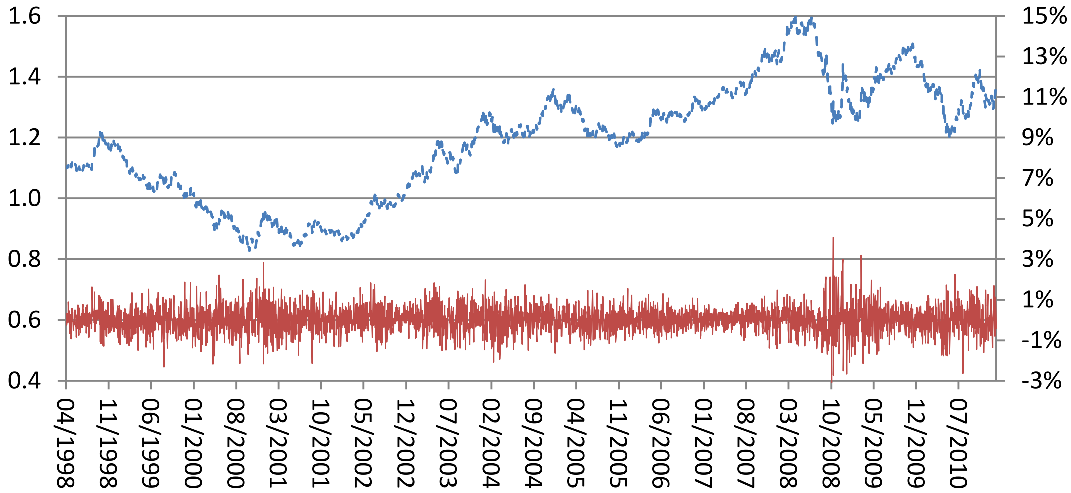
FTSE100 (20th August 1998 to 12th January 2011)



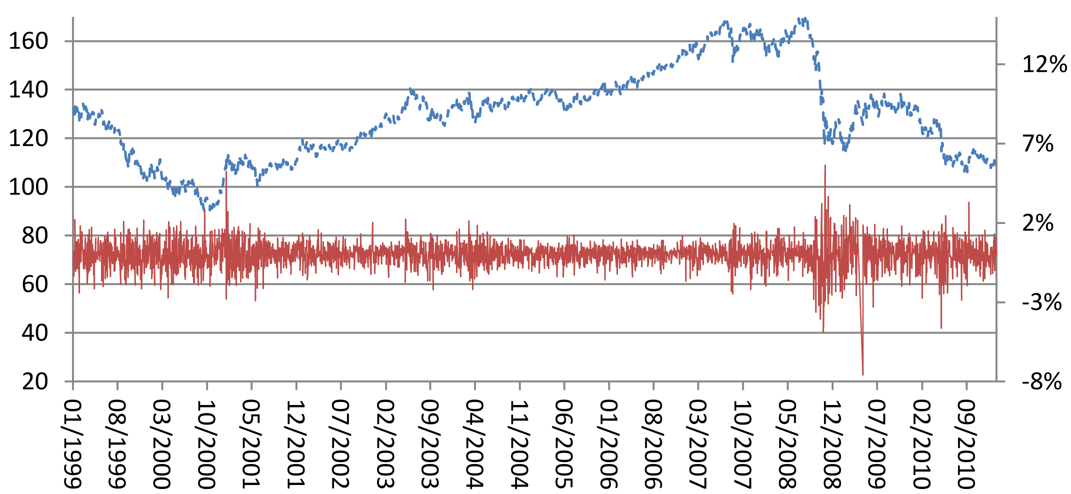
Euro to British Pound exchange rate (4th January 1999 to 21st January 2011)



Euro to United States Dollar exchange rate (20th April 1998 to 24th January 2011)



