# On Forecasting Daily Stock Volatility: the Role of Intraday Information and Market Conditions

Ana-Maria Fuertes[a,*], Marwan Izzeldin[b], Elena Kalotychou[a]

[a] *Faculty of Finance, Cass Business School*

[b] *Department of Economics, Lancaster University Management School*

First draft - October 2007. This version - June 2008.

**Abstract**

Several recent studies advocate the use of nonparametric estimators of daily price variability that exploit intraday information. This paper compares four such estimators, *realised volatility*, *realised range*, *realised power variation* and *realised bipower variation*, by examining their in-sample distributional properties and out-of-sample forecast ranking when the object of interest is the conventional conditional variance. The analysis is based on a 7-year sample of transaction prices for 14 NYSE stocks. The forecast race is conducted in a GARCH framework and relies on several loss functions. The realized range fares relatively well in the in-sample fit analysis, for instance, regarding the extent to which it brings normality in returns. However, overall the realised power variation provides the most accurate 1-day-ahead forecasts. Forecast combination of all four intraday measures produces the smallest forecast errors in about half of the sampled stocks. A market conditions analysis reveals that the additional use of intraday data on day $t-1$ to forecast volatility on day $t$ is most advantageous when day $t$ is a low volume or an up-market day. The results have implications for value-at-risk analysis.

**Keywords**: C53; C32; C14.

**JEL Classification**: Conditional variance; Quadratic variation; Nonparametric estimators; Intraday prices; Superior predictive ability.

* *Corresponding author*. e-mail: a.fuertes@city.ac.uk., T: +44 (0)20 7040 0186; F: +44 (0)20 7040 8881, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ. We thank Valentina Corradi, Costas Milas, Andrew Patton, and participants at the 3rd ESRC Workshop on Nonlinear Economics and Finance (February 2008) at Keele University, the 2008 BAA Annual Conference in Blackpool, the 2008 FMA European Conference in Prague, the 2008 CFE Workshop in Neuchatel, Switzerland, the 2008 Annual International Symposium on Forecasting in Nice, and seminar participants at City University, Department of Economics, March 2008, Essex University, School of Accounting, Finance and Management, December 2007, and Cass Business School, October 2007, for useful comments. We acknowledge Anthony Murphy for help with processing the intraday prices.

# 1 Introduction

Over the past decade there has been an enormous interest among academics and practitioners in modeling and forecasting the conditional variance of stock market returns. Volatility is a crucial concept for portfolio management, option pricing and financial market regulation, inter alios. One problematic issue is that, unlike prices or returns, the volatility process is unobserved even *ex post*. In a seminal paper, Andersen and Bollerslev (1998) focus on the problem of how the choice of proxy for the latent population measure of volatility can affect the quantitative assessment of volatility forecasting models. They illustrate that if the squared daily returns are used as proxy for the day's variance in the forecast evaluation, GARCH models do have very poor forecasting properties whereas using the sum of intraday squared returns which employs more information, the GARCH forecasts turn out to be far more accurate.[1] The rationale behind this is that the squared return is an extremely noisy (albeit unbiased) estimator of *ex post* volatility.

One related empirical question is how to obtain better daily volatility forecasts using intraday data. The literature branches in two broad directions. Several studies extend the daily GARCH model to incorporate the intraday information as an additional regressor. Instances include as augmentation variable the daily high-low price range (Parkinson, 1980; Taylor, 1987), the number of intraday price changes (Laux and Ng, 1993), daily trading volume (Bessembinder and Seguin, 1993), and the standard deviation of intraday returns (Taylor and Xu, 1997). Another group of studies focus on different ways of modeling directly the intraday data as a way of providing better out-of-sample forecasts of daily volatility. Two instances are Martens (2001) who models the intraday returns directly using GARCH models and Koopman et al. (2005) who compare daily GARCH models with ARFIMA and Unobserved Components models fitted to a *realised volatility*

[1] A risk measure developed in recent years is Value-at-Risk (VaR), a quantile of the conditional distribution of returns given past information, which gives the worst expected loss. According to the 2008 *European Investment Practices Survey*, the majority of asset managers use parametric VaR approaches derived from traditional location-scale models such as ARMA-GARCH which means that VaR calculations are based on forecasts of the old measure of investment risk, the conditional volatility. There is a close relation between GARCH(1,1) and the exponentially-weighted-average (of squared returns) historical volatility advocated by J.P. Morgan's Riskmetrics for daily and monthly VaR estimation. If normality is assumed, VaR adds no extra information over the old volatility measure.

measure based on the sum of 5-minute squared returns.

Aside from forecasting issues, much emphasis has been given in recent years to the use of nonparametric estimators of daily volatility that exploit intraday prices. The theoretical properties of these estimators have been investigated using advanced and novel asymptotic theory in stochastics and econometrics. For instance, the *realised variance* has been thoroughly studied by Barndorff-Nielsen and Shephard (BN-S, 2002a, 2002b). The sum of intraday high-low price ranges or *realised range* has been scrutinized by Christensen and Podolskij (2005) and Martens and van Dijk (2006). Two other intraday volatility estimators, introduced by BN-S (2004a, 2004b), are the *realised power variation*, based on summing powers of the intraday absolute returns, and the *realised bipower variation*, the sum of products of consecutive intraday absolute returns.

The paper complements the literature in several directions. First, it investigates the relative merit of the above nonparametric (intraday) volatility estimators from two perspectives. On the one hand, their in-sample distributional properties are compared, for instance, by gauging their efficiency, persistence and whether they can normalize the daily returns. In the Mixture of Distributions Hypothesis (MDH) literature, which builds on the tenet that volatility and trading volume are jointly driven by the latent information flow, the performance of volatility measures is typically assessed by the extent to which they bring return normality. On the other hand, we compare their ability to enhance the out-of-sample daily GARCH forecasts. For completeness, another updating variable is the daily volume computed by summing the number of shares traded over all intraday intervals. Following a large body of recent literature, the forecast race is based on 1-day-ahead predictions and the latent conditional variance is proxied by the 5-min realised variance.[2]

Second, given that market microstructure issues bedevil the above intraday volatility estimators in different ways, the paper addresses the question of whether forecast combining is fruitful. For this purpose, a rolling-window approach is adopted that allows for time-varying combination weights.

[2] Although regulators and fund managers might be mostly interested in longer horizons, derivative traders are interested in daily losses. One-day-ahead volatility forecasts are, for instance, relevant for VaR measurement since banks may wish to update their estimates of potential loss on a daily basis to determine capital requirements.

Third, an important question that has not been addressed as yet is whether the importance of updating daily conditional volatility (GARCH) models with intraday data depends on market conditions. This paper compares the forecast value of the four intraday volatility estimators during up- versus down-market days, and low- versus high-volume days. Finally, the study contributes to the existing literature by analysing 14 individual NYSE stocks whereas most related studies focus on FX data or stock market indices.

The sample spans 7 years of trading data over 02/01/97 to 31/12/03. The statistical properties of four nonparametric volatility estimators of daily price variation alongside the squared returns and the GARCH volatility are assessed according to criteria motivated by the MDH. The realised power variation and realised range are the top performers. Next, rolling out-of-sample forecasts are generated with a GARCH model augmented with either of the lagged nonparametric volatility measures or volume. Different forecast accuracy criteria are used which include asymmetric loss functions and the Mincer-Zarnowitz levels regression. Pairwise comparisons of forecast accuracy are conducted via the Diebold-Mariano (1995) test in the case of non-nested models and by the Harvey et al. (1998) encompassing test for nested models. The results reveal significant forecast gains from using intraday price information but not trading volume. GARCH updated with realised power variation is in the lead, followed closely by the realised range, and with realised bipower variation at the other extreme. Finally, a joint forecast comparison is performed using Hansen (2005) superior predictive ability test. For most stocks, the realised power variation is not beaten by any of the alterantive models. Combining the predictive information of the competing GARCH-augmented models is worthwhile. Finally, exploiting intraday returns at $t-1$ to forecast next day volatility is most fruitful when $t-1$ is a low-volume or up-market day.

The rest of the paper is organized as follows. Section 2 provides a review of the large literature on intraday volatility measuring which is by no means exhaustive. Section 3 presents the variable definitions and forecasting framework. Section 4 discusses the results and Section 5 concludes.

## 2 Background literature

The GARCH modeling framework introduced by Engle (1982) is still widely used to analyse the dynamics of daily return variation in all areas of finance by academics and practitioners alike. Several studies have documented that out-of sample regressions of squared returns on GARCH forecasts produce low $R^2$ statistics below 10% (see, inter alios, Franses and van Dijk, 1996, and Brooks, 1998). However, Andersen and Bollerslev (1998) for FX data and Blair et al. (2001) for stock indexes show that GARCH forecasts, when compared with the sum of intraday squared returns as the conditional volatility proxy are far more accurate with an $R^2$ of about 50%.

A weakness of GARCH models though is that the future variance of returns is cast as a polynomial of current and past squared returns. If on day $t-1$ the return is zero, the squared return at $t-1$ will also equal zero ignoring any within-day price fluctuations. One way forward is to augment the GARCH equation with variables that carry predictive power for future volatility. Lamoureux and Lastrapes (1990), Najand and Yung (1991), and Bessembinder and Seguin (1993) include contemporaneous volume in GARCH models and document an improvement in the in-sample fit. A problem with this approach is that volume cannot be assumed to be exogenous since, according to the MDH, volume and volatility are simultaneously influenced by the latent information arrival process. Brooks (1998) and Donaldson and Kamstra (2004) show that augmenting GARCH with lagged volume leads to no improvement in forecast performance. However, Donaldson and Kamstra (2004) show for the S&P100 index that trading volume has a switching role in forecasting. If volume on day $t-1$ is low relative to the recent past, then one-day-ahead ARCH forecasts are at least as effective as option implied volatilities (VIX). Conversely, if volume at $t-1$ is high, the best volatility forecast for day $t$ can be obtained by placing more weight on market expectations.

With the increasing availability of high frequency data the research focus has shifted towards exploiting nonparametric estimators of daily volatility based on intraday returns. A large number of papers advocate the *realised variance* (RV) for the modeling and prediction of volatility of

FX returns (Taylor and Xu, 1997; Andersen and Bollerslev, 1998) and equity returns (Andersen, Bollerslev, Diebold and Ebens, ABDE, 2001). Luu and Martens (2002) find support for the assumptions underlying the MDH model when RV is used instead of daily squared returns. Pong et al. (2004) compare the forecasting ability of short memory (ARMA) and long memory (ARFIMA) models of RV, and the implied volatilities from OTC foreign currency options for horizons ranging from one day to 3 months. The models provide more accurate forecasts than the implied volatilities for short (one-day and one-week) horizons and this is attributed to the use of intraday returns rather than to the long memory specification. At the one- and 3-month horizons, the models of RV do not provide incremental information that is not already incorporated in the implied volatilities. Using an equity price index and two currencies, Galbraith and Kisinbay (2002) find that forecasts from AR fitted to daily RV outperform the forecasts from GARCH for a 1-day horizon whereas at 30 days the two methods become indistinguishable.

For the S&P100 index, Koopman et al. (2005) generate one-day-ahead forecasts from ARFIMA and Unobserved Components models fitted to RV, and from stochastic volatility (SV) and GARCH models fitted to daily returns and augmented with lagged RV and implied volatility.[3] Long memory models seem to provide the most accurate forecasts. Engle and Gallo (2006) develop a multiplicative-error model which combines several daily volatility indicators (absolute returns, squared high-low range and RV) and show that it forecasts quite well 1-month-ahead the VIX.

A second group of empirical studies advocate different nonparametric volatility estimators as an alternative or complement to the popular RV. Ghysels et al. (2006) introduce the MIDAS (MIxed DAta Sampling) regression approach and compare several daily volatility estimators based on FX data sampled at different intraday frequencies. They find that realised power variation (RPV) outperforms RV and that (intra-)daily absolute returns outperform, respectively, the corresponding squared returns. Using Yen/US$ and DM/US$ rates and the Spyder Exchange-Trade Fund that represents ownership in the S&P500 index, Liu and Maheu (2005) fit HAR-log and ARFIMA

[3] In the ARCH class of models, the expected volatility is parameterized as a function of past returns only. In contrast, the parameterized expectations in the SV class of models explicitly rely on latent state variables.

models to RV and augment them with lagged RPV and realized bipower variation. Only with RPV they find robust improvements in the 1-day-ahead forecasts of FX rates and the S&P500. For DM/US$ rates, the S&P500 and the 30-year US T-bond yield, Andersen et al. (2007) document that only the continuous part of the return process carries predictive power for future volatility.

# 3 Methodology

## 3.1 Population measures of volatility

In most of the volatility forecasting literature, the population measure of volatility is the *conditional variance.* Let $r_t$ denote the daily stock return, its conditional variance is denoted $var(r_t|\mathcal{F}_{t-1}) \equiv \sigma_t^2$ where $\mathcal{F}_{t-1}$ is the sigma field containing all relevant information up to time $t-1$, which naturally refers to $r_{t-j}, j > 1$ but it may also include other variables. It is assumed that $E(r_t|F_{t-1}) \equiv E_{t-1}(r_t) = 0$ such that $\sigma_t^2 = E_{t-1}(r_t^2)$ is the object of interest.

But there are other possible population measures of variance. To define them, let the price process belong to the class of semimartingales with jumps. The dynamics of the log price change in continuous time can be characterized by the stochastic differential equation

$$dp(t) = \mu(t)dt + s(t)dW(t) + k(t)dq(t) \qquad 0 \leq t \leq T \tag{1}$$

where $\mu(t)$ denotes the drift term, $s(t)$ is the instantaneous or spot volatility process which is assumed to be stationary and independent of the standard Brownian motion $W(t)$, $dq(t)$ is a counting process with $dq(t) = 1$ if a jump occurs at time $t$ and $k(t)$ is the jump size. Equation (1) embodies the intuitive idea that there are two types of randomness driving the stock returns. One is a Brownian motion generating the continuous sample path and small movements and the other consists of large but infrequent (discrete) jumps.

The *quadratic variation* (QV) or notional variability of the return process is defined as

$$QV_t = \int_{t-1}^{t} s^2(u)du + \sum_{t-1<j\leq t} k^2(j) = IV_t + J_t \tag{2}$$

where the first term is called *integrated variance* ($IV_t$) and corresponds to the continuous part of the log price process and the second term ($J_t$) reflects the contribution of the discrete jumps.[4] In a recent paper, BN-S (2004a) define the *integrated power variation* (IPV) of order $z$ as

$$IPV_t(z) = \int_{t-1}^{t} s^z(u)du, \ 0 < z \leq 2 \tag{3}$$

which for $z = 2$ becomes the integrated variance.

In this paper, the population measure of interest is the conditional variance due to its prevalence in applied and theoretical forecasting work in the past two decades. Andersen et al. (2002) argue instead in favour of QV as the relevant notion of variability. However, both population measures of volatility are closely related since the conditional variance of future returns is the conditional expectation of QV as shown in BN-S (2002b). This result provides further theoretical underpinning for the widespread use in empirical finance of GARCH models.

## 3.2 Daily GARCH models and intraday updating variables

Let the conditional mean and conditional variance of daily returns be captured, respectively, by and ARMA$(p, q)$ and GARCH$(r, s)$ equation[5]

$$r_t = \theta_0 + \sum_{i=1}^{p} \theta_i r_{t-i} + \sum_{j=1}^{q} \lambda_j u_{t-j} + u_t, u_t|\mathcal{F}_{t-1} \sim iid(0, h_t) \tag{4a}$$

$$h_t = \omega + \sum_{i=1}^{r} \alpha_i u_{t-1}^2 + \sum_{j=1}^{s} \beta_j h_{t-j} \tag{4b}$$

where $r_t$ are the daily returns and $u_t^2$ are the squared whitened returns. The lag orders of the (conditional) mean and variance equations will be appropriately selected so as to remove all the

[4] Most modern finance theory is based on semimartingales. If the return process is a semimartingale, then it has an associated $QV_t$ process. The latter plays a central role in the option pricing literature. In particular, in the absence of jumps, $QV_t$ equals the $IV_t$ highlighted in the stochastic volatility models first proposed by Hull and White (1987) as an alternative to the classical Black-Scholes formulae for option pricing. However, there is an increasing body of empirical work in finance which concludes that continuous-time models must incorporate jumps or discontinuities in order to provide a satisfactory characterization of the daily return process.

[5] We do not consider asymmetric GARCH models because the asymmetric relation between price movements and volatility (e.g. rationalized as 'leverage effect') has been shown to be rather weak or absent in individual stock price series as compared to broad stock price index series (see, for instance, Kim and Kon, 1994; Tauchen et al., 1996).

return autocorrelation and volatility clustering. The Ljung-Box and ARCH LM tests, respectively, will be used for these purposes. The degree of volatility persistence is given by $\lambda = \Sigma\alpha_i + \Sigma\beta_j$.

The selected GARCH model for each stock is then augmented as follows

$$h_t = \omega + \sum_{i=1}^{r} \alpha_i u_{t-1}^2 + \sum_{j=1}^{s} \beta_j h_{t-j} + \gamma v_{t-1} \tag{5}$$

where $v_{t-1}$ is a nonparametric estimator based on intraday prices at day $t-1$. In our case $v_{t-1}$ is the realised variance (RV), realised range (RR), realised power variation (RPV), realised bipower variation (BPV) or trading volume (VOL). For this purpose, the time dimension is discretized and the daily time interval is divided into $M$ equally-spaced subintervals of length $\delta$. The price at the start of the $j$th intraday interval is computed as the average of the closing and opening prices of intervals $j-1$ and $j$, respectively. The $j$th intraday return (on day $t$) is computed as

$$r_{t,j} = 100\left(\frac{\log(p_{t,j}^c) + \log(p_{t,j+1}^o)}{2} - \frac{\log(p_{t,j-1}^c) + \log(p_{t,j}^o)}{2}\right), j = 2, ..., M-1 \tag{6}$$

where each trading day [9:30am-4:00pm] amounts to a duration of $M \times \delta = 390$min, and $p_{t,j}^c$ ($p_{t,j}^o$) is the closing (opening) price of the $j$th intraday interval. For instance, $j = 2$ corresponds to 9:35am-9:40am. The extreme-interval returns are $r_{t,1} = 100\left(\frac{\log(p_{t,1}^c)+\log(p_{t,2}^o)}{2} - \log(p_{t,1}^o)\right)$ and $r_{t,M} = 100\left(\log(p_{t,M}^c) - \frac{\log(p_{t,M-1}^c)+\log(p_{t,M}^o)}{2}\right)$. For $\delta = 5\,\text{min}$, we have $M = 78$ intraday returns and one overnight return. However, a few trading days consist of $M < 78$ due to delayed openings and/or early closings of the NYSE. Overnight returns are not included due to the fact that the weight such a return should deserve is somewhat arbitrary as Hansen and Lunde (2006b) and Engle et al. (2006) argue. The usual logarithmic (or continuously compounded) daily returns used to estimate GARCH models amount to the aggregated intraday returns, $r_t = \sum_{j=1}^{M} r_{t,j} = \log(\frac{p_{t,M}^c}{p_{t,1}^o})$.

The most popular estimator, the *realised variance,* defined as the sum of intraday returns

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2, \quad t = 1, 2, ..., T \tag{7}$$

converges in probability to the quadratic variation ($RV_t \xrightarrow{p} QV_t$) under suitable conditions as the intraday sampling frequency increases ($M \to \infty$) and so RV is a consistent estimator of QV (see ABDE, 2001; BN-S, 2002a,b). If $M = 1$, then RV becomes the noisy daily squared return ($r_t^2$).

The *realised range* estimator introduced by Christensen and Podolskij (2005) is a generalization of the range estimator of Parkinson (1980) and defined as

$$RR_t = \frac{1}{4\log 2}\left[\sum_{j=1}^{M} 100 \times \left(\log(p_{t,j}^h) - \log(p_{t,j}^l)\right)^2\right] \quad t = 1, 2, ..., T \tag{8}$$

where $\log(p_{t,j}^h)$ and $\log(\log p_{t,j}^l)$ are the high and low prices of the $j$th interval, and the scaling factor $4\log 2$ is a bias-correction for market microstructure effects given by the second moment of the range of a Brownian motion $B_t$, that is, $E(s_B^2) = 4\log 2$ where $s_B = \sup_{0\leq t,s\leq 1}(B_t - B_s)$.

