The information content of implied volatility indexes for forecasting volatility and market risk

Pierre Giot*

December 17, 2002

*The author is from Department of Business Administration & CEREFIM at University of Namur, Rempart de la Vierge, 8, 5000 Namur, Belgium, Phone: +32 (0) 81 724887, Email: pierre.giot@fundp.ac.be, and Center for Operations Research and Econometrics (CORE) at Université catholique de Louvain, Belgium. We are grateful to Luc Bauwens, Peter Christoffersen, Sébastien Laurent and seminar participants at CORE who offered helpful comments.

The information content of implied volatility indexes for forecasting volatility and market risk

ABSTRACT

In this paper, we assess the efficiency, information content and unbiasedness of volatility forecasts based on the VIX/VXN implied volatility indexes, RiskMetrics and GARCH-type models at the 5-, 10- and 22-day time horizon. Our empirical application focuses on the S&P100 and NASDAQ100 indexes. We also deal with the information content of the competing volatility forecasts in a market risk (VaR type) evaluation framework. The performance of the models is evaluated using LR, independence, conditional coverage and density forecast tests. Our results show that volatility forecasts based on the VIX/VXN indexes have the highest information content, both in the volatility forecasting and market risk assessment frameworks. Because they are easy-to-use and compare very favorably with much more complex econometric models that use historical returns, we argue that options and futures exchanges should compute implied volatility indexes and make these available to investors.

Forecasting volatility has been and still is one of the major success story in the quantitative finance and financial econometrics literature. Indeed, volatility forecasting models have enjoyed a tremendous success since the early 1980's.[1] In financial econometrics, the seminal paper by Engle (1982) has spurred considerable research into ARCH-type models, i.e. the attempt to forecast volatility based on the information given by (past) squared returns. More simple techniques rely on the use of 'rolling window estimation' for the variance of the asset returns.[2] On the other hand, there is a growing trend in the applied finance literature to advocate the use of implied volatility as the best estimate of future volatility. In the framework of an option pricing model such as the Black and Scholes (1973) model, the expected volatility of the asset over the life of the option is the volatility embedded in the price of the option. If call or put option prices are available, then the Black and Scholes (1973) pricing formula can be 'inverted' such that the expected volatility over the life of the option is computed from the observed market prices of the call or put options. Indeed, when all the other option parameters are known, there is a one-to-one relationship between the option prices and underlying (expected) asset volatility. This yields the so-called implied volatility. Details are provided in Hull (2000). Because of the growing importance of modelling and predicting asset volatility, the relevance of implied volatility vs volatility forecasts based on historical returns in order to deliver unbiased and efficient forecasts of future realized volatility is an important topic in modern finance.

While early papers (see a review in Figlewski, 1997) had to rely on somewhat crude datasets, more recent studies use improved databases of actively traded options to evaluate the information content of implied volatility vs volatility computed from historical returns. However, the empirical evidence is rather mixed as to which volatility forecast performs best, although a broad survey of recent papers by Poon and Granger (2003) indicates that, broadly speaking, forecasts based on implied volatility beat forecasts based on historical returns. For example, Day and Lewis (1992) compare the information content of implied volatility of call options on the S&P100 index to GARCH type conditional volatility. Their evidence is rather mixed. Xu and Taylor (1995) focus on the informational efficiency of the PHLX currency options market. According to Jorion (1995) who deals with FOREX data, implied volatility is an efficient but biased forecast of future volatility. Canina and Figlewski (1993) (see also Figlewski, 1997) show that there is almost no correlation between implied volatility and future realized volatility. Christensen and Prabhala (1998) argue that the use of overlapping data and the inclusion of the October 1987 market crash in the Canina and Figlewski (1993) paper is one of the main explanation as to why implied volatility was found inefficient and biased and compared so poorly with volatility forecasts based on historical returns. They show that implied volatility indeed outperforms past volatility in forecasting future volatility and features a high information content. For the S&P100 index and VIX implied volatility index, Blair, Poon, and Taylor (2001) show that historical returns do not provide much incremental information compared to the information given by the VIX index of implied volatility. For three class of assets (stock index, exchange rate and oil), Martens and Zein (2002) show that implied volatility measures do provide superior volatility forecasts compared to daily GARCH-type models. However the switch to high-frequency intraday returns and realized volatility modelled using long memory models alters the outcome of the tests as "long memory volatility forecasts can compete with implied volatility". For foreign exchange volatility and using intraday returns, Neely (2002) argues that implied volatility is a biased estimator of future realized volatility and that volatility forecasts from econometric models should be taken into account. Ederington and Guan (2002) examine the relevance of implied volatility forecasts using S&P500 futures options data and conclude that "implied volatility has strong predictive power and generally subsumes the information in historical volatility". Giot (2003) compares the incremental information content of lagged implied volatility to GARCH models of conditional volatility for a collection of agricultural commodities traded on the New York Board of Trade and shows that past squared returns only marginally improve the information content provided by the lagged implied volatility.

Our study draws on the previous work highlighted above as we assess the efficiency, information content and absence of bias of competing volatility forecasts with respect to ex-post observed realized volatility computed from daily returns. We extend the analysis of the latter papers by focusing specifically on the volatility forecasts based on the VIX and VXN implied volatility indexes and by providing an application to the quantification of VaR type market risk using these VIX and VXN indexes. To assess the VaR performances, we use a range of recent statistical tests which span LR tests, independence and conditional coverage tests and tests set in the density forecast framework. Therefore, we appraise the relevance of the volatility forecasts both from a volatility forecasting point of view and in the framework of an application (VaR) that takes as inputs the volatility forecasts.