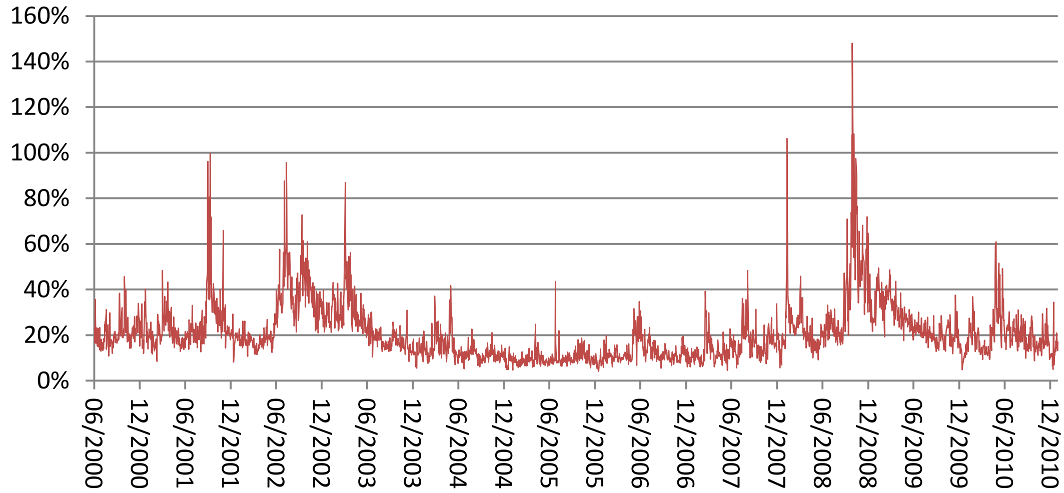
Euro to Japanese Yen exchange rate (4th January 1999 to 24th January 2011)



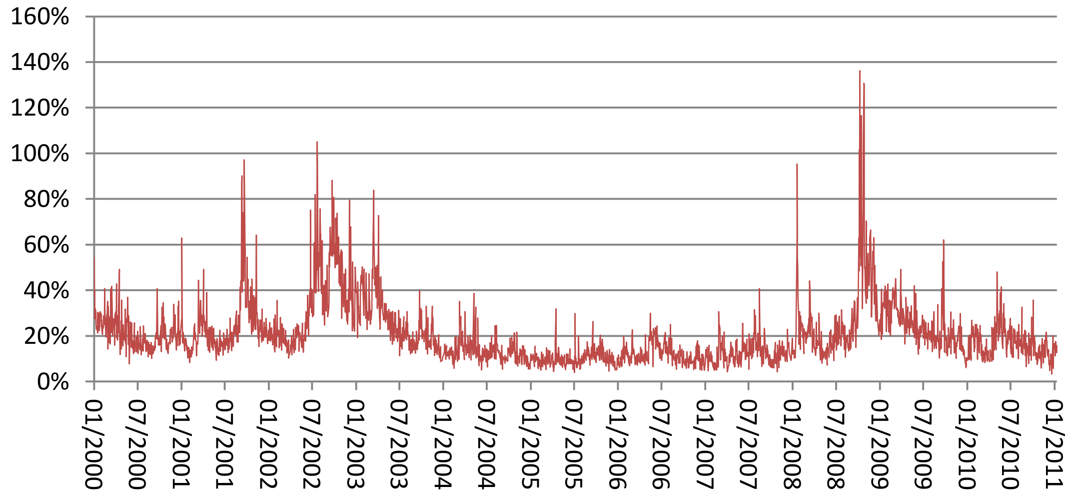
* The figures in the left column present the daily prices (dash line presented in the LHS axis) and the log-returns (solid line presented in the RHS axis).

Figure 2. The annualized adjusted realized volatilities, $\sqrt{252RV_t^{(\tau)}}$ at the optimal sampling frequencies

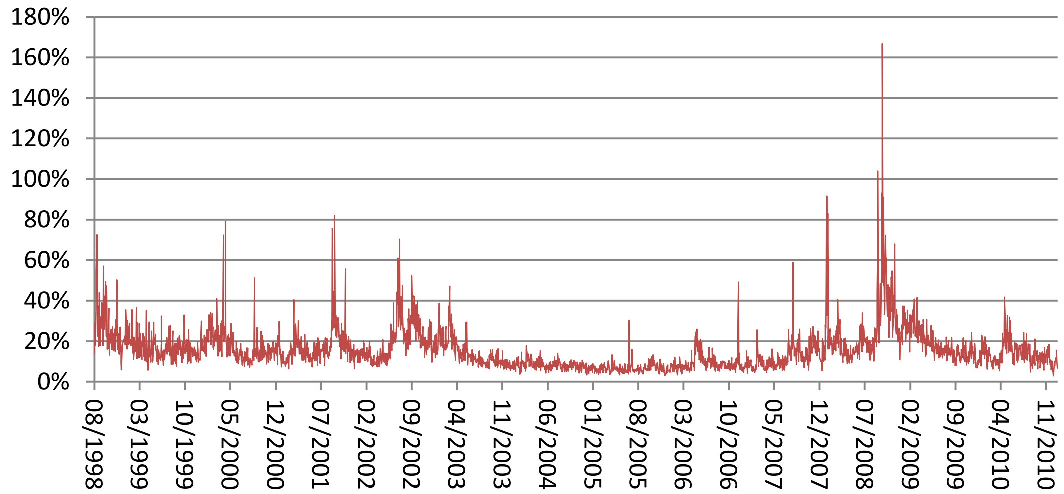
CAC 40 (13th June 2000 to 12th January 2011)