BN-S (2002a) and Christensen and Podolskij (2005) show that RR is a more efficient estimator than RV. In an ideal world without market frictions (no bid-ask bounce, discontinuous trading or jumps) the asymptotic variance of the RR estimator is $0.4\int_{t-1}^{t}\sigma(u)^4 du$, where the integral is called integrated quarticity, which is 5 times smaller than the variance of RV at $2\int_{t-1}^{t}\sigma(u)^4 du$. Hence, theoretically the RR estimator is more efficient than other variance estimators based on squared returns. Christensen and Podolskij (2005) and Martens and van Dijk (2006) show that, in the absence of jumps, as $M \to \infty$ the realised range converges in probability to the quadratic variation ($RR_t \underset{\to}{p} QV_t$). This results does not hold, however, in a jump-difussion setting; Theorem 1 in Christensen-Podolskij establishes that with jumps, RR is not a consistent estimator of QV. For a DGP without jumps, Martens and van Dijk (2006) accommodate the bid-ask bounce in Monte Carlo simulations to show that: *i*) both RR and RV are upward biased but the former suffers more; *ii*) infrequent trading induces a downward bias in RR but not in RV.

Another estimator introduced by BN-S (2004a), the *realised power variation* of order $z$, is

$$RPV_t(z) = \mu_z^{-1}\delta^{1-z/2}\sum_{j=1}^{M}|r_{t,j}|^z, \quad 0 < z < 2, \quad t = 1, 2, ..., T \tag{9}$$

where

$$\mu_z = E|\mu|^z = 2^{z/2}\frac{\Gamma(\frac{1}{2}(z+1)}{\Gamma(\frac{1}{2})}, \ \mu \sim N(0,1)$$

which for $z = 1$ becomes the realised absolute variation. BN-S (2004a) demonstrate its consistency by showing that as $M \to \infty$, it converges in probability to the integrated power variation, $RPV_t(z)$

$\xrightarrow{p} IPV_t(z)$, and so it is robust to jumps. Liu and Maheu (2005) study the 1-day-ahead forecasting properties of (9) for orders $z = \{0.25, 0.5, ..., 1.75\}$ and find that $0.5, 1$, and $1.5$ yield the lowest RMSE. Absolute return measures are more persistent than the squared counterparts so RPV could outperform RV in forecasting financial risk. Also RPV may provide better predictions than RV when the sample period contains large jumps. Further discussion on the RPV estimator can be found in Ghysels et al. (2006) and Forsberg and Ghysels (2007).

In a similar fashion, BN-S (2004a) define the *realised bipower variation* estimator as

$$RBP_t = \mu_1^{-2} \sum_{j=2}^{M} |r_{t,j}| \, |r_{t,j-1}| \tag{10}$$

where $\mu_1 = E(|\mu|) = \sqrt{2}/\sqrt{\pi} \simeq 0.79788$ and $\mu \sim N(0,1)$. BN-S (2004a) show that as RBP converges in probability to the integrated variance ($RBP_t \xrightarrow{p} IV_t$) and so it is also immune to jumps. The result that $RV_t - RBP_t \xrightarrow{p} J_t$ where $J_t$ is the jump component in (2) is exploited by BN-S (2006) alongside the joint asymptotic distribution of the two estimators (under the null of a continuous sample path) to develop a non-parametric test for jumps.

The asymptotics (as $M \to \infty$) of these nonparametric volatility estimators were derived under suitable theoretical conditions such as no market microstructure noise. Unfortunately, in real-world settings the semimartingale property of prices breaks down at ultra-high frequencies because the influence of market microstructure factors such as bid-ask bounce (Ross, 1984), screen fighting (Zhou, 1996), price discreteness and irregular trading become overwhelming. This means that, in practice, intraday measures of volatility calculated at very high frequencies become biased.[6]

As noted, we adopt $\delta = 5$ and the motivation for this choice is twofold. First, a 5-minute grid is short enough for the daily volatility dynamics to be picked up with reasonable accuracy and long enough for the adverse effects of market microstructure frictions not to be overwhelming.[7] Second,

[6]Several methods, mostly nonparametric, have been proposed to account for microstructure bias. Martens and Van Dijk (2006) suggest a bid-ask bias correction for the RR estimator, eq.(8), by scaling it with the ratio of the average level of the daily range and the average level of the RR over the previous $q$ trading days. Adding autocovariances to the RV estimator, eq.(7), has been suggested as a way of mitigating bid-ask bounce biases (Barndorf-Nielsen et al., 2004; Hansen and Lunde, 2006b). Jungbacker and Koopman (2005) develop a parametric model-based approach that accounts for microstructure noise and intra-daily seasonality.

[7]ABDE (2001), BN-S (2002a,b), Taylor and Xu (1997) and Fleming and Paye (2006), inter alios, advocate this

it will enable meaningful comparisons with previous studies, most of which are based on 5-min data. Nevertheless, we check how sensitive the main results are to using 15- and 30-min grids.

Finally, the daily volume ($VOL$) measure adopted is the total number of shares traded each day computed as $VOL_t = \sum_{j=1}^{M} vol_{t,j}$, where $vol_{t,j}$ is the number of shares traded over the $j$th interval. This is the measure of volume used in Lamoureux and Lastrapes (1990).

### 3.3 Forecast evaluation and market conditions

The sample size is divided into an estimation period ($T - T_1$) of fixed length 1261 days, and a holdout or evaluation period ($T_1$) of 500 days. Hence, each model is estimated over an initial window, denoted $[1, t]$, and a 1-day-ahead ex post volatility forecast is generated. The window is rolled forward one day to $[2, t+1]$ to obtain the second forecast and so forth until 500 iterations.

The population volatility measure ($\sigma_t^2$) is the *conditional variance* and its proxy ($\tilde{\sigma}_t^2$) for forecast evaluation is the 5-min realized variance because it is an asymptotically conditionally unbiased estimator of the conditional variance — a further appealing property of the realized variance is that it converges in probability to the QV which plays a central role in the option pricing literature. The precision of model $m$ forecasts, $\{h_{t,m}\}_{t=1}^{T_1}$, is gauged through several loss functions:

| | |
|---|---|
| *Mean absolute error* | $MAE = \frac{1}{T_1}\sum_{t=1}^{T_1}\left\lvert\tilde{\sigma}_t^2 - h_{t,m}\right\rvert$ |
| *Mean squared error* | $MSE = \frac{1}{T_1}\sum_{t=1}^{T_1}(\tilde{\sigma}_t^2 - h_{t,m})^2$ |
| *Heteroskedasticity-adjusted MAE* | $HMAE = \frac{1}{T_1}\sum_{t=1}^{T_1}\left\lvert 1 - \tilde{\sigma}_t^{-2}h_{t,m}\right\rvert$ |
| *Heteroskedasticity-adjusted MSE* | $HMSE = \frac{1}{T_1}\sum_{t=1}^{T_1}(1 - \tilde{\sigma}_t^{-2}h_{t,m})^2$ |
| *Adjusted mean absolute percentage error* | $AMAPE = \frac{1}{T_1}\sum_{t=1}^{T_1}\left\lvert\frac{\tilde{\sigma}_t^2 - h_{t,m}}{\tilde{\sigma}_t^2 + h_{t,m}}\right\rvert$ |
| *Theil-U* | $Theil\text{-}U = \sum_{t=1}^{T_1}(\tilde{\sigma}_t^2 - h_{t,m})^2 / \sum_{t=1}^{T_1}(\tilde{\sigma}_t^2 - h_{t,N}^2)^2$ |
| *Mean mixed error (U)* | $MME(U) = \frac{1}{\#U}\sum I_U \cdot e_{t,m}^2 + \frac{1}{\#O}\sum I_O \cdot \lvert e_{t,m}\rvert$ |
| *Mean mixed error (O)* | $MME(O) = \frac{1}{\#U}\sum I_U \cdot \lvert e_{t,m}\rvert + \frac{1}{\#O}\sum I_O \cdot e_{t,m}^2$ |
| *Logarithmic loss* | $LL = \frac{1}{T_1}\sum_{t=1}^{T_1}(\ln\tilde{\sigma}_t^2 - \ln h_{t,m})^2$ |
| *Gaussian maximum likelihood error* | $GMLE = \frac{1}{T_1}\sum_{t=1}^{T_1}\left(\ln h_{t,m} + \tilde{\sigma}_t^2 h_{t,m}^{-1}\right)$ |

In the MME(U) and MME(O) criteria, $e_{t,m} = \tilde{\sigma}_t^2 - h_{t,m}$ denotes the forecast error for model $m$. $\#U$ is the number of underpredictions and $I_U = 1$ if $e_{t,m} < 0$; likewise for $\#O$ and $I_O$.

grid also because daily returns standardized by 5-min realised volatility are approximately normal. In the forecasting literature, studies that use 1-, 5-, 15- and 30-min data report mixed results but overall they also tend to favour the 5-min sampling (Martens and van Dijk, 2006; Pong et al., 2004; Ghysels et al., 2006; Galbraith and Kisinbay, 2002).

MAE, MSE, HMAE, HMSE, AMAPE and Theil-U belong to the family of symmetric loss functions, in the sense that they equally penalize over- and under-predictions. The most widely adopted, MSE, proposed by Bollerslev et al. (1994) is based on a quadratic loss function and so it is particularly good where large forecast errors are disproportionately more worrisome than smaller errors. MAE is less sensitive to severe mispredictions than MSE whereas AMAPE, proposed by Makridakis (1993), is an interesting alternative in percentage. The heteroskedasticity-adjusted version of MSE and MAE, introduced by Bollerslev and Ghysels (1996), is used by Martens (2001) and Koopman et al. (2005) inter alios.[8] Theil-U is calculated as the ratio of MSE for the model at hand to the MSE of the naive model, typically a random-walk type model, $h_{t,N} = \tilde{\sigma}^2_{t-1}$.

A number of asymmetric loss functions have been employed in the volatility literature. Examples include the two mean mixed error statistics proposed by Brailsford and Faff (1996), MME(U) and MME(O), the logarithmic loss (LL) introduced by Pagan and Schwert (1990) and the Gaussian maximum likelihood error (GMLE) of Bollerslev et al. (1994) which corresponds to the loss function implied by a Gaussian likelihood. MME(U), LL and GMLE penalize under-predictions more heavily than over-predictions whereas MME(O) does the opposite. For instance, in option pricing it is well established that the higher the volatility the higher the value of the call option so the under-prediction (overprediction) of volatility is unattractive for the seller (buyer). In addition, we also utilize the $R^2$ of Mincer-Zarnowitz level regressions (MZ-$R^2$), also called unbiasedness-regressions in the literature, a measure of the informational content of the volatility forecasts.[9,10]

[8] The HMAE can also be referred to as mean absolute percentage error (MAPE) since it can be rewritten as $HMAE = \frac{1}{T_1}\sum_{t=1}^{T_1}\left|\frac{\sigma_t^2 - \hat{\sigma}^2_{t,m}}{\sigma_t^2}\right|$. Likewise, the HMSE is the mean squared percentage error (MSPE).

[9] The MZ levels regression is $\tilde{\sigma}_t^2 = a + bh_{t,m} + e_t, t = 1, ..., T_1$. Hence, $h_t$ will be unbiased for the true variance $\sigma_t^2$ if $a = 0$, $b = 1$ and $E(e_t) = 0$. The $R^2$ from this regression (called MZ-$R^2$) reflects the variance but not the bias-squared component of MSE, that is, it corrects for bias.

[10] Hansen and Lunde (2006a) study the distortion in model ranking from replacing $E[L(\sigma_t^2, h_{t,m}]$ by $E[L(\tilde{\sigma}_t^2, h_{t,m}]$. It is called *objective bias* to distinguish it from the *sampling error,* due to estimating $E[L(\tilde{\sigma}_t^2, h_{t,m}]$ by a sample average, which vanishes as $T_1$ increases. Taking the conditional variance as the latent volatility, $\sigma_t^2$, they derive a set of conditions under which the empirical model ranking obtained under $L(\tilde{\sigma}_t^2, h_{t,m})$ is consistent for the true model ranking under $L(\sigma_t^2, h_{t,m})$. The conditions are met by MSE, GMLE and the levels *MZ-*$R^2$. For a range of loss functions, Patton (2006) derives analytically the objective bias using the daily high-low range, daily square return and 5- and 30-min RV as conditional volatility proxies. They illustrate that the more precise the proxy, the less relevant the objective bias; in particular, the objective-bias, if any, is very small for the 5-min RV proxy.

The statistical significance of differences in forecast accuracy is assessed by means of the Diebold-Mariano (1995) test statistic for non-nested models and the Harvey et al. (1998) encompassing test for nested models. In the former, the null hypothesis is that of equal predictive accuracy of models A and B, that is, $H_0 : E(L_{t,A}) - E(L_{t,B}) = 0$, against the alternative $H_A : E(L_{t,A}) - E(L_{t,B}) \neq 0$. The DM test statistic is

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(d_t)/T_1}} \overset{p}{\longrightarrow} N(0,1),$$

where $\bar{d} = \frac{1}{T_1}\sum_{t=1}^{T_1}(d_t - \bar{d})$, $d_t$ is the loss differential and $\hat{V}(d_t)$ is a heteroskedasticiy and autocorrelation robust (HAC) estimator of the asymptotic variance of $d_t$.[11] The DM test can be employed under a variety of loss functions. For instance, $d_t^{MAE} = |\tilde{\sigma}_t^2 - h_{t,A}| - |\tilde{\sigma}_t^2 - h_{t,B}|$ and $d_t^{MZ\text{-}R^2} = \frac{(\hat{u}_t^A)^2 - (\hat{u}_t^B)^2}{T_1^{-1}\sum(\tilde{\sigma}_t^2 - \bar{\sigma}^2)^2}$ where $\hat{u}_t^A$ are the residuals of a regression of $\tilde{\sigma}_t^2$ on $h_{t,A}$; likewise for $\hat{u}_t^B$.

Let A denote a model which is nested in a larger model B. For nested models, the DM statistic is non-normal resulting in undersized tests with low power. In this paper, we deploy the Harvey et al. (1998) encompassing t-test (ENC-T) developed in the context of the MSE loss differential. Inferences are based on the critical values tabulated by Clark and McCracken (2001) for $(\pi, k_2) = (0.4, 1)$ where $\pi = \frac{T_1}{T - T_1}$ and $k_2$ is the number of excess parameters in model B. $H_0$ is as in the DM test, meaning here that the additional parameters in B do not help prediction (equal MSE) and $H_A$ is that B has smaller MSE than A. Essentially, the ENC-T test statistic is a $t$-statistic for the covariance between $L_{t,A}$ and $L_{t,A} - L_{t,B}$ as follows

$$ENC - T = (T_1 - 1)^{1/2} \frac{\bar{C}}{\sqrt{T_1^{-1}\sum(C_t - \bar{C})^2}}$$

where $C_t = (L_{t,A} - L_{t,B}) \times L_{t,A} = (\tilde{\sigma}_t^2 - h_{t,A})^2 - (\tilde{\sigma}_t^2 - h_{t,A}) \times (\tilde{\sigma}_t^2 - h_{t,B})$.

When several forecasting models are compared through pairwise tests, data mining (snooping) may hinder the significance of the outcome. In contrast to the DM test of equal predictive ability,

[11] The DM statistic does not converge to a standard normal if the evaluation period grows at the same rate as the estimation period because the effect of parameter estimation error does not vanish; a HAC estimator that captures the contribution of parameter uncertainty is then required. However, the distortion from ignoring the latter depends on $\pi = \frac{T_1}{T - T_1}$ and it gets larger, the larger $\pi$ is. In our analysis $\pi = 0.4$ which is not considered large and also the number of parameters to be estimated is relatively small so we do not account for parameter uncertainty.

the test of *superior predictive ability* (SPA) proposed by Hansen (2005) involves a composite hypothesis, thereby being less prone to data mining. The SPA test is designed to address the hypothesis $H_0$ :'any alternative forecast is not better than the benchmark' and it requires bootstrap critical values. Like the DM test, the SPA test can be conducted for any loss function.[12]

A motivation for combining forecasts from different models is that they are likely to capture distinct subtle aspects of the true volatility process, and the relative prominence of such aspects may vary over time. The four nonparametric volatility estimators considered suffer to different extents from market microstructure bias. Hence, it may pay to combine their information content while allowing for their relative role (weight) to time-vary through a rolling estimation approach (fixed window size at 1261 days) as follows. The combining weights for the $t^{th}$ forecast, $\hat{\alpha}_0(t), ..., \hat{\alpha}_4(t)$, $t = 1, ..., T_1$ ($T_1 = 500$) are obtained by regressing the volatility proxy on the in-sample GARCH-RV, GARCH-RR, GARCH-RPV and GARCH-RBP fitted variances over the relevant window $[t-1-(1260), t-1]$. The $t^{th}$ out-of-sample combined forecast is then computed as

$$h_{t,C} = \hat{\alpha}_0(t) + \hat{\alpha}_1(t)h_{t,RV} + \hat{\alpha}_2(t)h_{t,RR} + \hat{\alpha}_3(t)h_{t,RPV} + \hat{\alpha}_4(t)h_{t,RBP}, \quad t = 1, ..., 500$$

where $h_{t,RV}$ denotes the $t^{th}$ out-of-sample forecast from the GARCH-RV model and so forth.

To the best of our knowledge, the issue of different market conditions or regime-switching in the context of volatility forecasting using intraday data has not been addressed. We compare the value-added of intraday information for one-day-ahead volatility forecasting during 'up-market' (U) versus 'down-market' (D) days, and 'high-volume' (H) versus 'low-volume' (L) days. For this purpose, we classify and average the forecast errors into those incurred during up-(high-) or down-market (low-volume) days. Our definition of up/down market days is a short-term one based on the moving average of the daily return over the most recent 5-day window. Since the goal is to forecast the volatility on day $t$, the switching variable is a one-day-lagged ($t$-1) indicator function

$$S_{t-1} = \begin{cases} 1 & if \quad \frac{1}{5}\sum_{i=1}^{5} r_{t-i} > 0 \quad (\textit{Up-market} \text{ day}) \\ 0 & else \qquad\qquad\qquad\quad (\textit{Down-market} \text{ day}) \end{cases} \tag{11}$$

[12] The SPA test is implemented in OxMetrics 5 using Peter Hansen's code which we gratefully acknowledge. We focus the analysis on the two predefined loss functions in the code, MSE and MAE.

which equals 1, signifying a positive direction of the market if the moving average of the daily returns over the most recent 5-day period is positive.

For the high- versus low-volume days comparison the short-term indicator function is

$$V_{t-1} = \begin{cases} 1 \quad if \quad VOL_{t-1} > \frac{1}{5}\sum_{t=1}^{5} VOL_{t-1-i} & (\textit{High-volume } \text{day}) \\ 0 \quad else & (\textit{Low-volume } \text{day}) \end{cases} \tag{12}$$

Two questions are asked: (a) Does the ranking of intraday augmentation measures differ over market conditions?, (b) Do the benefits from exploiting intraday data differ over market conditions?

# 4 Empirical Results

## 4.1 Data and distributional properties

The transaction price and number-of-shares traded data is from Tick Data.[13] The observations pertain to 14 stocks traded on the NYSE and span the period 02/01/97 to 31/12/03, a total of 1761 days. The stocks are American Express (AXP), AT&T (ATT), Boeing (BA), Caterpillar (CAT), DELL, General Electric (GE), General Motors (GM), JP Morgan (JP), KO (Coca-Cola), McDonald (MCD), Microsoft (MSFT), Procter & Gamble (PG), WAL-MART (WMT) and IBM.

Table 1 shows the distributional properties of daily returns and trading volume. For all stocks, the returns are non-normally distributed, particularly, in the form of excess kurtosis.[14]

[Table 1 around here]

The Ljung-Box (LB) statistic suggests no autocorrelation in daily returns for many stocks — the exceptions are ATT, DELL, GM, IBM PG, and WMT returns. By using mean volume as measure of trading activity, stocks can be ranked from more to less active as: MSFT, DELL, GE, IBM, JPM, WMT, AXP, MCD, KO, BA, GM, PG, ATT and CAT. Trading volume is stationary as borne out by the ADF test and the Robinson $d$ statistic but the latter suggests long memory in

[13] *www.tickdata.com* provides high frequency data on a commercial basis for equity and commodity markets.