The two implied volatility indexes are computed and made available by the Chicago Board of Options Exchange (CBOE) and have enjoyed a widespread acceptance in the community of both academics and practitioners.[3] By construction, the VIX (VXN) index is a weighted average of the implied volatilities computed from a total of eight call and put near-the-money, nearby and second nearby American option contracts on the underlying S&P100 (NASDAQ100) index. The weighting method ensures that VIX/VXN give the implied volatility of a hypothetical at-the-money option with a constant maturity of 22 trading days to expiry. In an efficient market where option prices reflect all available information, the level of VIX (VXN) "is the market's best assessment of the expected volatility of the underlying stock index over the remaining life of the option", a life of 22 trading days and the underlying S&P100 (NASDAQ100) index in this case (quote is from Whaley, 2000, p. 1). Details regarding the construction of the VIX index are available in Whaley (1993), Fleming and Whaley (1995) or Whaley (2000). Note that, by construction, these implied volatility indexes take into account early exercise and dividend payments features, and they do not use as inputs market prices from options that are not actively traded (thus avoiding the troublesome problem of stale quotes for deep out-of-the money or in-the-money options). Thus, the VIX and VXN indexes deliver easy-to-use information regarding future volatility and should be less prone to computation errors than previous measures of implied volatility.

More specifically, we consider volatility forecasts based on the VIX/VXN implied volatility indexes, RiskMetrics and GJR-GARCH models and assess their relevance, i.e. efficiency, information content and absence of bias. For each volatility forecast, we consider three forward-looking horizons which are successively 5, 10 and 22 days. First, we compute descriptive statistics (mean absolute error, root mean square error and proportion of variance) for the three volatility forecasts vs the ex-post observed realized volatility. Secondly, we use Ordinary Least Squares analysis to regress the realized volatility against each of the three volatility forecasts, both in a basic regression setting and in the so-called encompassing regression framework. Thirdly and for the implied volatility forecast only, we perform the encompassing regression analysis using Instrumental Variable least squares to assess the validity of our regression results taking into account the potential Error In Variable (EIV) problem. Next we estimate GARCH models with and without lagged implied volatility and compare the estimated log-likelihoods to assess the information content of the implied volatility. Finally, we focus on the information content of the VIX/VXN based volatility forecasts in a market risk evaluation framework. We estimate Value-at-Risk models that take as volatility inputs the three competing volatility forecasts. Their performance is evaluated using first LR tests based on the proportion of VaR violations, then independence and conditional coverage tests and finally tests set in the density forecast framework (based on the probability integral transformations of the residuals).

For the S&P100 and NASDAQ100 indexes and the 1-, 5- and 10-day horizons, our results show that (a) no volatility estimator (with the exception of the VXN based at a 10-day horizon) is both unbiased and efficient; (b) lagged realized volatility does not really improve the fit of the models in the encompassing regressions; (c) volatility forecasts based on the VIX/VXN implied volatility indexes are the closest to being unbiased, have the highest information content and their statistical fit increases with the time horizon; (d) GARCH-type forecasts have little incremental information over that contained in the implied volatility indexes. When VaR models are estimated, our tests show that volatility forecasts based on the VIX/VXN indexes are meaningful inputs in VaR models as the number of VaR violations is correctly modelled in most cases, the null hypotheses of independence (for the VaR violations) and conditional coverage are usually not rejected and the probability integral transformations are not correlated. As such, our results, along with recent results given by Blair, Poon, and Taylor (2001), Claessen and Mittnik (2002) or Giot (2003), show that implied volatility indexes do provide accurate and meaningful information as to future volatility forecasts for the underlying indexes. Because they are easy-to-use and compare very favorably with much more complex econometric models that use historical returns, we argue that options and futures exchanges should develop implied volatility indexes and make these available to investors.

The rest of the paper is structured as follows. After this introduction, we detail the volatility forecasts in Section I. We characterize the ex-post observed realized volatility in Section 9. Section III shows how we assess the information content of the competing volatility forecasts. In Section IV we present an application to the computation of Value-at-Risk models. Finally, Section V concludes.

I. Volatility forecasts

This section details the three measures of volatility forecasts that are used throughout the paper. For a time horizon of h days, we first present the forecasts based on the implied volatility indexes and then describe the measures of volatility forecasts based on historical returns, i.e. the RiskMetrics and GJR-GARCH volatility forecasts.

A. Implied volatility

Our measure of implied volatility is given by the level on day t of the VIX and VXN implied volatility indexes, for the S&P100 and NASDAQ100 respectively. By definition and as explained in the introduction of the paper, the forward-looking time horizon is equal to 22 trading days and the implied volatility indexes are expressed in annualized terms. Unfortunately, there are no implied volatility term structures given by the CBOE. Therefore, we use the 'square root of time rule' to switch from a time horizon of 22 days to the required h-day interval. Hence and for a h-day forward-looking horizon, the implied volatility forecast on day t for the S&P100 index is equal to:

$$\sigma_{imp,h,t} = \sqrt{\frac{h}{360}} VIX_t. \tag{1}$$

Thus $\sigma_{imp,h,t}$ is the expected volatility over the $[t+1, t+h]$ period. For the NASDAQ100 index, the corresponding expression is $\sigma_{imp,h,t} = \sqrt{\frac{h}{360}} VXN_t$.