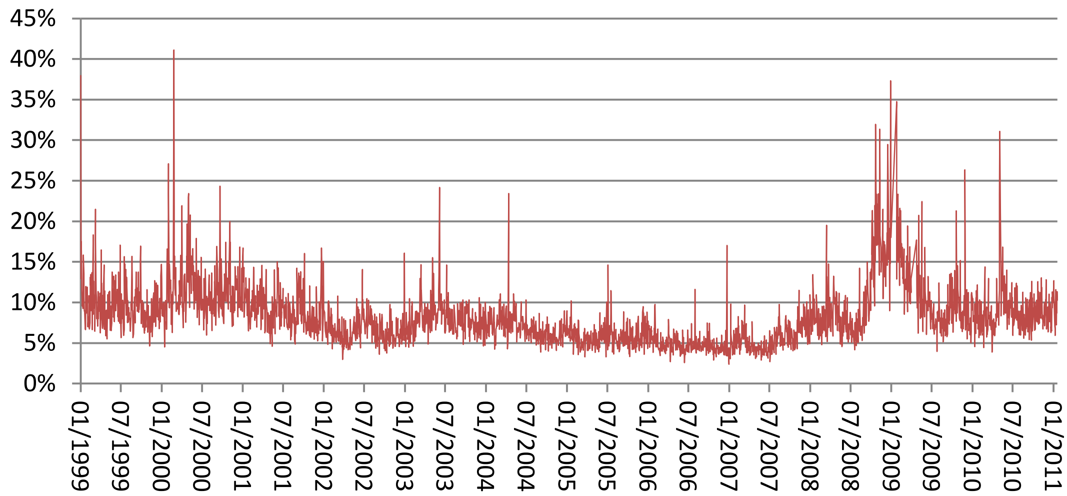
DAX 30 (3rd January 2000 to 12th January 2011)



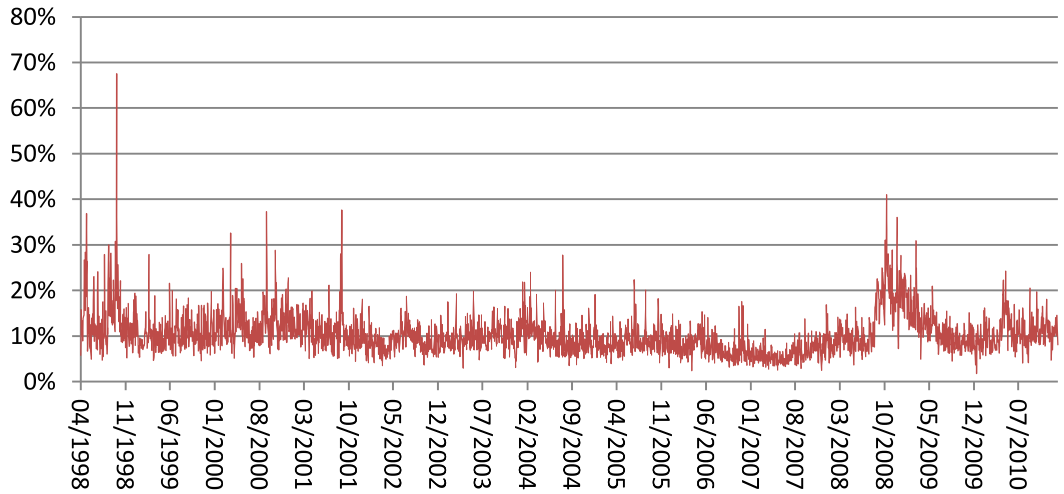
FTSE100 (20th August 1998 to 12th January 2011)