[14] Due to space constraints, an extensive collection of our empirical results are given in an Appendix which can be downloaded from *http://www.cass.city.ac.uk/faculty/a.fuertes*.

volume.[15] DELL's volume is the most persistent whereas the smallest persistence is shown by BA's volume. The assumption that volume follows a lognormal distribution was first advocated in Clark's (1973) MDH model and is still widely used. The Kolmogorov-Smirnov (KS) test suggests volume log-normality for half of the sampled stocks: BA, CAT, GE, GM, JPM, KO and MCD.[16]

Table 2 reports summary statistics for five distinct daily volatility estimators.

[Table 2 around here]

In line with the MDH theory, studies by Clark (1973), Tauchen and Pitts (1983), Richardson and Smith (1994) and Andersen and Bollerslev (1997a,b) document several interesting stylized facts about the unobserved, latent information flow driving the volatility process. These include small variation relative to its mean, lognormality, high persistence, correlation with volume and bringing normality in returns. We adopt these stylized facts as in-sample criteria to compare the estimators.

RV and RBP have approximately the same mean (e.g. for IBM, the mean RV and RBP are, respectively, 3.572 and 3.340). The mean of RR is generally smaller than that of RV with two exceptions only (DELL and MSFT). This is in line with the findings in Martens and van Dijk (2006) which illustrate that infrequent trading induces a downward bias in the RR, while it does not affect RV. The mean of RPV(for $z = 1.5$) is slightly higher than those of the other intraday-estimated volatility measures. But RPV is not in the same units as the other three measures, so any comparison of their moments has to be interpreted with caution.[17] Relative to its mean, RPV has generally the lowest dispersion (standard deviation) which suggests that it is the least

[15] A fractional integration parameter $0 < d < 0.5$ characterizes stationarity with long memory so that the autocorrelation function decays at a hyperbolic rate rather than exponentially as in short-memory ($d = 0$) processes.

[16] In contrast with the JB test that focuses on the skewness and kurtosis only, the KS test compares the cumulative distributions of the input data and the fitted distribution and so it has been shown to be more powerful that the JB test. The KS statistic is computed as $KS = max(D^+, D^-)$ with $D^+ = \max(\frac{t}{T} - F(VOL_t))$ and $D^- = \min(\frac{t}{T} - F(VOL_t)), t = 1, .., T$ where $F(\cdot)$ is the fitted lognormal distribution.

[17] The order chosen for the RPV is $z = 1.5$ throughout the paper. Building on the results in Liu and Maheu (2005), to make this choice we compare RPV(0.5), RPV(1) and RPV(1.5) according to their distributional properties, in-sample model-fit and out-of-sample forecasting properties. Firstly, daily returns standardized by RPV ($z = 0.5$) become normal at the 10%, 5% or 1% level in none of the stocks, 7 stocks ($z = 1$), and 9 stocks ($z = 1.5$). Second, the model fit of GARCH-RPV is clearly superior, according to the loglikelihood, AIC and SBC, for $z = 1.5$ also. Third, for the majority of stocks according to virtually all loss functions considered, the forecast errors of GARCH-RPV are smaller for $z = 1.5$. Detailed results can be found in Appendix Tables A and B.

noisy in the present context followed by RR. At the other extreme, the crude squared return has a StDev/Mean ratio about five times larger than RPV.

The ADF test suggests that all five unconditional volatility measures are stationary. But the degree of persistence (Robinson $d$) of the intraday-estimated measures is substantially higher than that of daily squared returns. RPV and RR are the most persistent (followed by RV and then RBP) and so they may provide a better signal for future volatility. All measures show positive skewness and large kurtosis with squared returns having the largest kurtosis. The KS test suggests that lognormality for the intraday measures but not for daily squared returns. This is in line with ABDE (2001) and ABDL (2001), inter alios, who show that the lognormal distribution provides a good fit for realised volatility. Figure 1 plots for the least traded (CAT) and most traded (MSFT) stock the different volatility estimators alongside volume (scaled by $10^7$).

[Figure 1 around here]

Normality of returns is an assumption that underlies many financial theories, for instance, the Black-Scholes option pricing model and some VaR approaches. But daily stock returns are clearly non-normal. Several studies have recovered normality by subordinating returns to the 'financial clock' using RV as standardization variable (ABDE, 2001; ABDL, 2001; BN-S, 2002a; Bandi and Russell, 2006; and Areal and Taylor, 2001). On this basis, they conclude that RV reflects well the information flow to the market.[18] We standardize the daily returns by the various uncondi-

[18]Daily returns are non-normal because information is available to traders at a varying rate so the price process evolves at different rates during identical time intervals. When no information is available, trading is slow, and the price evolves slowly. When new information arrives, trading is brisk and the price process evolves much faster. The upshot is that the number of individual random effects added together to give the daily price change (return) is non-constant, rendering the Central Limit Theorem inapplicable. The theoretical motivation for expecting returns to be normal when standardized by volume starts from Clark (1973) and Monroe (1978). Clark argues that returns are non-normal when sampled at intervals which are equidistant in calendar time but are normal in the time scale of the latent trading activity (called 'financial time'), that is, when computed over intervals with equivalent trading activity. Clark further shows that trading volume can be taken as an instrument for the true operational time or 'imperfect clock' measuring the speed of evolution of the price change process. By standardizing returns with volume, in effect, the returns 'clock' is subordinated to that of volume. Monroe provides further justification by showing that any semi-martingale can be written as a time-changed Brownian motion. Monroe's result in essence tells us that there exists a time-filtration mechanism that can restore return normality. Andersen et al. (2005) show that using high-frequency data for the construction of the proxy for the 'financial time' (they sample at intervals of

tional (intraday) and conditional (GARCH) volatility measures and analyze the extent to which standardized returns become normal. Table 3 sets out the results.

[Table 3 around here]

The findings are consistent with the literature where it has been documented that historical GARCH (and SV) models do not adequately capture all the leptokurtosis in daily returns. The standardization by GARCH mitigates but does not eliminate all the factors that induce non-normality.[19] RR is the most successful in bringing normality (for all stocks except MSFT) followed by RPV, RBP and RV. The lognormality of the intraday-estimated variability measures together with their ability to bring normality of standardized returns provides support for Clark's (1973) contention that asset returns follow a normal-lognormal mixture in the context of cotton futures.

In the MDH literature, volume is taken as a proxy for the latent trading activity process (Clark, 1973). The MDH posits that volume and volatility are positively correlated because they are simultaneously influenced by the rate of news arrival. The extent of this correlation provides a further ranking for the volatility measures — Appendix C reports, for each stock, the pairwise correlations among volume and the nonparametric and GARCH volatility measures. The correlation of volume with each of the volatilities is positive but the nonparametric measures show higher correlation with volume than GARCH. The highest correlation occurs with RR and RPV at 43.4% and 42.04%, respectively, on average across stocks. The correlation between the four intraday measures is high (above 90%) but drops to about 50% between the intraday measures and GARCH.

To sum up, in the context of the MDH stylised facts, RPV and RR emerge as superior intraday-estimated measures of the daily price variation given that they display the longest memory, the smallest standard deviation relative to their mean, and the highest correlation with volume. RR fares slightly better than RPV, however, in bringing returns normality.

equivalent QV) will appproximately time-scale returns and render them i.i.d. Gaussian. Their approach deviates from the MDH literature in that there is no trading activity proxy involved.

[19] The GARCH model used, fitted to the daily returns, is a GARCH(1,1) for most stocks. However, in some cases higher orders are needed to absorb all the volatility clustering. The models are described in Section 4.2.

## 4.2 Model estimates and in-sample model fit ranking

The estimation results by QML over the entire sample for the least traded (CAT) and most traded (MSFT) stocks are presented in Table 4, and for the remaining stocks in Appendix Table D.

[Table 4 around here]

For all stocks, the lagged intraday volatility ($v_{t-1}$) is strongly significant at the 1% level but, in line with the literature, lagged volume is insignificant with the exception of CAT, DELL, KO and MSFT — the significance of lagged volume in a GARCH equation would provide support for a simple MDH version (Luu and Martens, 2002). A uniform result across stocks is that the inclusion of an intraday volatility measure in the GARCH equation results in a substantial reduction in volatility persistence which, according to the MDH argument, suggests that they capture well the news arrival process — the MDH posits that the volatility clustering is partly explained by time dependence in the public information flow. Moreover, the intraday-estimated volatilities turn the ARCH coefficients from strongly significant at the 1% level to either insignificant or marginally significant at the 10% level. This suggests that the predictive information on future daily volatility contained in RV, RR, RPV or RBP encompasses the information in daily squared returns.

In order to assess in-sample model fit, the log-likelihood (lnL), AIC and SBC values can be compared across models for a given stock since they all refer to the same dependent variable. The lnL of the GARCH models augmented by each of the four intraday volatilities are greater than those of GARCH for all stocks. But this is not the case for many stocks with the volume measure. Second, the GARCH models augmented with intraday volatility measures rank top also according to the AIC and SBC. For the least traded (CAT) stock, the dynamics of daily returns seems to be best captured by the GARCH-RPV model as suggested by the lnL (largest) and AIC (smallest) values. The model ranking according to lnL, AIC and SBC is GARCH-RPV, GARCH-RV, GARCH-RBP, GARCH-RR, GARCH-VOL and GARCH.

The volatility dynamics of the most traded (MSFT) stock is best captured by GARCH-RR according to the lnL and AIC. However, as for CAT, the least improvement is brought by GARCH-VOL with a relatively small increase in lnL and a fall in the AIC. Furthermore, according to SBC, the baseline GARCH is superior to GARCH-VOL. The model-fit ranking based on lnL and AIC is GARCH-RR, GARCH-RPV, GARCH-RV, GARCH-RBP, GARCH-VOL and GARCH and the only change in this ranking according to the SBC is in the relative fit of the last two models. The ranking for the other stocks (see Appendix D) suggests that overall the GARCH-RPV provides the best fit in 8 out of 14 stocks, and GARCH-RR, GARCH-RV and GARCH-RBP in 2 stocks each.

## 4.3 Out-of-sample forecast ranking

Figure 2 provides a bar-chart summary of the forecast 'horse race'. The bar length is the proportion of stocks for which a given model provides the most accurate forecasts.

[Figure 2 around here]

The RPV estimator brings the largest forecast gains for most stocks and loss functions, e.g. HMSE, MME(O) and LL are smallest for GARCH-RPV in 71.4%, 71.4% and 64.3% of the stocks, respectively. GARCH-RPV is the top performer according to (virtually) all forecast criteria in seven stocks: AXP, CAT, GE, IBM, JPM, ATT and WMT. Appendix E and F provide further details.

The forecast error measures and the MZ-$R^2$ are set out in Table 5. For each stock, the last row reports the improvement that the best augmented-GARCH brings versus the GARCH.

[Table 5 around here]

For all stocks and loss functions, the inclusion of a lagged intraday volatility measure in the GARCH equation notably improves the forecasts. In particular, for some stocks the MZ-$R^2$ of the best augmented-GARCH more than trebles that of GARCH. For instance, for the least traded CAT stock, the MZ-$R^2$ of GARCH is 13.99% whereas that of the best forecasts, GARCH-RPV, is 43.64%. The forecast enhancement of GARCH-RV versus GARCH is in line with the findings in Martens (2001) and Koopman et al. (2005) for FX rates and the S&P100 index, respectively, and with Grané and Veiga (2007) for four DJIA stocks, American Express, Coca-Cola, Disney and Pitzer. Adding lagged trading volume to GARCH does not bring forecast gains which is consistent with the results in Brooks (1998) and Donaldson and Kamstra (2004) for stock market indexes. The ENC-T test based on the MSE for the comparison of GARCH and best augmented-GARCH (nested models) suggests that forecast improvement is significant at the 1% level throughout. Moreover, there tend to be significant differences in forecast accuracy between the alternative augmented-GARCH (non-nested) models as suggested by the DM test for most loss functions.

Considering the 11 loss functions and the 14 stocks, a total of 154 pairwise combinations, in 55% of them the GARCH-RPV model is the top forecaster, followed by GARCH-RR (19%), GARCH-RV (14%) and GARCH-RBP (12%).[20] For the least traded (CAT) stock, GARCH-RPV leads for virtually all loss functions followed by GARCH-RV. Exceptions are the asymmetric MME(U) and MME(O) for which the minimum loss is achieved by GARCH-RV but is closely followed by GARCH-RPV. GARCH-VOL is ranked last. For the most traded (MSFT) stock, according to virtually all loss functions, the best forecasts are those from the GARCH-RR model followed by GARCH-RPV. The HMSE loss is an exception with a minimum that corresponds to GARCH-RPV.

DELL and MSFT show similar behavior in the sense that, virtually according to all loss functions, the GARCH-RR provides the best forecasts. This is in contrast with the earlier finding that DELL and MSFT are the two stocks for which the RR measure has most difficulty in bringing return normality (c.f. Table 3). This suggests that whether or not a given intraday volatility measure brings normality in daily returns may not necessarily tell us much regarding its forecasting power. Moreover, for the MSFT stock, RR is not lognormally distributed at the 5% level whereas RPV, RBP and RV are. Likewise for CAT stock, although RPV emerges as top forecaster, it is not lognormally distributed whereas RV, RR and RBP are. For IBM, unanimously across all loss

[20] Using 30-minute DM/US$ and Yen/US$ data, Martens (2001) shows that extending the daily GARCH model with the sum of intraday squared returns leads to similar improvement as modeling the intraday returns directly. Hence, our results indirectly suggest that extending the daily model with RPV is superior to the latter also.

functions, RPV is the best forecaster although, interestingly, it fails to bring normality in daily returns (c.f. Table 3). Moreover, for IBM the highest persistence is shown by RR ($d = 0.403$) followed by RPV ($d = 0.394$). Therefore, when scrutinizing the individual stocks, some mismatch is observed between the ranking from forecast and MDH-related criteria.[21]

Patton (2006) and Hansen and Lunde (2006a) show theoretically and via simulation that many criteria used in the literature are inconsistent when the evaluation is based on a volatility proxy (i.e. $\tilde{\sigma}_t^2$ instead of $\sigma_t^2$) so they may favour an inferior model with a probability that converges to one asymptotically as the holdout sample ($T_1$) increases. Exceptions are MSE, GMLE and the levels MZ-$R^2$. Patton shows that the more efficient (less noisy) the proxy, the smaller the degree of distortion in the ranking which depends also on the sampling frequency. Across various loss functions, he shows analytically that when 5-min RV is used as proxy almost all of the objective bias disappears. Interestingly, the model rankings are quite similar across loss functions in Table 5 but they become rather more unstable when the crude daily $r_t^2$ is used as volatility proxy (see Appendix H). For instance, for AXP virtually all loss functions chose the same (GARCH-RPV) model as best when the sum of 5-min squared returns is the volatility proxy in Table 5 whereas four models (GARCH, GARCH-RV, GARCH-RV and GARCH-RBP) emerge as best from one criteria or another when $r_t^2$ is used as proxy in Appendix H. Reassuringly, in a majority of cases the MSE, GMLE and the MZ-$R^2$ criteria tend to point to the same best model when the two different proxies are used — illustrative examples are AXP and MSFT for which GARCH-RPV and GARCH-RR, respectively, are selected according to all three criteria irrespective of the proxy.

[21] We investigated whether the forecast ranking of the nonparametric volatility estimators changes when they are based on 15- and 30-min data. Appendix G reports the results for the two least-traded (CAT, ATT) and the two most-traded stocks (MSFT, DELL). The top forecaster remains the RPV. As expected, the forecast losses (like-for-like models) rise as the sampling frequency decreases, in line with the results in Pong et al. (2004) for FX rates using 5- and 30-min frequencies. Futhermore, the value-added of the nonparametric volatilities is larger at the 5-min than at the 15- and 30-min. For instance, for CAT the forecast error reduction from GARCH to GARCH-RPV is 53.88%, 43.53% and 37.78%, respectively, at the 5-, 15- and 30-min frequencies. These findings corroborate that the 5-min sampling is more useful than 15- and 30-min from the viewpoint of predicting future daily volatility. The forecast comparison is conducted by proxying the forecast target (the conditional variance) by either the sum of 5-min squared returns or the sum of 15- (or 30-) min squared returns. The main finding is that the forecast errors of like-for-like models increase when the sampling frequency of the proxy decreases.

Table 5 also reports the average losses of the combined forecasts using the time-varying weighting approach described in Section 3.3 (denoted COMBINED). We also considered an equal-weights combining scheme (denoted COMB-EQW). Notwithstanding the high correlation between the GARCH-RV, GARCH-RR, GARCH-RPV and GARCH-RBP forecasts being combined (inefficiency of weight estimates), overall across stocks and loss functions the varying-weights approach gives better results than the equal-weights approach and so only the former are reported. It is tempting to attribute this to the fact that the regression-based combining approach accounts for bias through the intercept $\hat{\alpha}_0(t)$. But the bias corrected MZ-$R^2$ is only higher with the equal-weights approach for half of the stocks and there is evidence of forecast bias for all but 2 stocks (see Appendix I).[22]

The COMBINED model yields the smallest forecast error with virtually all loss functions for 4 stocks (CAT, DELL, GM and MSFT) and the gains relative to the best augmented-GARCH can be as much as 51%. For instance, the HMSE of the best augmented-GARCH is reduced by 45% for DELL, 51% for GM and 44% for MSFT. For 4 other stocks (AXP, GE, IBM and ATT) the COMBINED model is in the lead for at least half of the loss functions. The ENC-T test suggests that the MSE from COMBINED is significantly smaller than that from the best augmented-GARCH in 6 stocks. Hence, jointly exploiting the information from all four intraday measures can be fruitful which indirectly corroborates that they are affected differently by microstructure noise.

The results of the SPA test are set out in Table 6. The null hypothesis is that the GARCH-RPV model which was found to be in the lead for most stocks and across most forecast accuracy metrics (hence, taken as benchmark) is not worse than any of the alternative models. These are GARCH, GARCH-RV, GARCH-RR, GARCH-RBP, GARCH-VOL, COMBINED and COMB-EQW.

[Table 6 around here]

The SPA test $p$-values suggest that, for eight stocks, GARCH-RPV is not outperformed by any of the seven alternative models in terms of both the MSE and MAE losses. For four further stocks

[22] Appendix H further illustrates that: (a) the GARCH-RV, GARCH-RR, GARCH-RPV and GARCH-RBP forecasts suffer from upward (if any) biases, particularly the latter three, whereas GARCH-VOL is downward biased, (b) the equal-weight forecasts are upward biased but the estimated-weights combined forecasts are generally unbiased.

(CAT, DELL, GM, MSFT), GARCH-RPV is significantly beaten by the COMBINED model. For two stocks (ATT and IBM) the evidence is mixed with MSE and MAE favouring GARCH-RPV and the COMBINED model, respectively. The last column of the table summarises the SPA test results: GARCH-RPV significantly emerges most often as the model with superior predictive ability.

Tables 7 and 8 report the HMSE, AMAPE and LL criteria, respectively, for up- and down-market days, high- and low-volume days as defined in (11) and (12).[23] For each stock, the last row (Benefit %) reports the forecast error reduction that the best augmented-GARCH brings relative to GARCH. Italics are used to signify the regime in which the largest reduction is achieved.

[Table 7 around here]

The ranking of the augmented-GARCH models is virtually identical in both regimes and GARCH-RPV ranks top for most stocks. The forecast losses tend to be smaller for down-market days. This suggests that the daily stock volatility at day $t$ is relatively more difficult to forecast when $t-1$ is an up-market day. In the light of this finding, it is not surprising to see that the largest benefits from exploiting intraday data in order to generate a day $t$ volatility forecast tend to occur when $t-1$ represents an up-market day. For instance, for DELL the percentage reduction in the GARCH forecast errors is 5.59 (HMSE), 5.75 (AMAPE) and 10.30 (LL) over down-market days whereas it increases, respectively, to 37.90, 14.42 and 27.51 over up-market days.

Table 8 suggests that the forecast ranking is almost identical over high- and low-volume days with the GARCH-RPV model having the smallest forecast errors.