B. RiskMetrics

Let us define by $r_t = ln(P_t) - ln(P_{t-1})$ the daily return on the S&P100 or NASDAQ100 index. Then, the RiskMetrics specification for volatility on day t is:

$$RM_t = (1-\lambda) r_{t-1}^2 + \lambda RM_{t-1} \tag{2}$$

with $\lambda = 0.94$ for daily returns.[4] Note that RM_t is an unconditional measure of daily variance and there are thus no volatility term structure in this model either. Hence and for a h-day forward-looking horizon, the RiskMetrics volatility forecast on day t (i.e. expected volatility over the $[t+1, t+h]$ time period according to the RiskMetrics model) for the S&P100 or NASDAQ100 index is equal to:

$$RM_{h,t} = \sqrt{h RM_t}. \tag{3}$$

C. ARCH-type forecasts

ARCH-type models have become extremely popular in financial econometrics and quantitative finance for modelling and forecasting volatility. The first ARCH model was put forward by Engle (1982) and was followed by numerous extensions such as the GARCH model of

Bollerslev (1986), the EGARCH model of Nelson (1991), the GJR-GARCH of Glosten, Jagannathan, and Runkle (1993) or the APARCH model of Ding, Granger, and Engle (1993). See Engle (1995) or Bollerslev, Chou, and Kroner (1992) for a review of ARCH models and their applications. Because ARCH type models focus on the modelling of conditional variance, their volatility forecasts exhibit a dynamic structure that allows for a decreasing volatility term structure. In this paper, we work with the GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993). Our choice of GARCH model is motivated by the fact that we wish to deal with a simple-to-estimate model (because the implied volatility and RiskMetrics based volatility forecasts detailed above are quite simple to use and it is sensible to compare like with like) that however retains some flexibility. As explained below, the GJR-GARCH structure allows for conditional asymmetry in the conditional variance and also allows for multi-step variance forecasts using a simple recursion rule. As such the GJR-GARCH model is a credible alternative to implied volatility and RiskMetrics based volatility forecasts for market practitioners who do not wish to deal with lengthy estimation procedures.

More specifically and for $r_t = ln(P_t) - ln(P_{t-1})$ as the daily return on the S&P100 or NASDAQ100 index, the GJR-GARCH model can be written as:[5]

$$r_t = \varepsilon_t \sqrt{g_t} \tag{4}$$

where ε_t is the IID standardized error term[6] and g_t is defined as:

$$g_t = \omega + \alpha_1 r_{t-1}^2 + \alpha_n r_{t-1}^2 d_{t-1} + \delta_1 g_{t-1} \tag{5}$$

where α_n measures the asymmetric effect and d_t is a dummy variable which is equal to 1 when e_t is negative. Thus, good news (e_t is positive) has an impact of α_1, while bad news (e_t is negative) has an impact of $\alpha_1 + \alpha_n$. Empirical studies (Black, 1976; French, Schwert, and Stambaugh, 1987; Pagan and Schwert, 1990; Glosten, Jagannathan, and Runkle, 1993; Nelson, 1991) on financial returns usually indicate a negative leverage effect, i.e. negative returns increase the conditional volatility. On day t, the one-day-ahead forecast for the conditional variance is immediately given by:

$$g_{t+1} = \omega + \alpha_1 r_t^2 + \alpha_n r_t^2 d_{t-1} + \delta_1 g_t \tag{6}$$

where the parameters of the model are set equal to their estimated values. For $h > 1$, the following recursion is used:

$$g_{t+k} = \omega + (\alpha_1 + 0.5\alpha_n + \delta_1) g_{t+k-1} \tag{7}$$

and

$$GJR_{h,t} = \sqrt{\sum_{k=1}^{h} g_{t+k}}. \tag{8}$$

Thus $GJR_{h,t}$ is the expected volatility over the $[t+1, t+h]$ period according to the GJR-GARCH model. Hence, $GJR_{h,t}$ exhibits a volatility term structure as the volatility forecast up to day $t+h$ is no longer (as in the RiskMetrics methodology) the product of the square root of h with the volatility forecast for day $t+1$.

II. Realized volatility

Given daily returns $r_t = ln(P_t) - ln(P_{t-1})$ for the S&P100 or NASDAQ100 index, the forward-looking realized volatility over a time horizon of h days is computed by taking the square root of the sum of the (future) squared returns over this h-day period. At time t, the forward-looking realized volatility $RV_{h,t}$ for the time period $[t+1, t+h]$ is thus computed as:

$$RV_{h,t} = \sqrt{\sum_{j=1}^{h} r_{t+j}^2}. \tag{9}$$

Note that this volatility measure is computed ex-post, i.e. at time $t+h$ when all returns have been observed.[7]

A. Overlapping data

In a time series framework where t ranges from 1 to T, Equation (9) defines realized volatility using overlapping data. Indeed, successive measures of realized volatility $\{RV_{h,t}\}$, for $t = 1 \dots T$, share some squared returns. For example and with $h = 2$, $RV_{2,t}$ and $RV_{2,t+1}$ both use r_{t+2}^2.

B. Non-overlapping data

For our empirical analysis, we also define realized volatility computed from non-overlapping data. Indeed, the measure of realized volatility computed using Equation (9) and using all $\{RV_{h,t}\}$ for $t = 1 \dots T$ yields strongly correlated volatility measures. As pointed out in Christensen and Prabhala (1998), the use of realized volatility computed from overlapping data in regression analysis yields potentially big estimation problems and incorrect results (see below). Hence we also define realized volatility measures computed from non-overlapping squared returns data. While Equation (9) is still valid, we no longer compute it for all $t = 1 \dots T$ but for a subset of those times such that the newly defined $\{RV_{h,k}\}$ use unique data. In this case, it is straightforward to see that the sampling times k are: $\{1, h, 2h, \dots\}$.