Euro to British Pound exchange rate (4th January 1999 to 21st January 2011)



Euro to United States Dollar exchange rate (20th April 1998 to 24th January 2011)



Euro to Japanese Yen exchange rate (4th January 1999 to 24th January 2011)

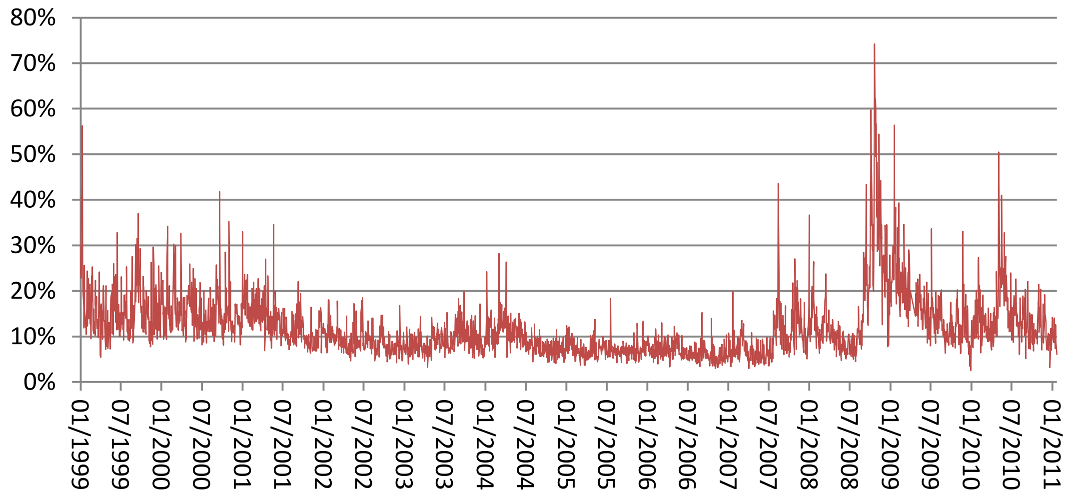
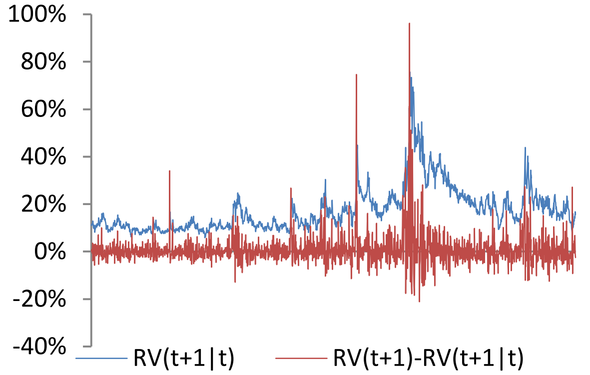


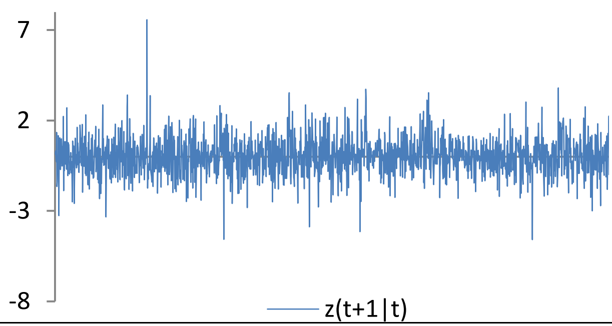
Figure 3. The $\sqrt{252RV^{(\tau)}_{t+1|t}}$, the discrepancy between $\sqrt{252RV^{(\tau)}_{t+1}}$ and $\sqrt{252RV^{(\tau)}_{t+1|t}}$ and the standardized forecast errors, $z_{t+1\backslash t}$, for the ARFIMA(1,d,1)-GARCH(1,1) model.