[Table 8 around here]

Generally, the losses tend to be somewhat smaller in the high-volume regime suggesting that the conditional volatility on day $t$ appears to be less forecastable if there was low volume on day $t-1$;

[23] Since one goal is to compare the forecast errors across market conditions, and down-market (high volume) days tend to be more volatile than up-market (low volume) days, measures such as MSE and GMLE can be misleading for this purpose. Unit-free measures such as the HMSE, AMAPE, LL and Theil-U will be more informative. The unreported (for space constaints) HMAE and Theil-U give similar results.

exceptions are IBM, MCD, MSFT and ATT. Consistent with the latter, the largest reduction in forecast errors from using intraday data (benefit %) is obtained for the low-volume days. For instance, for DELL the reduction in the GARCH forecast error by including RPV is 8.38% (HMSE), 3.76% (AMAPE) and 9.38% (LL) in the high-volume regime whereas it increases, respectively, to 39.68, 16.31 and 29.46 in the low-volume days. Another interesting case is GM (and to some extent BA) for which there is no forecast error reduction in high-volume days, but the reduction is 12.13% (HMSE), 3.68% (AMAPE) and 6.61% (LL) in low-volume days.

The upshot is that 1-day-ahead the conditional volatility is easier to forecast when the stock is underperforming ($t-1$ is a down-market day) and when trading volume is relatively high ($t-1$ is a high-volume day). Table 9 reports the average return, volatility and volume for days $t-1$ and $t$ over different market conditions. As expected, the volatility is higher in down-market and high-volume days. Our findings are in line with the view that high volatility and volume arise largely from news arrival: when $t-1$ is a high volatility (volume) day more public information is available which, in turn, helps to forecast volatility at $t$.[24] The same rationale applies if $t-1$ is a down market day, because trading volume (and volatility) is higher in down versus up days. This effect is exacerbated because, as Admati and Pfleiderer (1988) show, trades from both informed and discretionary liquidity traders come in clusters, with both groups preferring to trade during 'thick' markets. This clustering of trades, when trading activity is already high, triggers the release of even more information. Moreover, high trading activity may to some extent mitigate the microstructure noise (e.g. infrequent-trading effects) and this could also explain why the augmented-GARCH models tend to produce better forecasts during high volume (volatility) days.

[Table 9 around here]

Table 9 also reports the average correlation between the 'true' volatility on days $t-1$ and $t$. The stronger correlations for down-market and high-volume days are in line with the finding of smaller

[24] Using firm-specific announcements data, Kalev et al. (2002) show that public information flows drive volatility and volume simultaneously, in line with the MDH. But the observed volume may also partly reflect liquidity pressures or the 'game' played by strategic traders with heterogenous information and the revision of dispersed beliefs.

forecast errors during such market conditions. Hence, the use of intraday data is more crucial when markets are relatively tranquil (low volatility), that is, on up-market and low-volume days.

## 5 Conclusions

How to forecast daily volatility is a challenging question because, unlike prices and volume, volatility is not directly observable. A recent literature focuses on exploiting the intraday variation and proposes several nonparametric estimators called *realised* variance, range, power variation and bipower variation. This paper compares these estimators on the basis both of their distributional properties and their ability to forecast one-day-ahead the conditional variance of returns. A GARCH framework is adopted as platform to compare their incremental predictive content which is not to suggest that GARCH is the 'best' framework for volatility prediction. For completeness, a volume measure of intraday trading activity is also included in the horse race.

The popular realized variance estimator is dominated by the realised power variation and the realised range. The realized range fares relatively well in the in-sample distributional analysis regarding the extent to which it brings normality in standardized returns. However, overall across stocks and loss functions the realised power variation appears to be the top performer for short term forecasts of one day. This means that, among the four nonparametric volatility estimators, the lattter enhances the GARCH forecasting ability the most. A rationale for this finding is that the realized power variation is not only immune to jumps, like the realized bipower variation, but it is also the most persistent and less noisy. Nevertheless, forecast combining appears worthwhile for about half of the stocks which indirectly corroborates that the four intraday-estimated volatility measures are impacted by microstructure noise in different ways. The additional use of intraday data on day $t-1$ to forecast volatility on day $t$ is more advantageous when $t-1$ is an up-market or low volume day relative to the recent past. Since daily volatility forecasts are key inputs for VaR analysis, our findings may have important practical implications for this area of risk management.

# References


[1] Andersen, T., & Bollerslev, T. (1997*a*). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The Journal of Finance, 52*, 975-1005.

[2] Andersen, T., & Bollerslev, T. (1997*b*). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance, 4*, 115-158.

[3] Andersen, T., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review, 39*, 885-906.

[4] Andersen, T., Bollerslev,T, & Diebold, F.X. (2002). Parametric and non-parametric volatility measurement, in L.P. Hansen and Y. Ait-Sahalia (eds.), *Handbook of Financial Econometrics*, Amsterdam: North-Holland.

[5] Andersen, T., Bollerslev,T, & Diebold, F.X. (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics, 89*, 701-720.

[6] Andersen, T., Bollerslev, T., Diebold, F.X. & Ebens, H. (2001). The Distribution of Realised Stock Return Volatility. *Journal of Financial Economics, 61*, 43-76.

[7] Andersen, T., Bollerslev,T, Diebold, F.X, & Labys, P. (2001). The Distribution of Realised Exchange Rate Volatility. *Journal of the American Statistical Association, 96*, 42–55.

[8] Andersen, T., Bollerslev, T., & Dobrev, D. (2005). Nonparametric Exploration of Continuous Time Volatility Models with Leverage and Jumps. Kellogg School, Northwestern University, mimeo.

[9] Areal, N., & Taylor, S. (2001). The realised volatility of FTSE-100 futures prices. *Journal of Futures Markets, 22*, 627-648.

[10] Bandi, F.M., & Russell, J.R. (2006). Separating Microstructure Noise from Volatility. *Journal of Financial Economics*, *79*, 655-692.

[11] Barndorff -Nielsen, O. E., & Shephard, N. (2002a). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*, 253-280.

[12] Barndorff -Nielsen, O. E., & Shephard, N. (2002b). Estimating quadratic variation using realised variance. *Journal of Applied Econometrics, 17*, 457-477.

[13] Barndorff -Nielsen, O. E., & Shephard, N. (2004a). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics, 4*,1-30.

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

[15] Barndorff-Nielsen, O. E. & Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics, 4*, 1–30.

[16] Bessembinder, H., & Seguin, P.J. (1993). Price volatility, trading volume and market depth: Evidence from future markets. *Journal of Financial and Quantitative Analysis*, *28*, 21-39.

[17] Blair, J.B., Poon, SH., & Taylor, S.J. (2001). Forecasting S&P100 volatility: The incremental information content of implied volatilities and high frequency index returns'. *Journal of Econometrics* 105, 5-26.

[18] Bollerslev, T., & Ghysels, E. (1996). Periodic autoregressive conditional heteroskedasticity. *Journal of Business and Economic Statistics, 14*, 139-157.

[19] Bollerslev, T., Engle, R.F., & Nelson, D. (1994). 'ARCH models' in Engle, R.F. and McFadden, D.L. (Eds), *Handbook of Econometrics,* IV Elsevier (Chapter 49).

[20] Brailsford, T. J., & Faff, R. W. (1996). An Evaluation of Volatility Forecasting Techniques. *Journal of Banking and Finance, 20*, 419-438.

[21] Brandt, M.W., & Diebold, F.X. (2006). A no-arbitrage approach to range based estimation of return covariances and correlations. *Journal of Business, 79*, 61-74

[22] Brooks, C. (1998). Predicting stock index volatility using GARCH models: Can market volume help? *Journal of Forecasting*, *17*, 57-80.

[23] Christensen, K., & Podolskij, M. (2005). Realised range-based estimation of integrated variance. Aarhus School of Business, mimeo.

[24] Clark, P.K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometica, 41*, 135-155.

[25] Clark, T. E., & McCracken, M. W. (2001). Tests for equal forecast accuracy and encompassing for nested models. *Journal of Econometrics, 105*, 85-110.

[26] Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253-263.

[27] Donaldson, G., & Kamstra, M. (2004). Volatility forecasts, trading volume and the ARCH versus option-implied volatility trade-off. Federal Reserve Bank of Atlanta WP 04-06

[28] Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987-1006.

[29] Engle, R., & Gallo, G.M. (2006). A multiple indicators model for volatility using intradaily data. *Journal of Econometrics, 127*, 3-27.

[30] Engle, F., Sokalska, M.E., & Chanda, A. (2006). Forecasting intraday volatility in the US market. Stern School of Business, NYU, mimeo.

[31] Forsberg, L., & Ghysels, E. (2007). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, *5*,31-67.

[32] Franses, P. H., & van Dijk, D. (1996). Forecasting stock market volatility using non-linear GARCH models. *Journal of Forecasting*, *15*, 229-235.

[33] Galbraith, J, & Kisinbay, T. (2002). Information content of volatility forecasts at medium-term horizons. *CIRANO Working Papers 2002s-21.*

[34] Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics, 131*, 59-95.

[35] Grané, A., & Veiga, H. (2007). The effect of realised volatility on stock returns risk estimates. *Universidad Carlos III,* Madrid, Working Paper 07-63, .

[36] Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics, 23*, 365-380

[37] Hansen, P.R., & Lunde, A. (2006a). Consistent ranking of volatility models, *Journal of Econometrics*, *131*, 97-121.

[38] Hansen, P.R., & Lunde, A. (2006b). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, *24*, 127-161.

[39] Harvey, D. I., Leybourne, S. J., & Newbold, P. (1998). Tests for forecast Encompassing. *Journal of Business and Economic Statistics*, *16*, 254-259.

[40] Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finan*ce, *42*, 281-300.

[41] Jungbacker, B., & Koopman, J. (2005). Model-based measurement of actual volatility in high frequency data. Tinbergen Institute Amsterdam, mimeo.

[42] Kalev, P.S., Liu, W.-M., Pham, P.K., & Jarnecic, E. (2002). Public information arrival and volatility of intraday stock returns. *Journal of Banking and Finance*, *28*, 1441-1467.

[43] Koopman, S.J., Jungbacker, B., & Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, *12*, 445-475.

[44] Lamoureux, C. G., & Lastrapes, W. D. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance*, *45*, 221-229.

[45] Laux, P.A., & Ng, L.K. (1993). The sources of GARCH: empirical evidence from an intraday returns model incorporating systematic and unique risks. *Journal of International Money and Finance, 12*, 543–560.

[46] Liu, C., & Maheu, J. (2005). Modeling and forecasting realised volatility: The role of power variation. University of Toronto, Department of Economics, mimeo.

[47] Luu, J.C., & Martens, M. (2002). Testing the mixture of distributions hypothesis using realised volatility. *Journal of Futures Markets, 23*, 661-679.

[48] Makridakis, S. (1993). Accuracy measures: Theoretical and practical concerns. *International Journal of Forecasting*, 9, 527-529.

[49] Martens, M. (2001). Forecasting daily exchange return volatility using intraday returns. *Journal of International Money and Finance*, *20*, 1-23.

[50] Martens, M., & van Dijk, D. (2006). Measuring volatility with the realised range. *Journal of Econometrics, 138*, 181-207.

[51] Monroe, I. (1978). Processes that can be embedded in Brownian motion', *Annals of Probability, 6*, 42-56.

[52] Najand, M., & Yung, K. (1991). A GARCH examination of the relationship between volume and price variability in futures markets. *Journal of Futures Markets*, *11*, 613-621.

[53] Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatilities. *Journal of Econometrics,* 45, 267-290.

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

[55] Patton, A. (2006). Volatility forecast comparison using imperfect volatility proxies. *University of Technology Sidney*, QFRC WP175.

[56] Pong., S., M. Shackleton, S. Taylor, & Xu, X. (2004). Forecasting sterling/dollar volatility: A comparison of implied volatilities and AR(FI)MA models. *Journal of Banking and Finance*, *28*, 2541-2563.

[57] Richardson, M., & Smith, T. (1994). A direct test of the mixture of distributions hypothesis: measuring the daily flow of information. *Journal of Financial and Quantitative Analysis*, *29*, 101-116.

[58] Robinson, P.M. (1995). Log-periodogram regression of time series with long-range dependence. *The Annals of Statistics, 23*, 1048-1072.

[59] Ross, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, *39*, 1127-1139.

[60] Tauchen, G., H. Zhang, & Liu, M. (1996). Volume, volatility and leverage: A dynamic analysis. *Journal of Econometrics*, *74*, 177-208.

[61] Tauchen, G., & Pitts, M. (1983). The price variability -volume relationship on speculative markets. *Econometrica, 51*, 485–505.

[62] Taylor, S. J., & Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance, 4*, 317-340.

[63] Taylor, S.J., 1987. Forecasting the volatility of currency exchange rates. *International Journal of Forecasting, 3*, 159–170.

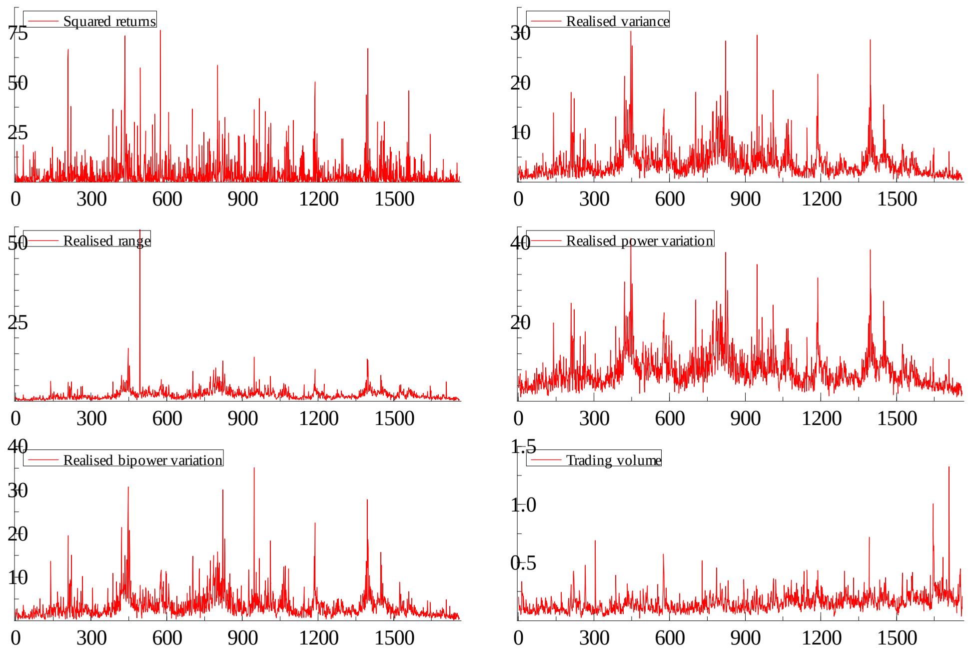
[64] Zhou, B. (1996). High-frequency data and volatility in foreign exchange rates'. *Journal of Business and Economic Statistics*, *14*, 45-52.

**Biographies:** Ana-Maria FUERTES (BSc Eng, MSc Control Eng, PhD Economics) is currently Professor in Financial Econometrics at the Faculty of Finance, Cass Business School, City University, London. Her research interests are in time series analysis, panel data methods, forecasting and empirical finance. She has published articles in the *International Journal of Forecasting, Journal of Economic Dynamics and Control*, *Economics Letters*, *Journal of International Money and Finance*, *Journal of Banking and Finance, Journal of Applied Econometrics* and *Computational Statistics and Data Analysis.*

Elena KALOTYCHOU (BA/MA Cantab Maths, MSc Operational Research, PhD Finance) is currently Lecturer in Finance at the Faculty of Finance, Cass Business School, City University, London. Her research interests are in applied econometrics, credit risk, financial modelling and forecasting. She has published in the *International Journal of Forecasting*, *Computational Statistics and Data Analysis, Financial Markets Institutions and Instruments, Journal of Multinational Financial Management* and *Applied Economics.*

Marwan IZZELDIN (BSc Economics, MA Economics, PhD Finance) is currently Lecturer in Economics and Econometrics at the Department of Economics, Lancaster University Management School. His research interests are in high frequency data, long memory, bootstrap methods and islamic banking. He has published in the *Journal of Applied Econometrics* and *The Economic and Social Review.*

A. Least traded stock (CAT, Caterpillar)




B. Most traded stock (MSFT, Microsoft)

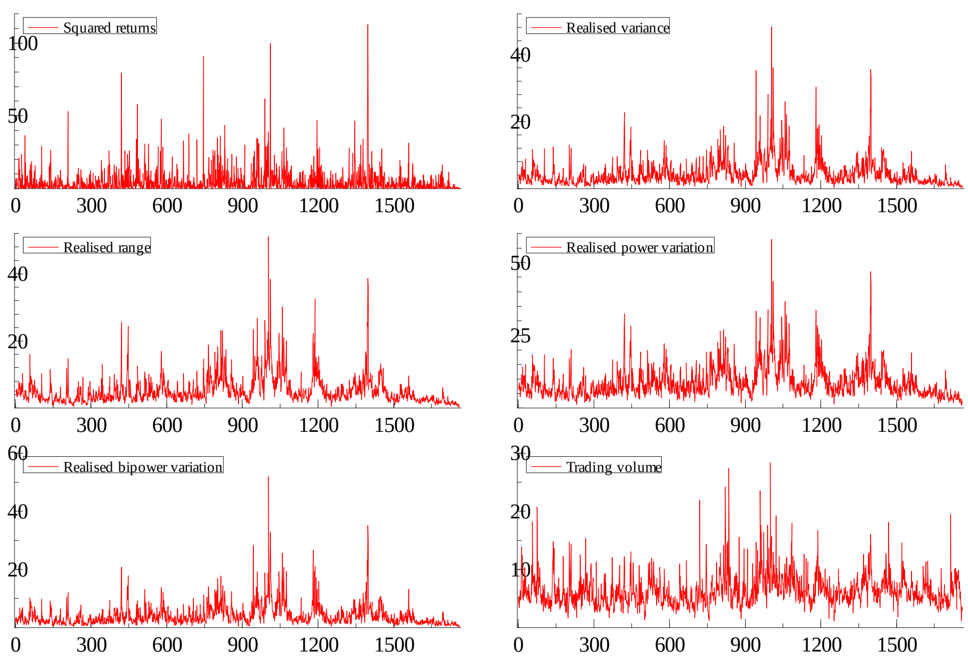



Fig. 1 Time series plots for daily volatility and volume

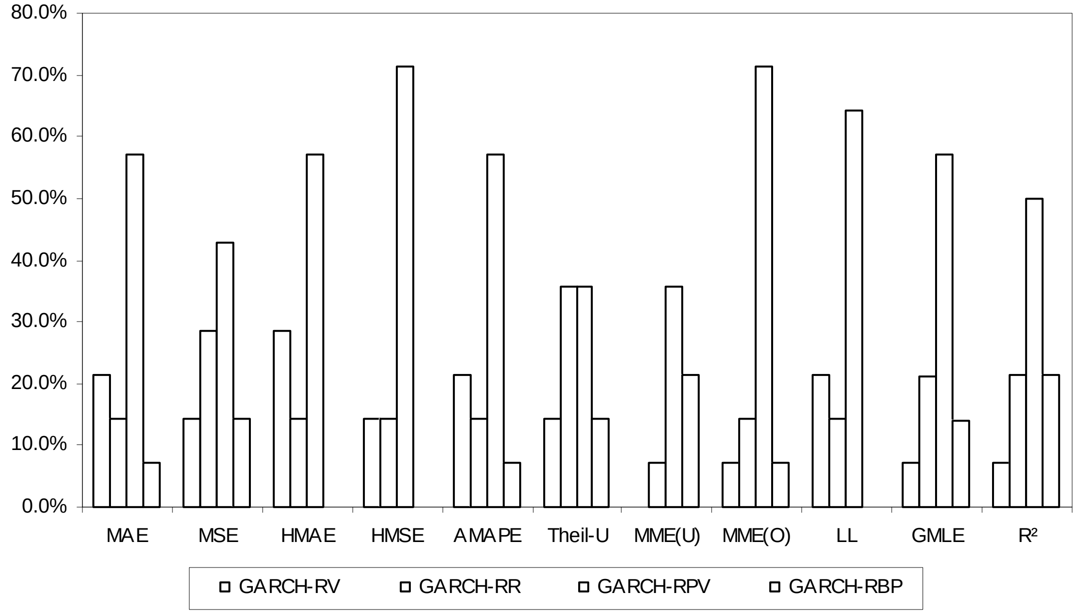



Fig. 2. Summary of forecast competition for the 14 individual NYSE stocks.