III. Assessing the information content of the volatility forecasts

In this section, we assess the information content of the three competing volatility forecasts presented in Section I for a 5-, 10- and 22-day forward-looking horizon. We deal with the S&P100 index and associated VIX implied volatility index, and the NASDAQ100 index and corresponding VXN implied volatility index. We first provide information on how the volatility forecasts are empirically computed and perform a graphical analysis. We then proceed with the statistical and econometric analysis whose four steps are: (a) descriptive statistics (mean absolute error, root mean square error and proportion of variance) for the three volatility forecasts vs the ex-post observed realized volatility; (b) OLS analysis (basic and encompassing regressions) of the realized volatility against each of the three volatility forecasts; (c) for the implied volatility forecast only, Instrumental Variable least squares analysis to assess the validity of our regression results taking into account the potential Error In Variable (EIV) problem; (d) in-sample GARCH model analysis with and without lagged implied volatility to compare the estimated log-likelihoods and assess the information content of the implied volatility.

For the S&P100 and NASDAQ100 indexes, complete return and implied volatility data is available from January 2nd, 1986 to June 28th, 2002, and from January 3rd, 1995 to June 28th, 2002, respectively. Prior to the regression analysis detailed in the following sections, we extract (and remove) an initial period from these samples for both indexes. Indeed we need a sample of about 1 to 2 years of daily data as a first estimation sample for the GJR-GARCH model.[8] Note that all models are evaluated in a purely out-of-sample framework. While this is straightforward for the RiskMetrics and implied volatility methods (Equations (1) and (2) are computed for all times t and immediately give out-of-sample volatility forecasts for the next h days), the GJR-GARCH method requires re-estimating the parameters of the model as new daily observations enter the information set. Thus, the GJR-GARCH model is first estimated on an initial period (from $t = 1$ to $t = T_0$ say) to get the first set of parameters (which allow the computation of the volatility forecasts at the h-day horizon using Equation (8)). Then the model is re-estimated on a weekly basis as the analysis is rolled forward, i.e. the successive estimation samples are increased by adding daily returns, i.e. considering successively the estimation of the model for $t = 1$ to $t = T_0 + 1$, $t = 1$ to $t = T_0 + 2, \ldots, t = 1$ to $t = T - 1$ where T is the end date of June 28th, 2002. Prior to the regression analysis, this initial analysis thus gives us daily out-of-sample forecasts for the expected volatility over the next h days according to the implied volatility, RiskMetrics and GJR-GARCH methods. Thereafter, the assessment of the volatility forecasts is performed on the August 1st, 1989 - June 28th, 2002 period (S&P100 index) and on the September 3rd, 1996 - June 28th, 2002 period (NASDAQ100 index). These are thus the samples where the realized volatility is compared to the volatility forecasts. Overlapping data and non-overlapping data realized volatility is computed for all times t (and all sampling times k) in these two time periods using Equation (9).

In Figure 1 we show the level of the S&P100 and VIX indexes over the August 1st, 1989 - June 28th, 2002 period. Over that period, the S&P100 index was characterized by a strong bull market, at least until March 2000. In the early 1990's, the VIX level was quite high but decreased significantly in the 1991-1996 period. Since the end of 1996, the VIX index has taken much larger values. Over the August 1st, 1989 - June 28th, 2002 sample, its average value is 20.32, its minimum value is equal to 9.04 and its maximum value is equal to 49.04. Figure 4 shows the level of the NASDAQ100 and VXN indexes over the September 3rd, 1996 - June 28th, 2002 period. While the NASDAQ100 index was characterized by a very strong bull market until March 2000, it subsequently lost more than 75% of its value. Note that the VXN index has been much higher since 2000 than in the 1990's. Over the September 3rd, 1996 - June 28th, 2002 sample, its average value is 45.23, its minimum value is equal to 24.44 and its maximum value is equal to 93.17. The annualized volatility of the NASDAQ100 index is thus much larger than the annualized volatility of the S&P100 index. Both patterns suggest a negative relationship between stock index returns and implied volatility, i.e. implied volatility increases sharply in bear markets when stock index returns are mostly negative.[9]

Figures 2, 3 (for the S&P100 index) and 5, 6 (for the NASDAQ100 index) plot the realized volatility and competing volatility forecasts at the 5- and 22-day horizon respectively. In each figure, the top panel shows the realized volatility and the VIX/VXN based volatility forecast, the middle panel shows the realized volatility and the RiskMetrics based volatility forecast while the bottom panel shows the realized volatility and the GJR-GARCH based volatility forecast. In all figures and panels, the volatility forecasts seem to track quite well the realized volatility with the exception of the GJR-GARCH forecasts in the 1991-1996 period (realized volatility was quite low at that time, and the GJR-GARCH volatility forecasts were much too high).

A. MAE, RMSE and proportion of variance

Given the realized volatility $RV_{h,t}$ and volatility forecasts $Y_{h,t}$ where $Y_{h,t}$ is successively the implied volatility, RiskMetrics and GJR-GARCH forecast:

- the mean absolute error (MAE) of the volatility forecast is equal to:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|RV_{h,t} - Y_{h,t}| \tag{10}$$

- the root mean square error (RMSE) of the volatility forecast is equal to:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(RV_{h,t} - Y_{h,t})^2} \tag{11}$$

- the proportion of variance explained by the volatility forecast is equal to:

$$P = 1 - \frac{\sum_{t=1}^{T}(RV_{h,t} - Y_{h,t})^2}{\sum_{t=1}^{T}(RV_{h,t} - RV)^2} \tag{12}$$

where RV is the mean of $RV_{h,t}$ over the sample.