CAC40
ARFIMA(1,d,1)-GARCH(1,1)
for $z_t \sim N(0,1)$

$\sqrt{252RV^{(\tau)}_{t+1|t}}$ and $\left(\sqrt{252RV^{(\tau)}_{t+1}}-\sqrt{252RV^{(\tau)}_{t+1|t}}\right)$

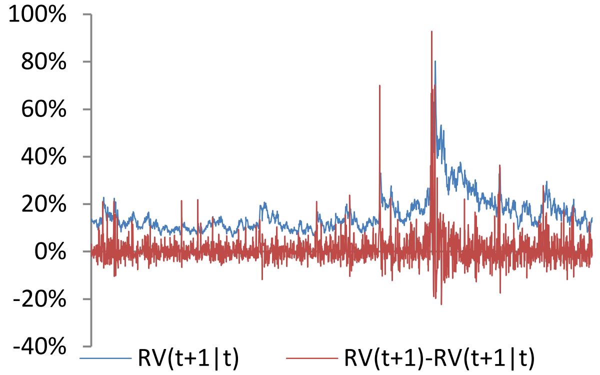


$z_{t+1\backslash t}$

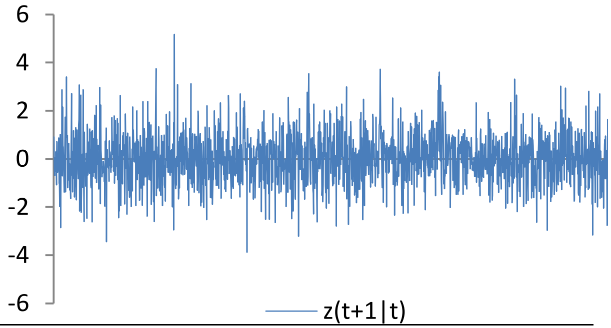


DAX30
ARFIMA(1,d,1)-GARCH(1,1)
for $z_t \sim Ged(0,1;v)$

$\sqrt{252RV^{(\tau)}_{t+1|t}}$ and $\left(\sqrt{252RV^{(\tau)}_{t+1}}-\sqrt{252RV^{(\tau)}_{t+1|t}}\right)$

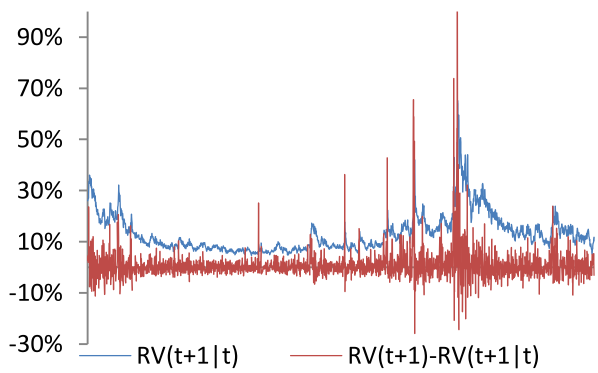


$z_{t+1\backslash t}$



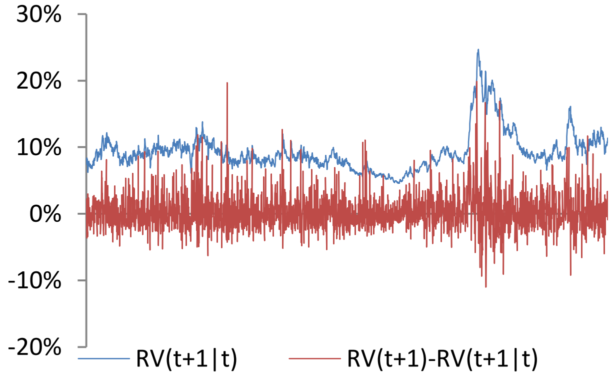
FTSE 100
ARFIMA(1,d,1)-GARCH(1,1)
for $z_t \sim N(0,1)$

$\sqrt{252RV^{(\tau)}_{t+1|t}}$ and $\left(\sqrt{252RV^{(\tau)}_{t+1}}-\sqrt{252RV^{(\tau)}_{t+1|t}}\right)$