The figure plots the frequency that each model had the smallest out-of-sample loss in terms of several loss functions and the $R^2$ of the levels Mincer-Zarnowitz regression.

Table 1. Unconditional daily stock return and trading volume distribution

| | **ATT** | **AXP** | **BA** | **CAT** | **DELL** | **GE** | **GM** | **IBM** | **JPM** | **KO** | **MCD** | **MSFT** | **PG** | **WMT** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | **daily returns** (%) | | | | | | | |
| Mean | -0.095 | -0.001 | 0.012 | -0.036 | 0.102 | -0.015 | -0.106 | -0.012 | -0.015 | 0.081 | 0.083 | 0.010 | 0.130 | -0.003 |
| StDev | 2.268 | 2.165 | 1.985 | 1.998 | 2.993 | 1.866 | 1.886 | 1.927 | 2.353 | 1.653 | 1.784 | 2.184 | 1.635 | 2.005 |
| Skewness | 0.194 | -0.026 | 0.044 | 0.132 | 0.218 | 0.269 | 0.142 | 0.172 | 0.960 | 0.027 | -0.174 | 0.210 | -0.062 | 0.130 |
| Kurtosis | 4.915 | 4.392 | 5.231 | 4.334 | 6.218 | 5.293 | 4.107 | 5.150 | 14.035 | 5.951 | 6.740 | 4.048 | 7.077 | 5.497 |
| JB test | 280.3 | 274.3 | 366.1 | 135.9 | 774.0 | 407.2 | 95.9 | 472.8 | 9207.1 | 639.2 | 1035.416 | 93.7 | 1220.3 | 462.8 |
| (p-value) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| LB(10) test | ***22.189*** | **13.276** | **12.340** | **10.545** | 25.033 | **12.894** | ***18.668*** | 32.738 | **15.675** | **15.413** | **6.822** | **12.626** | 34.987 | ***21.456*** |
| (p-value) | ***(0.014)*** | **(0.209)** | **(0.263)** | **(0.394)** | (0.005) | **(0.230)** | ***(0.045)*** | (0.000) | **(0.109)** | **(0.118)** | **(0.742)** | **(0.245)** | (0.000) | ***(0.018)*** |
| | | | | | | | **daily trading volume** ($VOL \times 10^{-7}$) | | | | | | | |
| Mean | 0.247 | 0.438 | 0.317 | 0.155 | 3.950 | 1.480 | 0.293 | 0.714 | 0.695 | 0.368 | 0.419 | 6.394 | 0.266 | 0.608 |
| StDev | 2.233 | 0.215 | 0.179 | 0.084 | 3.086 | 0.667 | 0.184 | 0.343 | 0.350 | 0.170 | 0.246 | 2.743 | 0.176 | 0.252 |
| Skewness | 7.775 | 2.309 | 3.815 | 3.570 | 2.182 | 2.006 | 2.191 | 2.722 | 2.398 | 1.842 | 3.027 | 2.032 | 9.788 | 1.752 |
| Kurtosis | 128.364 | 14.241 | 29.778 | 34.609 | 9.507 | 10.277 | 10.172 | 18.462 | 15.756 | 8.883 | 19.028 | 11.460 | 189.924 | 8.817 |
| ADF test | -4.77(13) | -8.81(6) | -14.06(3) | -6.96(11) | -4.76(24) | -5.99(12) | -3.83(21) | -6.31(19) | -6.21(10) | -8.13(8) | -4.27(23) | -6.00(18) | -8.22(9) | -4.59(19) |
| Robinson d | 0.385 | 0.337 | 0.230 | 0.293 | 0.434 | 0.371 | 0.407 | 0.340 | 0.371 | 0.325 | 0.330 | 0.316 | 0.309 | 0.327 |
| KS test | 0.0407 | 0.0457 | **0.0304** | ***0.0379*** | 0.0574 | **0.0198** | **0.0277** | 0.0458 | **0.0187** | **0.0223** | **0.0191** | 0.0546 | 0.0407 | 0.0441 |
| (p-value) | (0.006) | (0.001) | **(0.075)** | ***(0.012)*** | (0.000) | **(0.489)** | **(0.132)** | (0.001) | **(0.564)** | **(0.337)** | **(0.533)** | (0.000) | (0.006) | (0.002) |

JB denotes the Jarque-Bera test statistic for the null hypothesis of normality. LB($p$) denotes the Ljung-Box Q-statistic Q-statistic for the null of no autocorrelation up to a maximum lag of $p$ days. ADF is the augmented Dickey-Fuller statistic for the null of a unit root with 5% and 1% critical values of -2.862 and -3.433, respectively. The numbers in parenthesis indicate the truncation lag chosen based on the AIC. Robinson d is the fractional differencing parameter based on Robinson's (1995) estimator. KS denotes the Kolmogorov-Smirnov statistic for the null of log-normality. Bold and bold italics denote insignificant, respectively, at the 5% (or 10%) and the 1% level.

Table 2. Unconditional daily stock volatility distributions

| | **ATT** | **AXP** | **BA** | **CAT** | **DELL** | **GE** | **GM** | **IBM** | **JPM** | **KO** | **MCD** | **MSFT** | **PG** | **WMT** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | **Squared returns**($r_t^2$) | | | | | | | |
| Mean | 5.150 | 4.683 | 3.940 | 3.991 | 8.965 | 3.481 | 3.567 | 3.714 | 5.535 | 2.740 | 3.189 | 4.765 | 2.690 | 4.021 |
| StDev | 10.141 | 9.289 | 8.108 | 7.279 | 20.529 | 7.206 | 6.251 | 7.890 | 19.972 | 6.093 | 7.610 | 8.328 | 6.597 | 8.522 |
| StDev/Mean | 1.969 | 1.983 | 2.057 | 1.823 | 2.289 | 2.070 | 1.752 | 2.124 | 3.608 | 2.223 | 2.386 | 1.747 | 2.452 | 2.119 |
| Skewness | 5.434 | 5.820 | 5.568 | 4.352 | 9.916 | 7.059 | 4.100 | 8.178 | 20.758 | 8.601 | 9.084 | 5.035 | 8.003 | 6.898 |
| Kurtosis | 46.480 | 56.130 | 47.390 | 29.840 | 160.129 | 82.450 | 27.220 | 113.350 | 586.580 | 124.030 | 121.660 | 44.530 | 92.770 | 87.457 |
| ADF | -10.27(8) | -5.64(23) | -24.83(1) | -8.76(12) | -10.81(8) | -6.72(18) | -7.07(15) | -12.76(6) | -31.48(0) | -10.07(9) | -35.30(0) | -17.51(3) | -5.95(23) | -6.19(18) |
| Robinson d | 0.180 | 0.269 | 0.139 | 0.175 | 0.197 | 0.086 | 0.294 | 0.152 | 0.080 | 0.174 | 0.096 | 0.129 | 0.163 | 0.167 |
| KS | 0.0591 | 0.0731 | 0.0479 | 0.0684 | 0.053 | 0.0696 | 0.0647 | 0.0777 | 0.0653 | 0.066 | 0.0509 | 0.08 | 0.0631 | 0.0634 |
| (p-value) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | | | | | | | **Realised variance(RV)** | | | | | | | |
| Mean | 4.506 | 4.673 | 4.078 | 3.782 | 8.170 | 3.648 | 3.021 | 3.572 | 5.622 | 2.836 | 3.550 | 4.441 | 2.905 | 4.140 |
| StDev | 4.356 | 5.065 | 3.950 | 3.176 | 7.689 | 3.776 | 2.950 | 3.663 | 7.717 | 2.417 | 3.337 | 3.892 | 3.080 | 4.298 |
| StDev/Mean | 0.966 | 1.083 | 0.968 | 0.839 | 0.941 | 1.035 | 0.976 | 1.025 | 1.372 | 0.852 | 0.940 | 0.876 | 1.470 | 1.038 |
| Skewness | 4.190 | 4.673 | 6.773 | 3.054 | 3.570 | 4.629 | 4.145 | 6.442 | 8.931 | 3.206 | 4.511 | 3.486 | 5.152 | 6.340 |
| Kurtosis | 31.610 | 32.030 | 105.240 | 17.800 | 25.870 | 38.220 | 33.070 | 92.930 | 136.670 | 19.070 | 35.830 | 24.560 | 49.260 | 90.930 |
| ADF | -4.73(18) | -5.20(17) | -8.63(7) | -4.68(18) | -3.75(24) | -5.48(20) | -6.70(7) | -5.04(18) | -5.31(20) | -4.85(21) | -5.68(15) | -5.60(16) | -4.49(21) | -5.72(13) |
| Robinson d | 0.345 | 0.358 | 0.337 | 0.409 | 0.373 | 0.330 | 0.385 | 0.371 | 0.318 | 0.368 | 0.329 | 0.387 | 0.358 | 0.350 |
| KS | **0.0181** | **0.0253** | **0.0199** | **0.0203** | **0.0284** | ***0.0387*** | ***0.0331*** | **0.0211** | **0.0234** | **0.0265** | **0.0319** | **0.0226** | **0.0215** | **0.0225** |
| (p-value) | **(0.606)** | **(0.207))** | **(0..481)** | **(0.455)** | **(0.114)** | **(0.010)** | ***(0.042)*** | **(0.408)** | **(0.284)** | **(0.168)** | **(0.055)** | **(0.323)** | **(0.382)** | **(0.327)** |
| | | | | | | | **Realised range(RR)** | | | | | | | |
| Mean | 3.014 | 2.859 | 2.748 | 2.131 | 10.028 | 2.741 | 1.745 | 2.525 | 3.839 | 1.889 | 2.554 | 5.243 | 1.871 | 2.648 |
| StDev | 2.656 | 2.711 | 2.284 | 2.014 | 9.300 | 2.345 | 1.624 | 2.119 | 4.408 | 1.363 | 2.124 | 4.285 | 1.697 | 2.137 |
| StDev/Mean | 0.881 | 0.948 | 0.831 | 0.945 | 0.927 | 0.855 | 0.930 | 0.383 | 1.148 | 0.721 | 0.831 | 0.817 | 0.907 | 0.807 |
| Skewness | 3.671 | 3.239 | 3.509 | 11.102 | 3.830 | 3.765 | 3.357 | 3.557 | 7.050 | 2.624 | 3.662 | 3.144 | 4.556 | 2.872 |
| Kurtosis | 26.610 | 20.020 | 28.630 | 259.890 | 33.470 | 31.970 | 20.650 | 30.240 | 90.890 | 14.690 | 25.890 | 20.260 | 43.510 | 18.090 |
| ADF | -4.60(18) | -4.66(17) | -7.01(7) | -5.28(13) | -3.43(24) | -4.85(20) | -5.65(9) | -4.36(18) | -4.82(23) | -4.53(15) | -5.91(11) | -4.88(18) | -4.31(20) | -4.95(12) |
| Robinson d | 0.370 | 0.404 | 0.391 | 0.381 | 0.396 | 0.382 | 0.426 | 0.406 | 0.350 | 0.388 | 0.364 | 0.400 | 0.382 | 0.397 |
| KS | **0.0148** | **0.0239** | **0.0218** | **0.0199** | **0.0136** | **0.0302** | 0.0394 | **0.0211** | **0.0145** | **0.0202** | **0.0126** | ***0.0348*** | **0.0316** | **0.0135** |
| (p-value) | **(0.827)** | **(0.265)** | **(0.363)** | **(0.479)** | **(0.898)** | **(0.079)** | (0.008) | **(0.403)** | **(0.850)** | **(0.462)** | **(0.938)** | ***(0.028)*** | **(0.058)** | **(0.901)** |
| | | | | | | | **Realised power variation(RPV)** | | | | | | | |
| Mean | 8.680 | 8.863 | 8.111 | 7.468 | 13.888 | 7.573 | 6.356 | 7.408 | 10.112 | 6.315 | 7.315 | 8.887 | 6.280 | 8.121 |
| StDev | 5.679 | 6.351 | 5.064 | 4.471 | 9.158 | 5.145 | 4.172 | 4.770 | 8.256 | 3.741 | 4.486 | 5.408 | 4.293 | 5.491 |
| StDev/Mean | 0.654 | 0.716 | 0.624 | 0.598 | 0.659 | 0.679 | 0.656 | 0.643 | 0.816 | 0.592 | 0.613 | 0.608 | 0.683 | 0.676 |
| Skewness | 2.832 | 2.659 | 3.130 | 2.161 | 2.396 | 2.978 | 2.469 | 2.367 | 4.875 | 2.293 | 2.895 | 2.252 | 2.856 | 2.832 |
| Kurtosis | 17.017 | 14.370 | 24.786 | 10.819 | 13.350 | 17.554 | 12.818 | 13.373 | 47.943 | 11.694 | 17.751 | 12.366 | 17.600 | 20.182 |
| ADF | -4.42(18) | -4.71(17) | -6.79(10) | -4.73(12) | -3.53(24) | -4.98(20) | -5.87(9) | -4.37(18) | -4.66(20) | -4.21(18) | -5.71(12) | -5.25(16) | -5.34(12) | -4.93(13) |
| Robinson d | 0.370 | 0.389 | 0.377 | 0.421 | 0.392 | 0.363 | 0.409 | 0.394 | 0.380 | 0.385 | 0.365 | 0.392 | 0.396 | 0.393 |
| KS | **0.0156** | **0.0245** | **0.0233** | 0.0430 | **0.0117** | ***0.0333*** | **0.0302** | **0.0204** | **0.0219** | **0.0220** | **0.0246** | **0.0222** | **0.0295** | **0.0167** |
| (p-value) | **(0.779)** | **(0.238)** | **(0.286)** | (0.003) | **(0.968)** | ***(0.040)*** | **(0.078)** | **(0.449)** | **(0.358)** | **(0.354)** | **(0.232)** | **(0.343)** | **(0.091)** | **(0.706)** |
| | | | | | | | **Realised bipower variation(RBP)** | | | | | | | |
| Mean | 4.090 | 4.292 | 3.702 | 3.401 | 7.652 | 3.416 | 2.734 | 3.340 | 5.175 | 2.598 | 3.252 | 4.146 | 2.676 | 3.724 |
| StDev | 4.252 | 4.732 | 3.775 | 3.048 | 7.480 | 3.710 | 2.783 | 3.210 | 7.024 | 2.369 | 3.263 | 3.753 | 2.879 | 3.888 |
| StDev/Mean | 1.039 | 1.102 | 1.019 | 0.896 | 0.977 | 1.086 | 1.0179 | 0.961 | 1.357 | 0.911 | 1.003 | 0.905 | 1.075 | 1.044 |
| Skewness | 4.506 | 3.931 | 7.403 | 3.529 | 3.656 | 4.945 | 3.818 | 6.442 | 7.799 | 3.549 | 4.798 | 3.658 | 5.253 | 4.961 |
| Kurtosis | 34.320 | 25.60 | 124.150 | 23.790 | 25.350 | 42.870 | 25.730 | 92.930 | 101.450 | 22.800 | 40.060 | 28.230 | 51.290 | 51.810 |
| ADF | -4.79(18) | -6.46(11) | -8.67(7) | -4.71(19) | -3.82(24) | -5.55(24) | -6.69(7) | -7.00(7) | -4.98(21) | -5.00(21) | -6.33(12) | -6.96(6) | -4.42(21) | -5.01(17) |
| Robinson d | 0.340 | 0.364 | 0.336 | 0.402 | 0.364 | 0.327 | 0.387 | 0.383 | 0.331 | 0.363 | 0.328 | 0.387 | 0.350 | 0.361 |
| KS | **0.0180** | **0.0267** | **0.0170** | **0.0226** | **0.0225** | ***0.0359*** | **0.0262** | **0.0188** | **0.0229** | **0.0290** | **0.0233** | **0.0240** | **0.0221** | **0.0177** |
| (p-value) | **(0.613)** | **(0.159)** | **(0.683)** | **(0.324)** | **(0.329)** | ***(0.021)*** | **(0.174)** | **(0.551)** | **(0.309)** | **(0.102)** | **(0.288)** | **(0.259)** | **(0.351)** | **(0.634)** |

See note in Table 1. The daily RV, RR, RPV and RBP are based on prices sampled at the 5-min frequency. RPV is computed for z=1.5.

Table 3. Normality tests for daily returns and standardized daily returns

| | **ATT** | **AXP** | **BA** | **CAT** | **DELL** | **GE** | **GM** | **IBM** | **JPM** | **KO** | **MCD** | **MSFT** | **PG** | **WMT** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | $r_t$ | | | | | | |
| JB | 280.28 | 274.28 | 366.12 | 135.87 | 774.04 | 407.18 | 95.88 | 472.83 | 9207.09 | 639.23 | 1035.42 | 93.68 | 1220.25 | 462.83 |
| p-value | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | | | | | | | | $r_t/\sqrt{\mathbf{GARCH}_t}$ | | | | | | |
| JB | 214.92 | 76.34 | 175.91 | 67.52 | 246.85 | 146.56 | 19.89 | 105.05 | 268.85 | 169.67 | 931.94 | 43.82 | 313.67 | 53.33 |
| p-value | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | | | | | | | | $r_t/\sqrt{\mathbf{RV}_t}$ | | | | | | |
| JB | ***6.81*** | 12.80 | **4.43** | **4.65** | 12.25 | 14.66 | 12.24 | 15.42 | ***8.94*** | **4.42** | **2.52** | 22.43 | **4.94** | **5.66** |
| p-value | **(0.033)** | (0.000) | **(0.109)** | **(0.097)** | (0.002) | (0.000) | (0.002) | (0.000) | **(0.011)** | **(0.109)** | **(0.283)** | (0.000) | **(0.084)** | **(0.058)** |
| | | | | | | | | $r_t/\sqrt{\mathbf{RR}_t}$ | | | | | | |
| JB | **2.16** | ***6.34*** | **0.05** | **1.45** | ***10.19*** | ***8.49*** | **2.59** | **3.04** | **1.40** | **4.34** | **3.89** | 22.82 | **1.23** | **5.36** |
| p-value | **(0.339)** | ***(0.042)*** | **(0.977)** | **(0.485)** | ***(0.010)*** | ***(0.014)*** | **(0.274)** | **(0.219)** | **(0.497)** | **(0.111)** | **(0.143)** | (0.000) | **(0.540)** | **(0.069)** |
| | | | | | | | | $r_t/\sqrt{\mathbf{RPV}_t}$ | | | | | | |
| JB | **5.68** | 9.35 | **1.10** | **3.34** | ***6.79*** | 13.23 | 12.73 | 9.36 | ***7.37*** | **2.81** | **0.29** | 19.61 | **0.41** | **2.87** |
| p-value | **(0.058)** | (0.009) | **(0.576)** | **(0.188)** | ***(0.033)*** | (0.001) | (0.001) | (0.009) | ***(0.025)*** | **(0.244)** | **(0.867)** | (0.000) | **(0.816)** | **(0.283)** |
| | | | | | | | | $r_t/\sqrt{\mathbf{RBP}_t}$ | | | | | | |
| JB | **4.71** | 10.72 | **3.31** | **3.69** | 11.21 | 17.35 | 10.84 | 12.14 | ***6.10*** | ***7.36*** | **1.38** | 21.51 | **2.87** | **4.26** |
| p-value | **(0.095)** | (0.005) | **(0.191)** | **(0.158)** | (0.004) | (0.000) | (0.004) | (0.002) | ***(0.047)*** | ***(0.023)*** | **(0.501)** | (0.000) | **(0.241)** | **(0.118)** |

RV, RR, RBP and RPV are calculated using prices sampled at the 5-min frequency. RPV is calculated using power order $p = 1.5$. JB is the Jarque-Bera test. Bold and bold italics indicate that the normality null cannot be rejected at the 5% (or 10%) and the 1% level, respectively
.