Empirical results for the S&P100 index are given in Table I. The top panel of this table presents results for the realized volatility measure which uses overlapping data, while the bottom panel shows results for the non-overlapping case. There are no real differences between the results for the overlapping and non-overlapping data. In all cases and for all time horizons, the RMSE is the smallest for the forecast based on the VIX index while the proportion of variance (P) is the highest. Based on these two criteria, the volatility forecast based on the implied volatility index provides the best information as to the realized volatility over the next h days. Note also that the proportion of variance for the VIX based volatility forecast is an increasing function of the number of look-ahead days. Quite surprisingly, the RiskMetrics volatility forecast is the second-best candidate and it outperforms the GJR-GARCH volatility forecast for all h although it is based on a much simpler volatility structure. Results for the NASDAQ100 index are given in Table II and are quite similar to what has been discussed for the S&P100 index: volatility forecasts based on the VXN index have the smallest RMSE and largest P, with the proportion of variance P being largest for the 22-day time horizons. In this case however, GJR-GARCH volatility forecasts outperform RiskMetrics volatility forecasts on a short-term (5- to 10-day) basis. This preliminary evidence suggests that VIX/VXN based volatility forecasts do outperform volatility forecasts measures based on historical returns on a short to medium-term basis.

B. Ordinary regression analysis

On day t, the ex-post observed realized over the forward-looking h-day horizon is $RV_{h,t}$. The three volatility forecasts are $\sigma_{imp,h,t}$ (VIX/VXN based volatility forecast), $RM_{h,t}$ (RiskMetrics forecast) and $GJR_{h,t}$ (GJR-GARCH forecast). A basic regression analysis suggests estimating the three regressions:

$$RV_{h,t} = \beta_0 + \beta_1 \sigma_{imp,h,t} + e_t, \tag{13}$$

$$RV_{h,t} = \beta_0 + \beta_1 RM_{h,t} + e_t, \tag{14}$$

$$RV_{h,t} = \beta_0 + \beta_1 GJR_{h,t} + e_t \tag{15}$$

and analyzing the estimated coefficients, standard errors and R^2 of the regression. More importantly, the so-called encompassing regression analysis suggests three relevant statistical tests: (a) is the volatility forecast efficient? (b) is the volatility forecast unbiased? (c) is the volatility forecast unbiased and efficient? The encompassing regressions are:

$$RV_{h,t} = \beta_0 + \beta_1 \sigma_{imp,h,t} + \beta_2 RV_{h,t-1} + e_t, \tag{16}$$

$$RV_{h,t} = \beta_0 + \beta_1 RM_{h,t} + \beta_2 RV_{h,t-1} + e_t, \tag{17}$$

$$RV_{h,t} = \beta_0 + \beta_1 GJR_{h,t} + \beta_2 RV_{h,t-1} + e_t. \tag{18}$$

The first test (efficiency?) requires a t-statistic test on the β_2 coefficient. If $H0 : \beta_2 = 0$ cannot be rejected, then the volatility forecast is efficient with respect to the observed realized volatility $RV_{h,t}$. The second test requires a Fisher test as the volatility forecast is deemed unbiased if $H0 : \beta_0 = 0$ AND $\beta_1 = 1$ cannot be rejected. Finally, the volatility forecast is both unbiased and efficient if $H0 : \beta_0 = 0$ AND $\beta_1 = 1$ AND $\beta_2 = 0$ cannot be rejected.[10] Additionally, the usual t-statistic on the β_1 coefficient tells us if the volatility forecast provides some relevant information in addition of the lagged realized volatility. For all regression models, the error term e_t is the usual error term in regression analysis, i.e. $e_t \sim N(0, \sigma^2)$. We consider the Instrumental Variable regression analysis for the implied volatility forecast in the next sub-section.

As pointed out in Christensen and Prabhala (1998) as an answer to the previous analysis by Canina and Figlewski (1993), it is of paramount importance to estimate those models using the *non-overlapping* realized (and forecasted) volatility data. Indeed, if those regressions are estimated with the overlapping volatility data, they automatically exhibit a very large amount of autocorrelation which leads to potential incorrect estimation results. Hence in our empirical study, we estimate the basic and encompassing regressions using non-overlapping data.

Estimation results for the S&P100 index are given in Tables III (basic regression) and V (encompassing regression). Note that we focus immediately on and present results only for the regressions based on the non-overlapping data. A quick estimation of these regressions with overlapping data yields extremely strong correlated residuals (with a DW statistic around 0.2), which confirms the analysis of Christensen and Prabhala (1998). We also test the regression residuals for possible heteroscedasticity but the null hypothesis of no-heteroscedasticity is almost never rejected.[11] See also below for the estimation of the models on separate subperiods (S&P100 index).