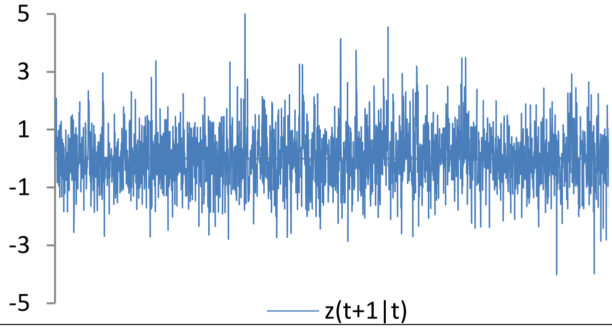


EURUSD
ARFIMA(1,d,1)-GARCH(1,1)
for $z_t \sim skT(0,1;v,g)$

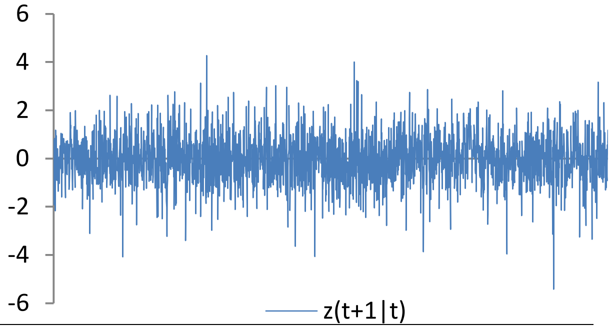
$\sqrt{252RV^{(\tau)}_{t+1|t}}$ and $\left(\sqrt{252RV^{(\tau)}_{t+1}}-\sqrt{252RV^{(\tau)}_{t+1|t}}\right)$



$z_{t+1\backslash t}$



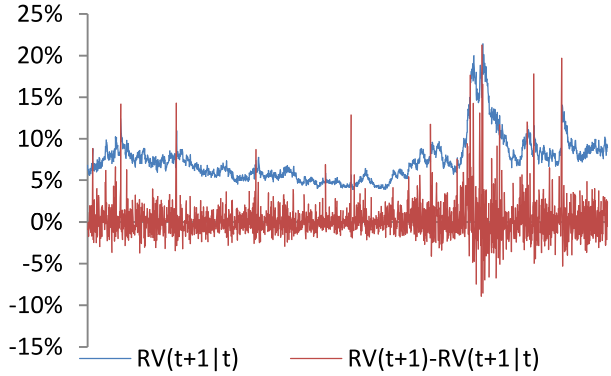
$z_{t+1\backslash t}$



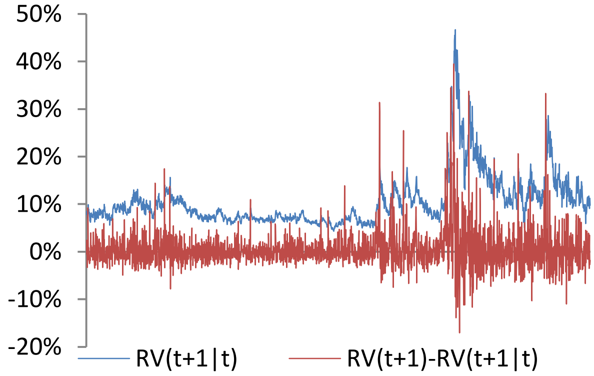
EURGBP
ARFIMA(1,d,1)-GARCH(1,1)
for $z_t \sim t(0,1;v)$

EURJPY
ARFIMA(1,d,1)-GARCH(1,1)
for $z_t \sim N(0,1)$

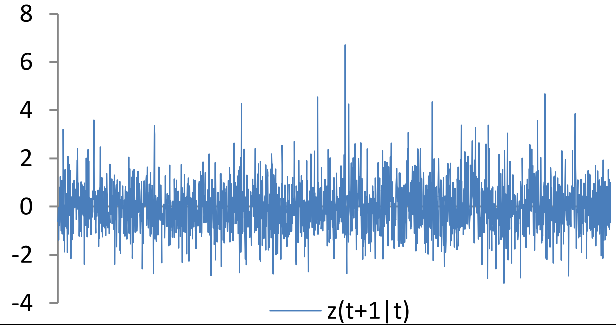
$\sqrt{252RV_{t+1|t}^{(\tau)}}$ and $\left(\sqrt{252RV_{t+1}^{(\tau)}}-\sqrt{252RV_{t+1|t}^{(\tau)}}\right)$



$\sqrt{252RV_{t+1|t}^{(\tau)}}$ and $\left(\sqrt{252RV_{t+1}^{(\tau)}}-\sqrt{252RV_{t+1|t}^{(\tau)}}\right)$



$z_{t+1\backslash t}$



$z_{t+1\backslash t}$