Table 4. GARCH estimation results from 02/01/97 to 31/12/03

| | **CAT (least traded stock)** | | | | | | **MSFT (most traded stock)** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GARCH | with RV | with RR | with RPV | with RBP | with VOL | GARCH | with RV | with RR | with RPV | with RBP | with VOL |
| $\theta_0$ | 0.0200 | -0.0084 | -0.0146 | -0.0077 | -0.0087 | 0.0260 | 0.0213 | -0.0095 | -0.0087 | -0.0074 | -0.0128 | 0.0251 |
| | *(0.0437)* | *(0.0439)* | *(0.0444)* | *(0.0438)* | *(0.0444)* | *(0.0450)* | *(0.0477)* | *(0.0461)* | *(0.0465)* | *(0.0472)* | *(0.0473)* | *(0.0484)* |
| $\omega$ | 0.0366 | **0.3260** | **0.6074** | 0.0903 | **0.3092** | **0.0894** | **0.2384** | **0.5931** | **0.5774** | 0.0298 | **0.5970** | 0.2866 |
| | *(0.0231)* | *(0.1426)* | *(0.2241)* | *(0.1151)* | *(0.1329)* | *(0.0329)* | *(0.0897)* | *(0.1852)* | *(0.2149)* | *(0.1656)* | *(0.1792)* | *(0.2080)* |
| $\alpha_1$ | **0.1108** | 0.0559 | **0.0788** | 0.0577 | 0.0645 | **0.1004** | **0.1015** | 0.0278 | 0.0170 | 0.0302 | 0.0426 | **0.1379** |
| | *(0.0401)* | *(0.0412)* | *(0.0420)* | *(0.0421)* | *(0.0431)* | *(0.0220)* | *(0.0217)* | *(0.0301)* | *(0.0265)* | *(0.0258)* | *(0.0259)* | *(0.0262)* |
| $\alpha_2$ | **-0.0844** | -0.0472 | -0.0380 | -0.0543 | -0.0544 | **-0.0768** | | | | | | |
| | *(0.0398)* | *(0.0358)* | *(0.0342)* | *(0.0351)* | *(0.0375)* | *(0.0227)* | | | | | | |
| $\beta_1$ | **0.9646** | **0.6891** | **0.5760** | **0.6871** | **0.7261** | **0.9656** | **0.8507** | **0.3730** | **0.2258** | **0.4006** | **0.4028** | **0.5585** |
| | *(0.0124)* | *(0.0996)* | *(0.1250)* | *(0.1033)* | *(0.0888)* | *(0.0094)* | *(0.0343)* | *(0.1111)* | *(0.0975)* | *(0.0775)* | *(0.0778)* | *0.0625* |
| $\gamma$ | | **0.2350** | **0.4317** | **0.1494** | **0.2211** | **-2.97e-08** | | **0.5176** | **0.5889** | **0.3013** | **0.5026** | **1.83e-08** |
| | | *(0.0823)* | *(0.1498)* | *(0.0523)* | *(0.0790)* | *(1.05e-08)* | | *(0.1108)* | *(0.0811)* | *(0.0431)* | *(0.0760)* | *(4.55e-09)* |
| $lnL$ | -3670.44 | -3643.614 | -3650.790 | -3641.110 | -3645.447 | -3663.152 | -3813.564 | -3762.681 | -3753.741 | -3761.323 | -3763.916 | -3810.397 |
| $AIC$ | 4.1743 | 4.1473 | 4.1554 | 4.1444 | 4.1494 | 4.1695 | 4.3357 | 4.2815 | 4.2713 | 4.2799 | 4.2829 | 4.3357 |
| $SBC$ | 4.1898 | 4.1659 | 4.1741 | 4.1631 | 4.1680 | 4.1882 | 4.3481 | 4.2970 | 4.2868 | 4.2955 | 4.2984 | 4.3512 |
| $LB(10)$ | 0.463 | 0.231 | 0.264 | 0.230 | 0.206 | 0.563 | 0.549 | 0.450 | 0.372 | 0.453 | 0.451 | 0.491 |
| $ARCH(10)$ | 0.488 | 0.481 | 0.460 | 0.478 | 0.508 | 0.451 | 0.342 | 0.714 | 0.454 | 0.660 | 0.651 | 0.525 |
| *Ranking* | 6 | 2 | 4 | 1 | 3 | 5 | 5 or 6 | 3 | 1 | 2 | 4 | 6 or 5 |

The reported estimation results are for the conditional mean equation $r_t = \theta_0 + u_t, u_t|\mathcal{F}_{t-1} \sim iid(0, h_t)$, and conditional variance equation $h_t = \omega + \alpha_1 u_{t-1}^2 + ... + \alpha_r u_{t-r}^2 + \beta_1 h_{t-1} + ... + \beta_s h_{t-s} + \gamma v_{t-1}$. $LB(10)$ is the Ljung-Box test for autocorrelation in $u_t$. $ARCH(10)$ is Engle's ARCH LM test. The intraday sampling frequency for RV, RR, RPV, RBP is 5 min. Bold denotes significance at the 1%, 5% or 10% level. Bollerslev-Wooldridge s.e. are reported in parentheses. The degree of volatility persistence is given by $\lambda = \sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j$. Rankings 1 and 6 mean top and bottom, respectively, according to in-sample model fit.

Table 5. Out-of-sample forecast evaluation

| | | Forecast accuracy measures | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Stock** | **Model** | MAE | MSE | HMAE | HMSE | AMAPE | Theil-U | MME(U) | MME(O) | LL | GMLE | MZ-$R^2$ |
| | GARCH | 3.0166 | 22.4841*** | 0.9492 | 2.2441 | 0.2831 | 0.8428 | 59.3852 | 13.6460 | 0.5699 | 2.5595 | 28.8653 |
| | GARCH-RV | 2.6179*** | 20.3382* | 0.7629*** | 1.3708** | 0.2442*** | 0.7624* | 48.8266 | 13.5681*** | 0.4401 | 2.5101*** | 36.9995** |
| | GARCH-RR | 2.8303*** | 21.3482** | 0.8636*** | 1.7020*** | 0.2626*** | 0.8002** | 52.8509** | 15.6177*** | 0.4941*** | 2.5215*** | 35.8314* |
| **ATT** | GARCH-RPV | **2.5203** | **18.8451** | **0.7224** | **1.1681** | **0.2390** | **0.7064** | 49.3295 | **11.0480** | **0.4118** | **2.5006** | **39.6859** |
| | GARCH-RBP | 2.6583*** | 20.8671** | 0.7779*** | 1.4241** | 0.2467*** | 0.7822** | **46.8354** | 15.0448*** | 0.4483*** | 2.5118*** | 37.2140** |
| | GARCH-VOL | 5.2470*** | 47.8291*** | 2.2905*** | 17.3949*** | 0.3869*** | 1.7781*** | 68.8293** | 47.6280*** | 1.3396*** | 2.7446*** | 2.5456*** |
| | COMBINED | 2.4400 | 20.0675 | 0.6141 | 0.8372 | 0.2223 | 0.7522 | 48.9171 | 10.8525 | 0.3564 | 2.4891 | 35.1258 |
| | *Benefit* (%) | 16.45 | 16.18 | 23.89 | 47.95 | 15.58 | 16.18 | 21.13 | 19.04 | 27.74 | 2.30 | 37.49 |
| | GARCH | 2.0818 | 16.4060*** | 0.8336 | 1.9394 | 0.2572 | 1.1495 | 49.3419 | 7.7150 | 0.4831 | 2.2114 | 29.0442 |
| | GARCH-RV | 1.7938*** | 10.6703** | 0.6812*** | 1.0788** | 0.2263 | 0.7476** | **28.7188** | 7.0649 | 0.3497* | 2.1425*** | 54.8918 |
| | GARCH-RR | 2.2248*** | 12.3740*** | 0.9791** | 2.0343*** | 0.2794* | 0.8668*** | 41.7935 | 9.6238*** | 0.5314*** | 2.1975*** | 53.2825 |
| **AXP** | GARCH-RPV | **1.6658** | **10.3212** | **0.6106** | **0.8284** | **0.2113** | **0.7230** | 29.6144 | **5.6148** | **0.3079** | **2.1287** | **56.9665** |
| | GARCH-RBP | 1.7594** | 10.4086 | 0.6790** | 1.0871** | 0.2242** | 0.7291 | 28.9881 | 6.8762 | 0.3487** | 2.1413** | 56.1080 |
| | GARCH-VOL | 2.3054** | 18.2394*** | 1.5165** | 38.9584*** | 0.2708*** | 1.2780** | 49.7656* | 10.3297* | 0.7109** | 2.2552** | 22.2397** |
| | COMBINED | 1.6789 | 10.3938 | ***0.5053*** | ***0.5119*** | ***0.2037*** | ***0.6963*** | ***22.4327*** | 6.2478 | 0.3227 | 2.2939 | 56.0890 |
| | *Benefit* (%) | 19.98 | 37.09 | 26.75 | 57.29 | 17.84 | 37.10 | 41.80 | 27.22 | 36.27 | 3.74 | 96.14 |
| | GARCH | 1.6993 | 7.8206*** | 0.4806 | 0.5322 | 0.1972 | 1.1670 | 15.2132 | 4.4878 | 0.2804 | 2.3894 | 35.4082 |
| | GARCH-RV | **1.4898** | **5.4187** | **0.4530** | 0.4560 | **0.1790** | **0.8086** | **11.1988** | 4.2482* | 0.2251*** | 2.3501 | 54.4613** |
| | GARCH-RR | 1.7047** | 5.8298** | 0.5712*** | 0.7002*** | 0.2010** | 0.8698** | 12.3551 | 5.7955*** | 0.2817*** | 2.3638*** | 53.5012* |
| **BA** | GARCH-RPV | 1.5012 | 5.4405 | 0.4562 | **0.4455** | 0.1799 | 0.8118 | 11.6785 | **4.0773** | **0.2241** | **2.3491** | **55.7621** |
| | GARCH-RBP | 1.5287** | 5.4633 | 0.4694*** | 0.4866** | 0.1826** | 0.8152 | 11.4482 | 4.5720*** | 0.2312** | 2.3515 | 53.6656** |
| | GARCH-VOL | 2.3346*** | 11.1985** | 0.7861*** | 1.6217*** | 0.2577*** | 1.6705** | 22.2042* | 8.8998*** | 0.4785* | 2.4588*** | 8.4811* |
| | COMBINED | 1.7468 | 6.0563 | 0.5482 | 0.5976 | 0.1990 | 0.9036 | 11.5966 | 6.3922 | 0.2674 | 2.3611 | 53.7150 |
| | *Benefit* (%) | 12.33 | 30.71 | 5.74 | 16.29 | 9.23 | 30.71 | 26.39 | 9.15 | 20.08 | 1.69 | 57.48 |
| | GARCH | 1.7797 | 6.1406*** | 0.9471 | 1.8327 | 0.2811 | 1.2597 | 16.8100 | 5.7511 | 0.5316 | 2.0682 | 13.9861 |
| | GARCH-RV | 1.3050** | 3.8605 | 0.6660** | 0.9716** | 0.2215* | 0.7921 | **11.2370** | **3.6301** | 0.3350*** | 1.9917 | 42.6145* |
| | GARCH-RR | 1.5758** | 4.7520*** | 0.8491*** | 1.4886*** | 0.2580 | 0.9750*** | 14.7864** | 4.8381*** | 0.4481** | 2.0293* | 37.0931** |
| **CAT** | GARCH-RPV | **1.2880** | **3.8056** | **0.6340** | **0.8451** | **0.2165** | **0.7809** | 12.0083 | 3.6810 | **0.3130** | **1.9850** | **43.6401** |
| | GARCH-RBP | 1.3401 | 3.9384*** | 0.6808*** | 0.9656 | 0.2263*** | 0.8081*** | 12.1031** | 3.8194** | 0.3420 | 1.9945 | 42.4124** |
| | GARCH-VOL | 2.2153*** | 8.9839*** | 1.2814*** | 4.6883*** | 0.3117*** | 1.8430*** | 19.9459** | 9.2635** | 0.7403** | 2.1294*** | 4.9856* |
| | COMBINED | 1.1176 | ***3.5523*** | 0.4617 | 0.5100 | 0.1813 | 0.7287 | 7.7120 | 3.2263 | 0.2373 | 1.9708 | 43.8049 |
| | *Benefit* (%) | 27.63 | 38.03 | 34.91 | 53.88 | 22.98 | 38.01 | 33.15 | 36.88 | 41.21 | 4.04 | 212.03 |
| | GARCH | 2.2559 | 10.2684*** | 0.6453 | 0.8808 | 0.2217 | 1.2250 | 17.7316 | 10.6282 | 0.3252 | 2.4342 | 42.3244 |
| | GARCH-RV | 2.3391*** | 9.7267*** | 0.6947*** | 0.9314*** | 0.2307*** | 1.1614*** | 19.2699** | 10.7855*** | 0.3372*** | 2.4332*** | 51.5926 |
| | GARCH-RR | **1.9811** | **8.0345** | **0.5513** | **0.6249** | **0.1983** | **0.9562** | 16.5241 | **8.1202** | **0.2569** | **2.4068** | **51.6494** |
| **DELL** | GARCH-RPV | 2.3758*** | 9.9858*** | 0.6863*** | 0.9127*** | 0.2311*** | 1.1919*** | **15.8615** | 11.2197*** | 0.3393*** | 2.4371*** | 51.2849 |
| | GARCH-RBP | 2.3606*** | 10.0715*** | 0.7020*** | 0.9877*** | 0.2311*** | 1.2025*** | 18.3796* | 11.0655*** | 0.3454*** | 2.4365*** | 49.6899*** |
| | GARCH-VOL | 7.3802*** | 119.4348*** | 2.1633*** | 13.8656*** | 0.3659*** | 14.2745*** | 16.1837 | 149.0813*** | 1.2698*** | 2.6959*** | 12.8040** |
| | COMBINED | 1.6742 | ***6.9393*** | 0.4129 | 0.3455 | 0.1767 | 0.8275 | 11.7587 | 6.6457 | 0.2293 | 2.4282 | 54.1353 |
| | *Benefit* (%) | 41.80 | 21.76 | 14.57 | 29.05 | 10.55 | 21.94 | 10.55 | 23.60 | 21.00 | 1.13 | 22.06 |

The target is the daily conditional variance proxied by the sum of 5-min squared returns. Bold indicates the top performer. Asterisks denote that the forecasts of the model are significantly worse (DM test), * at 10% , ** 5%, *** or 1% level, than those of the top performer. The GARCH vs augmented-GARCH comparison is based on the ENC-T test for nested models developed for the MSE loss. Likewise for the pairwise comparison of intraday-volatility.agurmented GARCH and the COMBINED model. For the latter, italics bold (under MSE) indicates that it beats the best individual forecast at the 10%, 5% or 1% level.

Table 5. Out-of-sample forecast evaluation (*cont.*)

| Stock | Model | Forecast accuracy measures | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | MAE | MSE | HMAE | HMSE | AMAPE | Theil-U | MME(U) | MME(O) | LL | GMLE | MZ-$R^2$ |
| | GARCH | 1.7272 | 10.1324$^{***}$ | 0.6340 | 0.8646 | 0.2287 | 0.9502 | 29.0961 | 5.0890 | 0.3619 | 2.1341 | 41.1487 |
| | GARCH-RV | 1.5628$^{***}$ | 7.9192 | 0.5758$^{**}$ | 0.7745 | 0.2099 | 0.7426 | 19.9198$^{*}$ | 5.1772$^{**}$ | 0.3023 | 2.0969 | 52.5671$^{*}$ |
| | GARCH-RR | 1.6990$^{***}$ | **7.6688** | 0.7016$^{***}$ | 1.0634 | 0.2327$^{***}$ | **0.7191** | 23.0784$^{**}$ | 5.9512$^{***}$ | 0.3675$^{***}$ | 2.1128$^{**}$ | **55.5787** |
| **GE** | GARCH-RPV | **1.4915** | 7.7464 | **0.5167** | **0.5886** | **0.1972** | 0.7264 | 19.8920 | **4.4162** | **0.2648** | **2.0835** | 55.2760 |
| | GARCH-RBP | 1.5755$^{***}$ | 7.9490 | 0.5757$^{**}$ | 0.7466 | 0.2109 | 0.7455 | **18.6959** | 5.4607$^{**}$ | 0.3007 | 2.0961 | 52.0659$^{**}$ |
| | GARCH-VOL | 2.0702$^{***}$ | 10.2605$^{***}$ | 0.8652$^{***}$ | 1.4644$^{***}$ | 0.2701$^{***}$ | 0.9615$^{***}$ | 30.9637$^{**}$ | 7.5387$^{***}$ | 0.4812$^{***}$ | 2.1544$^{***}$ | 41.9056$^{***}$ |
| | COMBINED | 1.5448 | ***7.3832*** | 0.4780 | 0.4221 | 0.2166 | 0.6840 | 13.0771 | 5.5854 | 0.3906 | 2.2940 | 55.6449 |
| | *Benefit* (%) | 13.65 | 23.55 | 18.50 | 31.92 | 13.77 | 24.32 | 35.74 | 13.22 | 26.83 | 2.37 | 35.07 |
| | GARCH | 1.4905 | 4.8797$^{***}$ | 0.7436 | 1.3186 | 0.2366 | 0.7623 | 11.9501 | 4.9465 | 0.3872 | 1.9968 | 53.5156 |
| | GARCH-RV | 1.4402 | 4.4830 | **0.7256** | 1.2516 | 0.2332 | 0.7004 | 11.9056$^{*}$ | 4.6104 | **0.3757** | 1.9934 | 57.3629$^{*}$ |
| | GARCH-RR | 1.9465$^{***}$ | 6.4279$^{***}$ | 1.0046$^{***}$ | 2.0327$^{***}$ | 0.2880$^{***}$ | 1.0043$^{***}$ | 13.0465$^{*}$ | 7.7755$^{****}$ | 0.5385$^{***}$ | 2.0432$^{**}$ | 57.0323$^{*}$ |
| **GM** | GARCH-RPV | 1.4508 | **4.4050** | 0.7296 | **1.2294** | 0.2350$^{*}$ | **0.6882** | 11.9278 | **4.4331** | 0.3768 | 1.9942 | 58.8185 |
| | GARCH-RBP | **1.4400** | 4.4190 | 0.7320 | 1.3035$^{*}$ | **0.2325** | 0.6903 | **10.9091** | 4.8677 | 0.3776$^{**}$ | **1.9925** | **59.0001** |
| | GARCH-VOL | 1.5957$^{**}$ | 5.8561$^{***}$ | 0.7498 | 1.4247 | 0.2397 | 0.9150$^{***}$ | 12.7515 | 5.8363$^{***}$ | 0.4093 | 2.0122$^{*}$ | 45.5787$^{***}$ |
| | COMBINED | 1.2868 | ***4.1314*** | 0.5276 | 0.6028 | 0.2030 | 0.6457 | 8.8911 | 3.9220 | 0.2859 | 1.9846 | 58.4042 |
| | *Benefit* (%) | 3.39 | 9.73 | 2.42 | 6.76 | 1.73 | 9.72 | 8.71 | 10.38 | 2.97 | 0.215 | 10.25 |
| | GARCH | 1.4223 | 4.0962$^{***}$ | 0.8016 | 1.1481 | 0.2601 | 1.3608 | 13.2574 | 4.3714 | 0.4228 | 1.8151 | 42.7414 |
| | GARCH-RV | 1.1013$^{**}$ | 2.5202$^{**}$ | 0.6505$^{**}$ | 0.8326$^{***}$ | 0.2199 | 0.8369$^{**}$ | 6.4769 | 3.1139$^{***}$ | 0.3195 | 1.7721$^{**}$ | 63.9620$^{**}$ |
| | GARCH-RR | 1.2717$^{***}$ | 2.8058$^{***}$ | 0.7823$^{***}$ | 1.1345$^{***}$ | 0.2483$^{***}$ | 0.9315$^{***}$ | 7.5023$^{**}$ | 3.7029$^{***}$ | 0.3981$^{***}$ | 1.7983$^{***}$ | 64.9736 |
| **IBM** | GARCH-RPV | **1.0543** | **2.3986** | **0.5879** | **0.6620** | **0.2070** | **0.7965** | **6.4115** | **2.7681** | **0.2784** | **1.7580** | **65.9708** |
| | GARCH-RBP | 1.1041$^{***}$ | 2.6050$^{**}$ | 0.6470$^{*}$ | 0.8170$^{*}$ | 0.2193 | 0.8651$^{**}$ | 6.6680 | 3.0939$^{***}$ | 0.3174$^{*}$ | 1.7717$^{**}$ | 62.7126$^{*}$ |
| | GARCH-VOL | 4.7177$^{***}$ | 27.0749$^{***}$ | 3.7685$^{***}$ | 31.2741$^{***}$ | 0.5174$^{***}$ | 8.9977$^{***}$ | 19.8335$^{**}$ | 30.7368$^{***}$ | 2.1877$^{***}$ | 2.2842$^{***}$ | 3.4678$^{**}$ |
| | COMBINED | 1.0223 | ***2.3060*** | 0.4975 | 0.4563 | 0.2139 | 0.7549 | 4.2532 | 2.8345 | 0.4336 | 2.1966 | 66.1373 |
| | *Benefit* (%) | 25.87 | 41.44 | 26.66 | 42.34 | 20.42 | 41.47 | 51.64 | 36.68 | 34.15 | 3.16 | 54.35 |
| | GARCH | 3.3925 | 57.9052$^{***}$ | 0.7523 | 1.1999 | 0.2506 | 0.9217 | **107.8462** | 44.1741 | 0.4308 | 2.4915 | 51.6710 |
| | GARCH-RV | 2.5663 | 52.7961 | 0.4741$^{**}$ | 0.4968$^{**}$ | 0.1869$^{*}$ | 0.8404 | 114.3802 | 23.3971$^{*}$ | 0.2504 | 2.4202$^{***}$ | 52.7068 |
| | GARCH-RR | 2.7198$^{***}$ | **49.2440** | 0.5943$^{***}$ | 0.6837$^{***}$ | 0.2132$^{*}$ | **0.7822** | 156.7707 | 16.4277$^{*}$ | 0.3080$^{***}$ | 2.4389$^{***}$ | 59.1868 |
| **JPM** | GARCH-RPV | **2.3625** | 50.1563 | **0.4191** | **0.3293** | **0.1777** | 0.7968 | 118.1784 | **9.3414** | **0.2216** | **2.4183** | **64.2387** |
| | GARCH-RBP | 2.5247 | 50.3707 | 0.4675$^{**}$ | 0.4917$^{*}$ | 0.1847 | 0.8018 | 110.5105 | 20.1947$^{*}$ | 0.2468 | 2.4188$^{**}$ | 55.5118 |
| | GARCH-VOL | 4.1993$^{***}$ | 63.0839$^{***}$ | 1.2677$^{***}$ | 10.7165$^{***}$ | 0.3127$^{**}$ | 1.0040 | 138.3577 | 47.5825$^{**}$ | 0.7148$^{***}$ | 2.5738$^{***}$ | 51.2399$^{**}$ |
| | COMBINED | 3.0559 | 49.9339 | 0.4982 | 0.4441 | 0.2380 | 0.6715 | 95.0156 | 20.3398 | 0.7358 | 4.8830 | 62.7711 |
| | *Benefit* (%) | 30.36 | 14.96 | 44.29 | 72.56 | 29.10 | 15.14 | 0.00 | 78.85 | 48.60 | 2.94 | 24.32 |
| | GARCH | 0.8576 | 2.7742$^{***}$ | 0.4397 | 0.4438 | 0.1842 | 1.3925 | 6.0087 | 1.8582 | 0.2427 | 1.6508 | 44.5086 |
| | GARCH-RV | **0.7269** | **1.6299** | **0.3880** | **0.3010** | **0.1646** | **0.8177** | **3.3919** | 1.6673 | **0.1834** | 1.6147 | **66.8741** |
| | GARCH-RR | 0.8568$^{*}$ | 1.7987$^{*}$ | 0.4953$^{***}$ | 0.4605$^{***}$ | 0.1893$^{**}$ | 0.9025$^{*}$ | 4.1824$^{*}$ | 2.0507$^{*}$ | 0.2312$^{***}$ | 1.6286$^{***}$ | 64.8503 |
| **KO** | GARCH-RPV | 0.7508$^{*}$ | 1.7307 | 0.3934 | 0.3043 | 0.1693 | 0.8665 | 3.6560 | **1.6245** | 0.2071 | 1.6518$^{*}$ | 65.8549$^{*}$ |
| | GARCH-RBP | 0.7483$^{**}$ | 1.6682 | 0.4133$^{***}$ | 0.3484$^{***}$ | 0.1684$^{**}$ | 0.8372 | 3.5920$^{*}$ | 1.8148$^{**}$ | 0.1903$^{***}$ | **1.6137** | 66.0930 |
| | GARCH-VOL | 1.2095$^{***}$ | 3.3809$^{***}$ | 0.7807$^{***}$ | 1.5950$^{**}$ | 0.2419$^{***}$ | 1.6859$^{***}$ | 6.1372$^{*}$ | 3.7124$^{***}$ | 0.4425$^{***}$ | 1.7046$^{***}$ | 35.9107$^{**}$ |
| | COMBINED | 0.8134 | 1.7910 | 0.4480 | 0.3778 | 0.1800 | 0.8986 | 4.0973 | 1.8946 | 0.2120 | 1.6271 | 64.4227 |
| | *Benefit* (%) | 15.24 | 41.25 | 11.76 | 32.18 | 10.64 | 41.28 | 43.55 | 12.58 | 24.43 | 2.25 | 50.25 |