A comparison of the R^2 of both basic and encompassing regressions for the three volatility forecasts confirm the fact that (a) volatility forecasts based on implied volatility indexes outperform the other volatility forecasts and (b) they get better and better as the time horizon h increases. For the basic regression $RV_{h,t} = \beta_0 + \beta_1 \sigma_{imp,h,t} + e_t$, β_0 and β_1 are not in most cases (individually) significantly different from 0 and 1 respectively, at the 5% level. Similarly, the encompassing regression $RV_{h,t} = \beta_0 + \beta_1 \sigma_{imp,h,t} + \beta_2 RV_{h,t-1} + e_t$ indicates that β_0, β_1 and β_2 are not (individually) significantly different from 0, 1 and 0 respectively. A Fisher test of unbiasedness in this encompassing regression, i.e. testing H0: $\beta_0 = 0$ AND $\beta_1 = 1$, is rejected for $h = 5$ and $h = 10$, but is not rejected for $h = 22$. The joint hypothesis H0: $\beta_0 = 0$ AND $\beta_1 = 1$ AND $\beta_2 = 0$ (i.e. unbiasedness and efficiency) is however rejected for all time horizons. Estimation results are quite different for the RiskMetrics volatility forecasts as β_0 and β_1 are often (individually) significantly from 0 and 1 respectively, but β_2 is (individually) not different from 0. Finally, for the GJR-GARCH volatility forecasts, (individual) results are acceptable at the 5-day horizon, but are much worse for the 10- and 22-day horizon. For all volatility forecasts (with the exception of the GJR-GARCH case where $h = 10$ and $h = 22$), it is however noteworthy that β_2 is in general not significantly different from 0. Hence lagged realized volatility does not really add valuable information in the encompassing regression. It is also obvious that β_0 and β_1 are the closest to 0 and 1 respectively when $\sigma_{imp,h,t}$ is the independent variable.

Estimation results for the NASDAQ100 index are given in Tables IV (basic regression) and VI (encompassing regression). Volatility forecasts based on the VXN index again display the largest R^2 in both basic and encompassing regressions, with GJR-GARCH being the second best when $h = 5$ and $h = 10$. Hence the GJR-GARCH volatility forecasts outperform the RiskMetrics volatility forecasts for short-term horizons in this case. Regarding VXN (encompassing regression), coefficients β_1 and β_2 are usually quite close to 1 and 0 respectively, although there are much larger discrepancies than for the S&P100 index. Coefficient β_0 is in most cases statistically smaller than zero. This would suggest that the volatility forecast based on the VXN index is always larger than the realized volatility at the give h-day time horizon. Regarding the RiskMetrics model, coefficients β_0, β_1 differ in most cases from 0 and 1 respectively. For the GJR-GARCH model, coefficients β_0, β_1 are quite close to 0 and 1 respectively, although not when $h = 22$ days. Note that the null hypothesis of unbiasedness and efficiency is always rejected except at the 10-day horizon for the VXN based volatility forecasts. The GJR-GARCH volatility forecast is almost unbiased at the 5-day horizon.

For the S&P100 index, we give in Tables VII and VIII results for the estimation of the encompassing regressions on the (distinct) August 1st, 1989 - December 31th, 1996 and January 2nd, 1997 - June 28th, 2002 time periods. Indeed, the first sub-period was characterized by a rather low volatility, while the second sub-period displayed a much larger return volatility which was caused by numerous events such as the Asian and Russian crisis (1997 and 1998), the end of the bull market (from 2000 onwards), the terrorist attacks on September 11th, 2001,... Furthermore, the volatility of the realized volatility and volatility forecasts is larger in the second sub-period than in the first as evidenced in Figures 1, 2 and 3.[12]

There are relatively few differences with the results presented for the global August 1st, 1989 - June 28th, 2002 time period. However, the R^2 of all encompassing regressions are much smaller in the second sub-period (major volatility shocks were not properly anticipated by the models) and the GJR-GARCH model clearly outperforms the RiskMetrics model. Note that when both sub-periods are tackled separately, the VIX based volatility forecasts are still the best game in town.

At this stage and lumping together the outputs for the S&P100 and NASDAQ100 indexes, our results can be summarized by (a) no volatility estimator (with the exception of VXN and $h = 10$ days) is both unbiased and efficient; (b) lagged realized volatility does not really improve the fit of the model in the encompassing regressions; (c) volatility forecasts based on the implied volatility indexes are the closest to being unbiased; (d) volatility forecasts based on the implied volatility indexes have the largest information content and their fit increases with the time horizon. Note that the criteria of unbiasedness, efficiency and information content need to be jointly (and not individually) appraised. Indeed and as pointed out in Poon and Granger (2003), an unbiased forecast is useless if the forecast errors are very large (i.e. no information content), while a biased forecast (that has a large information content) can help if the bias is taken into account.

C. Instrumental variable regression analysis

Implied volatility measures are computed from observed options prices. RiskMetrics and GARCH-type volatility forecasts are computed from historical returns. Hence the volatility forecasts based on implied volatility estimates are potentially affected by the so-called Error In Variable (EIV) problem, which does not affect RiskMetrics and GARCH-type volatility forecasts. Christensen and Prabhala (1998) list potential sources of measurement errors, such as a possible mismatch in option pricing formulae (no-dividend BS instead of American style options with dividends), a possible non-synchronicity between option prices and closing index levels or bid-ask spreads in options prices. Note however that most of these potential problems are clearly of limited nature in our study (compared to the implied volatility measures used in Christensen and Prabhala, 1998), as the VIX and VXN indexes take into account early exercise and dividend payments features, and use only actively traded options as inputs to the option pricing model.