Table 5. Out-of-sample forecast evaluation (*cont.*)

| **Stock** | **Model** | MAE | MSE | HMAE | HMSE | AMAPE | Theil-U | MME(U) | MME(O) | LL | GMLE | MZ-$R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Forecast accuracy measures** | | | | | | | | | | |
| | GARCH | 1.9435 | 17.3639$^{***}$ | 0.5624 | 0.9484 | 0.2227 | 0.8972 | 36.6720 | 6.0531 | 0.4201 | 2.3551 | 7.9065 |
| | GARCH-RV | 1.8887$^{***}$ | 14.6883$^{*}$ | 0.5322$^{***}$ | 0.7568$^{***}$ | 0.2122$^{***}$ | 0.7590 | 28.0535 | 8.0831 | 0.3630$^{***}$ | 2.2988$^{***}$ | 21.9645$^{*}$ |
| | GARCH-RR | 1.8626$^{***}$ | **13.7368** | 0.5460$^{**}$ | 0.7525$^{***}$ | 0.2088$^{**}$ | **0.7098** | 30.6672$^{**}$ | 6.8102$^{***}$ | 0.3441$^{***}$ | **2.2724** | 24.6457 |
| **MCD** | GARCH-RPV | **1.7801** | 13.8080 | **0.4685** | **0.5537** | **0.2018** | 0.7135 | **27.9033** | **4.9403** | **0.3315** | 2.2928$^{***}$ | **26.2765** |
| | GARCH-RBP | 1.8483$^{**}$ | 14.0844 | 0.5233$^{***}$ | 0.7236$^{***}$ | 0.2086$^{**}$ | 0.7278 | 28.1466$^{**}$ | 6.7899$^{***}$ | 0.3525$^{***}$ | 2.2915$^{***}$ | 23.8496$^{*}$ |
| | GARCH-VOL | 1.9901$^{***}$ | 17.8297$^{***}$ | 0.5564$^{*}$ | 0.9086$^{***}$ | 0.2294$^{**}$ | 0.9213 | 33.4877 | 6.1521 | 0.4452 | 2.3895$^{**}$ | 7.0180$^{*}$ |
| | COMBINED | 1.8361 | 13.8876 | 0.5504 | 0.7843 | 0.2077 | 0.7176 | 35.2170 | 6.0494 | 0.3431 | 2.2726 | 23.3070 |
| | *Benefit* (%) | 8.41 | 20.89 | 16.70 | 41.62 | 9.38 | 20.89 | 23.91 | 18.38 | 21.09 | 3.51 | 232.34 |
| | GARCH | 1.9524 | 8.6105$^{***}$ | 0.7942 | 1.2819 | 0.2515 | 1.5277 | 19.6059 | 8.4880 | 0.4235 | 2.2235 | 33.5199 |
| | GARCH-RV | 1.5284$^{***}$ | 5.3117$^{***}$ | 0.6150$^{***}$ | 0.7773$^{***}$ | 0.2106$^{***}$ | 0.9427$^{***}$ | 9.6405 | 6.0302$^{***}$ | 0.3000$^{***}$ | 2.1745$^{***}$ | 57.2189$^{**}$ |
| | GARCH-RR | **1.4355** | **4.7795** | **0.5647** | **0.6499** | **0.2007** | **0.8481** | **9.5153** | **5.1823** | **0.2701** | **2.1650** | **59.9528** |
| **MSFT** | GARCH-RPV | 1.5153$^{***}$ | 5.1317$^{***}$ | 0.5886$^{***}$ | 0.6887$^{**}$ | 0.2054$^{***}$ | 0.9107$^{***}$ | 10.3163 | 5.4635 | 0.2828$^{***}$ | 2.1687$^{**}$ | 58.1559$^{**}$ |
| | GARCH-RBP | 1.5164$^{***}$ | 5.3134$^{**}$ | 0.6054$^{**}$ | 0.7557$^{***}$ | 0.2091$^{***}$ | 0.9429$^{**}$ | 10.0238$^{*}$ | 5.8701$^{***}$ | 0.2947$^{***}$ | 2.1731$^{***}$ | 56.5251$^{*}$ |
| | GARCH-VOL | 3.4251$^{***}$ | 18.6881$^{***}$ | 1.5101$^{***}$ | 5.1865$^{***}$ | 0.3489$^{***}$ | 3.3161$^{***}$ | 22.8637$^{*}$ | 21.1438$^{***}$ | 0.8773$^{***}$ | 2.3555$^{***}$ | 16.8304$^{**}$ |
| | COMBINED | 1.3098 | ***4.6077*** | 0.4259 | 0.3648 | 0.1766 | 0.8162 | 8.2106 | 4.4980 | 0.2176 | 2.1642 | 59.6248 |
| | *Benefit* (%) | 26.48 | 44.49 | 28.90 | 49.30 | 20.20 | 44.49 | 51.47 | 38.95 | 36.22 | 2.63 | 78.86 |
| | GARCH | 0.6066 | 0.9080$^{***}$ | 0.5756 | 0.6819 | 0.2031 | 1.2261 | 2.2570 | 1.2715 | 0.2836 | 1.1973 | 54.4355 |
| | GARCH-RV | **0.4621** | 0.5863 | **0.3645** | **0.2629** | **0.1572** | 0.7918 | 1.2696 | 0.9195 | **0.1653** | **1.1577** | 68.7963 |
| | GARCH-RR | 0.5120$^{**}$ | 0.7127$^{**}$ | 0.4267$^{*}$ | 0.3677$^{***}$ | 0.1725$^{**}$ | 0.9627$^{**}$ | 1.6092$^{*}$ | 1.0204$^{*}$ | 0.2010 | 1.1719$^{***}$ | 62.3984$^{**}$ |
| **PG** | GARCH-RPV | 0.4681$^{*}$ | 0.5720$^{***}$ | 0.3716$^{*}$ | 0.2734$^{*}$ | 0.1637$^{***}$ | 0.7724$^{***}$ | **1.1931** | **0.9138** | 0.1855$^{**}$ | 1.1739 | 69.5375$^{**}$ |
| | GARCH-RBP | 0.4720$^{*}$ | **0.5600** | 0.3874$^{***}$ | 0.2965$^{***}$ | 0.1614$^{***}$ | **0.7561** | 1.2313$^{**}$ | 0.9487$^{*}$ | 0.1722$^{***}$ | 1.1581 | **70.3142** |
| | GARCH-VOL | 4.3179$^{***}$ | 22.7805$^{***}$ | 4.5441$^{***}$ | 36.3093$^{***}$ | 0.5733$^{***}$ | 30.6737$^{***}$ | 5.3454$^{***}$ | 24.8107$^{***}$ | 2.6818$^{***}$ | 1.8456$^{***}$ | 18.1998$^{*}$ |
| | COMBINED | 0.5323 | 0.6801 | 0.4435 | 0.3924 | 0.1907 | 0.9145 | 1.2516$^{**}$ | 1.1585 | 0.2861 | 1.2643 | 66.6211$^{**}$ |
| | *Benefit* (%) | 23.82 | 38.33 | 36.67 | 61.45 | 22.59 | 38.33 | 47.14 | 28.13 | 41.71 | 3.31 | 29.17 |
| | GARCH | 1.2801 | 6.0674$^{***}$ | 0.6151 | 0.7723 | 0.2243 | 1.3870 | 16.6024 | 3.5017 | 0.3426 | 1.7990 | 39.8960 |
| | GARCH-RV | 1.0887$^{**}$ | 4.1895 | 0.4930$^{*}$ | 0.5157 | 0.1907$^{***}$ | 0.9575 | 9.4215$^{*}$ | 3.3163 | 0.2470$^{*}$ | 1.7519$^{**}$ | 57.9931 |
| | GARCH-RR | 1.1819$^{**}$ | 4.4596$^{**}$ | 0.5554$^{***}$ | 0.6405$^{*}$ | 0.2037$^{*}$ | 1.0189$^{**}$ | 10.7409$^{*}$ | 3.6044$^{*}$ | 0.2757$^{***}$ | 1.7586$^{***}$ | 56.1337 |
| **WMT** | GARCH-RPV | **1.0543** | 4.1195 | **0.4562** | **0.4326** | **0.1835** | 0.9417 | **9.1475** | **2.9934** | **0.2295** | **1.7479** | 58.5898 |
| | GARCH-RBP | 1.0659 | **4.0894** | 0.4933$^{**}$ | 0.5289 | 0.1893$^{**}$ | **0.9347** | 9.6257$^{***}$ | 3.0971 | 0.2460$^{*}$ | 1.7508 | **58.9156** |
| | GARCH-VOL | 1.2223$^{**}$ | 6.2784$^{**}$ | 0.5272$^{*}$ | 0.5731$^{**}$ | 0.2122$^{***}$ | 1.4352$^{*}$ | 14.6043$^{*}$ | 3.0023 | 0.3137 | 1.8040$^{***}$ | 38.7879$^{**}$ |
| | COMBINED | 1.1002 | 4.2943 | 0.4746 | 0.4642 | 0.1927 | 0.9793 | 9.0753 | 3.3252 | 0.2832 | 1.8698 | 56.9535 |
| | *Benefit* (%) | 17.64 | 32.60 | 25.83 | 43.99 | 18.19 | 32.61 | 44.90 | 14.52 | 33.01 | 2.84 | 47.67 |

Table 6. Hansen's Superior Predictive Ability (SPA) test

| | Benchmark model: GARCH-RPV | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Alternative models | | | | SPA | |
| **Stock** | **Loss** | **Best Performing** | ***p*-val** | **Most Significant** | ***p*-val** | ***p*-val** | **Superior** |
| **ATT** | MAE | COMBINED | 0.0060 | COMBINED | 0.0060 | 0.0060 | COMBINED |
| | MSE | COMB-EQW | 0.8990 | COMBINED | 0.7780 | 0.9970 | GARCH-RPV |
| **AXP** | MAE | COMBINED | 0.0200 | COMBINED | 0.0200 | 0.0200 | COMBINED |
| | MSE | COMBINED | 0.1520 | COMBINED | 0.1520 | 0.2350 | GARCH-RPV |
| **BA** | MAE | GARCH-RV | 0.2340 | GARCH-RV | 0.2340 | 0.3780 | GARCH-RPV |
| | MSE | COMB-EQW | 0.1710 | COMB-EQW | 0.1710 | 0.3350 | GARCH-RPV |
| **CAT** | MAE | COMBINED | 0.0000 | COMBINED | 0.0000 | 0.0000 | COMBINED |
| | MSE | COMBINED | 0.0010 | COMBINED | 0.0010 | 0.0020 | COMBINED |
| **DELL** | MAE | COMBINED | 0.0000 | COMBINED | 0.0000 | 0.0000 | COMBINED |
| | MSE | COMBINED | 0.0000 | COMBINED | 0.0000 | 0.0000 | COMBINED |
| **GE** | MAE | COMBINED | 0.8890 | COMBINED | 0.8890 | 0.8900 | GARCH-RPV |
| | MSE | COMBINED | 0.0880 | COMBINED | 0.0880 | 0.1590 | GARCH-RPV |
| **GM** | MAE | COMBINED | 0.0000 | COMBINED | 0.0000 | 0.0000 | COMBINED |
| | MSE | COMBINED | 0.0000 | COMBINED | 0.0000 | 0.0000 | COMBINED |
| **IBM** | MAE | COMBINED | 0.0230 | COMBINED | 0.0230 | 0.0230 | COMBINED |
| | MSE | COMBINED | 0.1030 | COMBINED | 0.1030 | 0.1960 | GARCH-RPV |
| **JPM** | MAE | COMB-EQW | 0.9510 | GARCH-RBP | 0.9020 | 0.9240 | GARCH-RPV |
| | MSE | COMBINED | 0.1840 | COMBINED | 0.1840 | 0.3240 | GARCH-RPV |
| **KO** | MAE | GARCH-RV | 0.0750 | GARCH-RV | 0.0750 | 0.1170 | GARCH-RPV |
| | MSE | COMB-EQW | 0.1810 | COMB-EQW | 0.1810 | 0.5020 | GARCH-RPV |
| **MCD** | MAE | COMB-EQW | 0.9400 | COMBINED | 0.9030 | 0.9960 | GARCH-RPV |
| | MSE | GARCH-RR | 0.4180 | GARCH-RR | 0.4180 | 0.9740 | GARCH-RPV |
| **MSFT** | MAE | COMBINED | 0.0000 | COMBINED | 0.0000 | 0.0000 | COMBINED |
| | MSE | COMBINED | 0.0060 | COMBINED | 0.0060 | 0.0090 | COMBINED |
| **PG** | MAE | GARCH-RV | 0.1320 | GARCH-RV | 0.1320 | 0.2780 | GARCH-RPV |
| | MSE | GARCH-RPB | 0.2610 | GARCH-RPB | 0.2610 | 0.6160 | GARCH-RPV |
| **WMT** | MAE | GARCH-RPB | 0.8070 | GARCH-RPB | 0.8070 | 0.8260 | GARCH-RPV |
| | MSE | GARCH-RPB | 0.3830 | GARCH-RPB | 0.3830 | 0.7950 | GARCH-RPV |

The table presents the consistent SPA test $p$-values for $H_0$: 'any alternative model is not better than the benchmark' based on $B$=1000 bootstrap replications and $q$=0.5 time-dependence parameter for the block bootstrap length. Best performer is the alternative model with minimum loss. Most significant is the alternative model giving the highest $t$-statistic in the pairwise comparisons with the benchmark. The $p$-values of the pairwise comparisons ($H_0$ :'best performer/most significant model is not better than benchmark') that do not control for the full set of alternatives are also reported. COMBINED is the combined model using time-varying weights as outlined in Section 3.3. COMB-EQW is the equal-weights combined model. The last column reports the model with superior predictive ability according to the SPA test.

Table 7. Forecast evaluation and market conditions: up-market vs down-market days

| | | **Forecast accuracy measures** | | | | | |
|---|---|---|---|---|---|---|---|
| | | Up market | | | Down market | | |
| **Stock** | **Model** | $HMSE^U$ | $AMAPE^U$ | $LL^U$ | $HMSE^D$ | $AMAPE^D$ | $LL^D$ |
| **ATT** | GARCH-RV | 1.3371 | 0.2362 | 0.4086 | 1.4154 | 0.2522 | 0.4701 |
| | GARCH-RR | 1.7217 | 0.2532 | 0.4601 | 1.7095 | 0.2720 | 0.5272 |
| | GARCH-RPV | **1.1766** | **0.2319** | **0.3831** | **1.1736** | **0.2462** | **0.4385** |
| | GARCH-RBP | 1.3882 | 0.2370 | 0.4125 | 1.4695 | 0.2560 | 0.4814 |
| | Benefit (%) | *49.36* | 12.88 | 26.33 | 47.02 | *17.66* | *29.03* |
| **AXP** | GARCH-RV | 1.5765 | 0.2394 | 0.4079 | 0.5877 | 0.2138 | 0.2934 |
| | GARCH-RR | 2.8559 | 0.3006 | 0.6294 | 1.2300 | 0.2590 | 0.4368 |
| | GARCH-RPV | **1.1585** | **0.2199** | **0.3498** | **0.5004** | **0.2027** | **0.2666** |
| | GARCH-RBP | 1.6098 | 0.2365 | 0.4105 | 0.5689 | 0.2121 | 0.2881 |
| | Benefit (%) | *61.55* | *18.24* | 35.71 | 43.11 | 17.70 | *37.38* |
| **BA** | GARCH-RV | 0.5706 | **0.1860** | 0.2439 | **0.3047** | **0.1700** | 0.1990 |
| | GARCH-RR | 0.8777 | 0.2159 | 0.3202 | 0.4768 | 0.1834 | 0.2334 |
| | GARCH-RPV | **0.5463** | 0.1870 | **0.2427** | 0.3048 | 0.1705 | **0.1972** |
| | GARCH-RBP | 0.5953 | 0.1899 | 0.2513 | 0.3430 | 0.1733 | 0.2039 |
| | Benefit (%) | *23.15* | 6.51 | 14.62 | -5.44 | *12.78* | *27.15* |
| **CAT** | GARCH-RV | 1.1568 | 0.2379 | 0.3742 | 0.7687 | 0.2029 | 0.2921 |
| | GARCH-RR | 1.7984 | 0.2838 | 0.5136 | 1.1500 | 0.2290 | 0.3766 |
| | GARCH-RPV | **0.9993** | **0.2312** | **0.3459** | **0.6747** | **0.1997** | **0.2766** |
| | GARCH-RBP | 1.1291 | 0.2431 | 0.3795 | 0.7867 | 0.2072 | 0.3008 |
| | Benefit (%) | *56.00* | *24.31* | *42.60* | 50.07 | 21.46 | 39.29 |
| **DELL** | GARCH-RV | 1.1349 | 0.2483 | 0.3841 | 0.7107 | 0.2111 | 0.2866 |
| | GARCH-RR | **0.7562** | **0.2099** | **0.2846** | **0.4764** | **0.1839** | **0.2245** |
| | GARCH-RPV | 1.1342 | 0.2505 | 0.3887 | 0.6689 | 0.2091 | 0.2848 |
| | GARCH-RBP | 1.2166 | 0.2510 | 0.3974 | 0.7366 | 0.2086 | 0.2880 |
| | Benefit (%) | *37.90* | *14.42* | *27.51* | 5.59 | 5.75 | 10.30 |
| **GE** | GARCH-RV | 0.7266 | 0.2165 | 0.3115 | 0.8220 | 0.2057 | 0.2982 |
| | GARCH-RR | 0.9816 | 0.2465 | 0.3854 | 1.1415 | 0.2233 | 0.3577 |
| | GARCH-RPV | **0.4909** | **0.2016** | **0.2622** | **0.6706** | **0.1944** | **0.2690** |
| | GARCH-RBP | 0.6961 | 0.2193 | 0.3148 | 0.7963 | 0.2056 | 0.2931 |
| | Benefit (%) | *39.32* | *15.43* | *27.98* | 27.02 | 12.69 | 26.17 |
| **GM** | GARCH-RV | 1.7644 | **0.2494** | 0.4439 | 0.8799 | 0.2218 | 0.3272 |
| | GARCH-RR | 2.7927 | 0.3074 | 0.6368 | 1.4871 | 0.2752 | 0.4698 |
| | GARCH-RPV | **1.7034** | 0.2509 | **0.4426** | 0.8864 | 0.2239 | 0.3300 |
| | GARCH-RBP | 1.8966 | 0.2525 | 0.4595 | **0.8710** | **0.2181** | **0.3184** |
| | Benefit (%) | *18.23* | *5.81* | *10.47* | -14.92 | -0.71 | -2.82 |
| **IBM** | GARCH-RV | 0.9333 | 0.2241 | 0.3351 | 0.7536 | 0.2162 | 0.3077 |
| | GARCH-RR | 1.3039 | 0.2615 | 0.4356 | 0.9965 | 0.2362 | 0.3665 |
| | GARCH-RPV | **0.7941** | **0.2129** | **0.3019** | **0.5523** | **0.2013** | **0.2587** |
| | GARCH-RBP | 0.9498 | 0.2239 | 0.3379 | 0.7097 | 0.2150 | 0.3013 |
| | Benefit (%) | *47.68* | *25.61* | *40.16* | 34.59 | 15.67 | 27.52 |

The up- and down-market classification is for day $t-1$ and the forecast is for day $t$. Bold denotes the best forecasting model. Benefit (%) indicates the percentage forecast error reduction that the best forecasting model brings relative to the baseline GARCH. Italics in the last row (Benefit %) denotes the regime where the largest forecast error reduction is achieved.

Table 7. Forecast evaluation and market conditions: up- vs down-market days (*cont.*)

| | | **Forecast accuracy measures** | | | | | |
|---|---|---|---|---|---|---|---|
| | | Up market | | | Down market | | |
| **Stock** | **Model** | $HMSE^U$ | $AMAPE^U$ | $LL^U$ | $HMSE^D$ | $AMAPE^D$ | $LL^D$ |
| **JPM** | GARCH-RV | 0.4908 | 0.1797 | 0.2214 | 0.5099 | 0.1953 | 0.2810 |
| | GARCH-RR | 0.6533 | 0.2116 | 0.2870 | 0.7197 | 0.2158 | 0.3304 |
| | GARCH-RPV | **0.2927** | **0.1695** | **0.1886** | **0.3673** | **0.1873** | **0.2549** |
| | GARCH-RBP | 0.5136 | 0.1785 | 0.2209 | 0.4786 | 0.1919 | 0.2745 |
| | Benefit (%) | *80.09* | *35.87* | *58.64* | 61.98 | 22.00 | 38.32 |
| **KO** | GARCH-RV | 0.3356 | **0.1685** | **0.1908** | **0.2435** | **0.1577** | **0.1704** |
| | GARCH-RR | 0.4895 | 0.1929 | 0.2400 | 0.4181 | 0.1836 | 0.2182 |
| | GARCH-RPV | **0.3326** | 0.1739 | 0.2277 | 0.2570 | 0.1626 | 0.1762 |
| | GARCH-RBP | 0.3895 | 0.1717 | 0.1984 | 0.2834 | 0.1623 | 0.1766 |
| | Benefit (%) | 28.46 | 4.53 | 16.02 | *38.40* | *18.06* | *34.33* |
| **MCD** | GARCH-RV | 0.6195 | 0.2036 | 0.3431 | 0.9406 | 0.2241 | 0.3927 |
| | GARCH-RR | 0.6494 | 0.2026 | 0.3304 | 0.8911 | 0.2177 | 0.3651 |
| | GARCH-RPV | **0.4322** | **0.1975** | **0.3146** | **0.7127** | **0.2086** | **0.3567** |
| | GARCH-RBP | 0.5575 | 0.2028 | 0.3303 | 0.9412 | 0.2168 | 0.3842 |
| | Benefit (%) | 29.09 | 5.64 | 9.38 | *48.78* | *13.65* | *31.18* |
| **MSFT** | GARCH-RV | 0.8588 | 0.2158 | 0.3137 | 0.7290 | 0.2085 | 0.2939 |
| | GARCH-RR | **0.6800** | **0.2024** | **0.2755** | **0.6370** | **0.2009** | **0.2695** |
| | GARCH-RPV | 0.7535 | 0.2099 | 0.2951 | 0.6511 | 0.2036 | 0.2774 |
| | GARCH-RBP | 0.8689 | 0.2138 | 0.3133 | 0.6838 | 0.2071 | 0.2849 |
| | Benefit (%) | *52.06* | *22.50* | *38.35* | 46.96 | 18.38 | 34.57 |
| **PG** | GARCH-RV | **0.2634** | **0.1556** | **0.1627** | **0.2664** | **0.1612** | **0.1722** |
| | GARCH-RR | 0.4085 | 0.1723 | 0.2070 | 0.3063 | 0.1742 | 0.1940 |
| | GARCH-RPV | 0.2733 | 0.1611 | 0.1808 | 0.2778 | 0.1694 | 0.1961 |
| | GARCH-RBP | 0.2965 | 0.1592 | 0.1697 | 0.3005 | 0.1661 | 0.1784 |
| | Benefit (%) | *62.59* | *24.79* | *43.60* | 59.40 | 18.44 | 38.27 |
| **WMT** | GARCH-RV | 0.5214 | 0.1916 | 0.2355 | 0.5056 | 0.1896 | 0.2564 |
| | GARCH-RR | 0.6356 | 0.2080 | 0.2738 | 0.6358 | 0.1987 | 0.2747 |
| | GARCH-RPV | **0.4364** | **0.1831** | **0.2193** | **0.4281** | **0.1839** | **0.2391** |
| | GARCH-RBP | 0.5556 | 0.1882 | 0.2371 | 0.5024 | 0.1903 | 0.2540 |
| | Benefit (%) | *53.87* | *21.67* | *37.14* | 31.15 | 15.13 | 29.71 |

Table 8. Forecast evaluation and market conditions: high volume vs low volume days

| | | **Forecast accuracy measures** | | | | | |
|---|---|---|---|---|---|---|---|
| | | High-volume regime | | | Low-volume regime | | |
| **Stock** | **Model** | $HMSE^H$ | $AMAPE^H$ | $LL^H$ | $HMSE^L$ | $AMAPE^L$ | $LL^L$ |
| **ATT** | GARCH-RV | 1.3273 | 0.2577 | 0.4662 | 1.4227 | 0.2356 | 0.4256 |
| | GARCH-RR | 1.8214 | 0.2734 | 0.5240 | 1.6410 | 0.2565 | 0.4781 |
| | GARCH-RPV | **1.1375** | **0.2512** | **0.4356** | **1.2052** | **0.2313** | **0.3980** |
| | GARCH-RBP | 1.3726 | 0.2602 | 0.4753 | 1.4812 | 0.2381 | 0.4328 |
| | Benefit (%) | 23.62 | 10.67 | 17.36 | *57.59* | *19.28* | *34.69* |
| **AXP** | GARCH-RV | 0.8144 | 0.2344 | 0.3456 | 1.3061 | 0.2208 | 0.3564 |
| | GARCH-RR | 1.7002 | 0.2855 | 0.5268 | 2.3333 | 0.2757 | 0.5404 |
| | GARCH-RPV | **0.6148** | **0.2156** | **0.2958** | **1.0078** | **0.2082** | **0.3194** |
| | GARCH-RBP | 0.8121 | 0.2313 | 0.3455 | 1.3210 | 0.2190 | 0.3538 |
| | Benefit (%) | 53.96 | *19.03* | *38.87* | *58.95* | 17.22 | 34.57 |
| **BA** | GARCH-RV | 0.4284 | **0.1814** | 0.2190 | 0.4649 | **0.1770** | 0.2272 |
| | GARCH-RR | 0.6494 | 0.2070 | 0.2821 | 0.7289 | 0.1966 | 0.2796 |
| | GARCH-RPV | **0.4148** | 0.1824 | **0.2174** | **0.4522** | 0.1775 | **0.2258** |
| | GARCH-RBP | 0.4565 | 0.1866 | 0.2289 | 0.4982 | 0.1793 | 0.2306 |
| | Benefit (%) | -19.46 | 7.02 | 14.65 | *30.11* | *10.78* | *23.68* |
| **CAT** | GARCH-RV | 0.8655 | 0.2254 | 0.3234 | 1.0636 | 0.2190 | 0.3462 |
| | GARCH-RR | 1.3967 | 0.2649 | 0.4394 | 1.5799 | 0.2539 | 0.4593 |
| | GARCH-RPV | **0.7333** | **0.2237** | **0.3090** | **0.9375** | **0.2113** | **0.3176** |
| | GARCH-RBP | 0.8283 | 0.2309 | 0.3296 | 1.0793 | 0.2231 | 0.3534 |
| | Benefit (%) | 49.81 | 21.01 | 36.98 | *56.19* | *24.80* | *44.17* |
| **DELL** | GARCH-RV | 0.8471 | 0.2303 | 0.3322 | 1.0195 | 0.2321 | 0.3456 |
| | GARCH-RR | **0.5771** | **0.1970** | **0.2555** | **0.6703** | **0.1985** | **0.2582** |
| | GARCH-RPV | 0.8208 | 0.2312 | 0.3381 | 1.0051 | 0.2317 | 0.3435 |
| | GARCH-RBP | 0.9178 | 0.2299 | 0.3443 | 1.0626 | 0.2329 | 0.3497 |
| | Benefit (%) | 8.38 | 3.76 | 9.38 | *39.68* | *16.31* | *29.46* |
| **GE** | GARCH-RV | 0.5550 | 0.2078 | 0.2939 | 0.9635 | 0.2122 | 0.3119 |
| | GARCH-RR | 0.8313 | 0.2299 | 0.3558 | 1.2681 | 0.2362 | 0.3814 |
| | GARCH-RPV | **0.4492** | **0.1963** | **0.2679** | **0.7086** | **0.1983** | **0.2644** |
| | GARCH-RBP | 0.5633 | 0.2128 | 0.3034 | 0.9064 | 0.2103 | 0.3016 |
| | Benefit (%) | 27.52 | *15.04* | *27.27* | *34.12* | 13.18 | 26.89 |
| **GM** | GARCH-RV | 1.1687 | 0.2261 | **0.3465** | 1.3335 | **0.2394** | **0.4014** |
| | GARCH-RR | 1.7939 | 0.2746 | 0.4821 | 2.2502 | 0.3001 | 0.5881 |
| | GARCH-RPV | **1.1490** | 0.2270 | 0.3469 | **1.3085** | 0.2419 | 0.4028 |
| | GARCH-RBP | 1.2192 | **0.2241** | 0.3469 | 1.3863 | 0.2394 | 0.4040 |
| | Benefit (%) | -2.17 | -0.45 | -2.63 | *12.13* | *3.68* | *6.61* |
| **IBM** | GARCH-RV | 0.9134 | 0.2215 | 0.3289 | 0.7810 | 0.2187 | 0.3145 |
| | GARCH-RR | 1.2293 | 0.2455 | 0.4051 | 1.0733 | 0.2498 | 0.3942 |
| | GARCH-RPV | **0.6869** | **0.2062** | **0.2793** | **0.6491** | **0.2071** | **0.2787** |
| | GARCH-RBP | 0.8400 | 0.2177 | 0.3165 | 0.8090 | 0.2202 | 0.3202 |
| | Benefit (%) | 32.65 | 15.82 | 26.73 | *48.72* | *24.18* | *39.46* |

See note in Table 7.

Table 8. Forecast evaluation and market conditions: high vs low volume (*cont.*)

| | | **Forecast accuracy measures** | | | | | |
|---|---|---|---|---|---|---|---|
| | | High volume | | | Low volume | | |
| **Stock** | **Model** | $\text{HMSE}^H$ | $\text{AMAPE}^H$ | $\text{LL}^H$ | $\text{HMSE}^L$ | $\text{AMAPE}^L$ | $\text{LL}^L$ |
| **JPM** | GARCH-RV | 0.4093 | 0.1885 | 0.2537 | 0.5723 | 0.1877 | 0.2518 |
| | GARCH-RV | 0.6559 | 0.2158 | 0.3086 | 0.7141 | 0.2126 | 0.3111 |
| | GARCH-RPV | **0.3390** | 0.1862 | **0.2439** | 0.3264 | 0.1734 | 0.2076 |
| | GARCH-RBP | 0.4020 | **0.1854** | 0.2482 | 0.5688 | 0.1861 | 0.2497 |
| | Benefit (%) | 62.82 | 22.95 | 38.66 | *77.37* | *33.50* | *55.18* |
| **KO** | GARCH-RV | 0.2859 | **0.1655** | **0.1901** | **0.3066** | **0.1626** | **0.1759** |
| | GARCH-RR | 0.4086 | 0.1861 | 0.2254 | 0.5009 | 0.1910 | 0.2351 |
| | GARCH-RPV | **0.2835** | 0.1676 | 0.1958 | 0.3153 | 0.1702 | 0.2148 |
| | GARCH-RBP | 0.3291 | 0.1699 | 0.1943 | 0.3587 | 0.1658 | 0.1851 |
| | Benefit (%) | *34.66* | *12.32* | *28.32* | 28.39 | 8.37 | 19.18 |
| | GARCH-RV | 1.0436 | 0.2341 | 0.4230 | 0.5718 | 0.1987 | 0.3267 |
| **MCD** | GARCH-RR | 0.8801 | 0.2217 | 0.3873 | 0.6744 | 0.2010 | 0.3182 |
| | GARCH-RPV | **0.6893** | **0.2128** | **0.3591** | **0.4679** | **0.1959** | **0.3167** |
| | GARCH-RBP | 0.9416 | 0.2265 | 0.4054 | 0.5840 | 0.1977 | 0.3204 |
| | Benefit (%) | 38.57 | 6.67 | *22.15* | *44.75* | *11.50* | 20.53 |
| | GARCH-RV | 0.7841 | 0.2138 | 0.3101 | 0.7844 | 0.2095 | 0.2958 |
| **MSFT** | GARCH-RR | **0.6805** | **0.2070** | **0.2881** | **0.6350** | **0.1969** | **0.2589** |
| | GARCH-RPV | 0.7343 | 0.2118 | 0.3009 | 0.6624 | 0.2015 | 0.2717 |
| | GARCH-RBP | 0.7753 | 0.2134 | 0.3076 | 0.7524 | 0.2069 | 0.2883 |
| | Benefit (%) | 34.33 | 14.38 | 27.52 | *57.73* | *24.59* | *42.54* |
| | GARCH-RV | **0.2609** | **0.1634** | **0.1702** | **0.2686** | **0.1533** | **0.1636** |
| **PG** | GARCH-RR | 0.3188 | 0.1716 | 0.1921 | 0.4138 | 0.1745 | 0.2111 |
| | GARCH-RPV | 0.2747 | 0.1656 | 0.1784 | 0.2763 | 0.1636 | 0.1942 |
| | GARCH-RBP | 0.2881 | 0.1658 | 0.1737 | 0.3074 | 0.1588 | 0.1729 |
| | Benefit (%) | 56.75 | 19.30 | 38.79 | *64.52* | *24.90* | *43.78* |
| | GARCH-RV | 0.4574 | 0.1910 | 0.2413 | 0.5610 | 0.1909 | 0.2514 |
| **WMT** | GARCH-RR | 0.5413 | 0.1976 | 0.2634 | 0.7157 | 0.2084 | 0.2842 |
| | GARCH-RPV | **0.3556** | **0.1812** | **0.2244** | **0.4966** | **0.1860** | **0.2347** |
| | GARCH-RBP | 0.4328 | 0.1838 | 0.2267 | 0.6083 | 0.1944 | 0.2625 |
| | Benefit (%) | *53.42* | *20.73* | *36.21* | 37.31 | 16.34 | 30.85 |

Table 9. Average daily statistics over different market conditions

| | **Regime classification of day $t-1$** | | | |
|---|---|---|---|---|
| | Up market | Down market | High volume | Low volume |
| $r_{t-1}$ | 0.6480 | -0.6511 | 0.0608 | -0.0531 |
| $r_t$ | -0.0131 | 0.0112 | 0.0257 | -0.0276 |
| $\tilde{\sigma}^2_{t-1}$ | 3.3091 | 3.9892 | 4.3227 | 3.1242 |
| $\tilde{\sigma}^2_t$ | 3.2125 | 4.0836 | 4.0673 | 3.3180 |
| $VOL_{t-1}$ | 11,162,392 | 11,531,944 | 13,534,828 | 9,620,517 |
| $VOL_t$ | 11,091,917 | 11,594,947 | 12,432,407 | 10,499,129 |
| $\rho(\tilde{\sigma}^2_{t-1}, \tilde{\sigma}^2_t)$ | 0.6687 | 0.7034 | 0.7118 | 0.6633 |

$r_t$ is the return on day $t$, $\tilde{\sigma}^2_t$ is the sum of 5-min squared returns, $VOL_t$ is volume. The last row reports the correlation between the population variance (proxied by the sum of 5-min squared returns) on days $t$ and $t-1$. The regime classification is based on (11) and (12).