Nevertheless, for the basic and encompassing regressions which use implied volatility forecasts as independent variables, we re-estimate the previously defined equations using a two-stage Instrumental Variable methodology as suggested in Christensen and Prabhala (1998). Thus, Equations (13) and (16) are no longer directly estimated by taking $\sigma_{imp,h,t}$ as an input but use instead $INS_{h,t}$ as an instrumental variable. The basic regression can now be written as:

$$RV_{h,t} = \beta_0 + \beta_1 INS_{h,t} + e_t \tag{19}$$

and $INS_t = \widehat{\sigma}_{imp,h,t}$ is the forecasted implied volatility as given by the preliminary (first-stage) regression:

$$\sigma_{imp,h,t} = \beta_0^{'} + \beta_1^{'} \sigma_{imp,h,t-1} + e_t^{'}. \tag{20}$$

The encompassing regression is:

$$RV_{h,t} = \beta_0 + \beta_1 INS_{h,t} + \beta_2 RV_{h,t-1} + e_t \tag{21}$$

and $INS_t = \widehat{\sigma}_{imp,h,t}$ is the outcome of:

$$\sigma_{imp,h,t} = \beta_0^{'} + \beta_1^{'} \sigma_{imp,h,t-1} + \beta_2^{'} RV_{h,t-1} + e_t^{'}. \tag{22}$$

Estimation results for the S&P100 index are given in Table IX. The top panel of the table gives results for the basic regression, while the bottom panel of the table presents results for the encompassing regression. There are no key differences with the previous results of Tables III and V: coefficients β_0, β_1 and β_2 are (individually) close to 0, 1 and 0 respectively, and the Fisher test of unbiasedness is not rejected when $h = 22$ days only. The Fisher test of unbiasedness and efficiency is always rejected. Estimation results for the NASDAQ100 index are given in Table X. As in Table VI, the VXN based volatility forecasts are both unbiased and efficient at the 10-day horizon, but this is now also the case for the 22-day horizon. Broadly speaking, there are thus no sharp differences between the results of the OLS and IV regressions. Volatility forecasts based on the implied volatility indexes are unbiased at the 22-day horizon for the S&P100 index, while they are both unbiased and efficient at the 10- and 22-day horizon for the NASDAQ100 index. Note however that the number of observations is smaller for the NASDAQ100 index than for the S&P100 index, hence a greater uncertainly in the estimation results.

D. Implied volatility in ARCH-type models

The analysis given in the previous sections assessed the information content of the competing volatility forecasts using descriptive statistics and encompassing regressions. This was also an 'out-of-sample' analysis as the regressions took as inputs h-day ahead forecasted variables with rolling estimation. The literature on implied volatility suggests another way of assessing the information content by focusing on the impact of lagged implied volatility as an additional regressor in the specification for the conditional variance in GARCH-type models. In this case, this is thus an in-sample analysis. See for example Day and Lewis (1992), Xu and Taylor (1995), Blair, Poon, and Taylor (2001) or Giot (2003). In this framework, the conditional volatility model (GARCH type) is first estimated without the lagged implied volatility, then with the lagged implied volatility as an additional variable in the conditional variance equation, and finally with the lagged implied volatility only (i.e. no GARCH effects in this third specification). Hence and for a GARCH(1,1) model, the conditional volatility is successively given by the three specifications:

$$g_t = \omega + \alpha_1 r_{t-1}^2 + \delta_1 g_{t-1}, \tag{23}$$

$$g_t = \omega + \alpha_1 r_{t-1}^2 + \delta_1 g_{t-1} + \eta \sigma_{imp,t-1}^2 \tag{24}$$

and

$$g_t = \omega + \eta \sigma_{imp,t-1}^2. \tag{25}$$

Among relevant statistical tests, one can highlight: (a) a t-statistic test on η in Equation (24) to assess if lagged implied volatility provides additional incremental relevant information w.r.t. historical lagged squared returns; (b) a joint test on α_1 and δ_1, i.e. a LR test for $H0 : \alpha_1 = 0$ AND $\delta_1 = 0$ by comparing the log-likelihoods after ML estimation of the models given by Equations (24) and (25). If this null hypothesis cannot be rejected, then the implied volatility forecast is deemed efficient.

Estimated log-likelihoods for the three specifications of the volatility model applied to daily returns for the two indexes are given in Table XI. The estimation (in-sample) period is August 1st, 1989 - June 28th, 2002 (S&P100 index) and September 3rd, 1996 - June 28th, 2002 (NASDAQ100 index). By construction, the log-likelihood of Model (24) is the largest. Note however that Model (25), i.e. where the conditional variance is specified as $g_t = \omega + \eta \sigma_{imp,t-1}^2$ ranks second, ahead of the GARCH(1,1) without the lagged implied volatility, i.e. $g_t = \omega + \alpha_1 r_{t-1}^2 + \delta_1 g_{t-1}$. More precisely, the LR test for $H0 : \alpha_1 = 0$ AND $\delta_1 = 0$ in Equation (24) indicates that this hypothesis is rejected for the S&P100 index, but not for the NASDAQ100 index. Indeed, twice the decrease in log-likelihood when switching from Equation (24) to (25) is equal to 25.53 (S&P100 index) and 5.99 (NASDAQ100 index). The critical value for the χ^2 with 2 degrees of freedom is equal to 5.99 at the five percent level. Hence, in the in-sample GARCH framework, the null hypothesis that past squared returns add no significant volatility information in addition to the lagged implied volatility is thus not rejected for the NASDAQ100 index.

IV. Volatility forecasts and the quantification of market risk

The analysis given in the previous sections has shown that volatility forecasts computed from implied volatility indexes have a high information content and outperform in most cases volatility forecasts based on historical returns. The last section of the paper deals with the information content of the VIX/VXN based volatility forecasts in a market risk evaluation framework. Indeed, volatility forecasting has found numerous applications in quantitative finance, such as portfolio management, option pricing or risk management. In these three fields, volatility forecasts are one of the main inputs to the relevant models and are thus of paramount importance in the empirical application.

In this section we assess the added value of the VIX/VXN based volatility forecasts when these forecasts are used to quantify short-term market risk. We consider the familiar Value-at-Risk framework which provides, at a given percentage level, the most likely loss for a financial institution. For example, the VaR at level α for a time horizon of h days is the nominal h-day loss that will not be exceeded in $100 \cdot \alpha$ portfolio realizations out of 100. The literature on VaR models has grown remarkably since the middle of the 1990's because of the popularity of the RiskMetrics VaR specification of JP Morgan and the risk-adjusted measures of capital adequacy enforced by the Basel committee. Recent developments are presented in Jorion (2000), Saunders (2000) or Berkowitz and O'Brien (2002) who show that simple GARCH-type univariate models can adequately model the VaR of a global financial institution.

The definition "nominal h-day loss that will not be exceeded in $100 \cdot \alpha$ portfolio realizations out of 100" refers to the statistical notion of a quantile of a density distribution. More precisely, the VaR at time t for the forthcoming $[t+1, t+h]$ time horizon is the expected quantile at the α level for the density distributions of the h-day portfolio returns. In a parametric VaR framework of the GARCH-type, and given a collection of daily demeaned returns r_t, a conditional volatility model for g_t such that $r_t = \sqrt{g_t}\varepsilon_t$ and a density distribution for the standardized residuals ε_t, the parametric one-day VaR at time t is immediately given by:

$$VaR_t = Z_\alpha \sqrt{g_t} \tag{26}$$

where Z_α is the quantile at $100 \cdot \alpha$ percent of the standardized density distribution. Following Giot and Laurent (2002) we focus on the skewed Student density distribution which allows for excess kurtosis and skewness in the distribution of returns.[13] According to Lambert and Laurent (2001), the innovation process ε is said to be (standardized) skewed Student distributed if:

$$f(\varepsilon|\xi,\upsilon) = \begin{cases} \frac{2}{\xi+\frac{1}{\xi}} sg\left[\xi\left(s\varepsilon+m\right)|\upsilon\right] & \text{if} \quad \varepsilon < -\frac{m}{s} \\ \frac{2}{\xi+\frac{1}{\xi}} sg\left[\left(s\varepsilon+m\right)/\xi|\upsilon\right] & \text{if} \quad \varepsilon \geq -\frac{m}{s} \end{cases} \tag{27}$$

where $g(.|\upsilon)$ is the symmetric (unit variance) Student density and ξ is the asymmetry coefficient;[14] m and s^2 are respectively the mean and the variance of the non-standardized skewed Student.

Because the goal of the paper is to compare the information content of the volatility forecasts based on the VIX/VXN indexes, RiskMetrics and GJR-GARCH methods, we thus compare the performance of the competing VaR models:

VIX/VXN

$$VaR_t = Z_\alpha \sqrt{g_t} \tag{28}$$

where $g_t = \omega + \eta\,\sigma^2_{imp,t-1}$;

RiskMetrics

$$VaR_t = Z_\alpha \sqrt{g_t} \tag{29}$$

where $g_t = 0.04r^2_{t-1} + 0.94g_{t-1}$;

GJR-GARCH

$$VaR_t = Z_\alpha \sqrt{g_t} \tag{30}$$

where $g_t = \omega + \alpha_1 r^2_{t-1} + \alpha_n r^2_{t-1} d_{t-1} + \delta_1 g_{t-1}$.

In all cases, Z_α is the relevant quantile at the $100 \cdot \alpha$ % level from the skewed Student distribution as given in Equation (27).

In order to the back-test the VaR results, we first use the familiar Kupiec (1995) LR test. Given the ex-post (i.e. at date $t+1$) observed returns $\{r_{t+1}\}$ and ex-ante (i.e. at date t) forecasted $\{VaR_t\}$, the empirical failure rate $\widehat{f}$ is given by the number of returns smaller than the VaR. If the VaR model is correctly specified, this proportion must be equal to α. More precisely and in the binomial framework, Kupiec (1995) shows that the hypothesis $H_0 : f = \alpha$ against $H_1 : f \neq \alpha$, can be tested with the LR statistic $LR = -2\,ln(\alpha^{T-N}(1-\alpha)^N) + 2\,ln\Big((1-(N/T))^{T-N}(N/T)^N\Big)$, where N is the number of VaR violations, T is the total number of observations and f is the theoretical failure rate. Under the null hypothesis that f is the true failure rate, the LR test statistic is asymptotically distributed as a $\chi^2(1)$.

Secondly we use the independence and conditional coverage tests put forward by Christoffersen (1998). Indeed, the Kupiec (1995) LR test assesses only the equality between the proportion of VaR violations and the expected α level (what is refereed to as the unconditional coverage in Christoffersen, 1998). In a risk management framework, it is also of paramount importance that the VaR violations be uncorrelated over time which leads to the independence and conditional coverage tests based on the evaluation of interval forecasts. Using the same notation as Christoffersen (1998), we defined the indicator sequence of VaR violations as $\{I_t\}$ where I_t is a dummy variable that is equal to 1 if there is a VaR violation at time t (i.e. r_t is smaller than VaR_t) and is equal to 0 if there is no VaR violation at time t. If $\pi_{i,j}$ is the transition probability for two successive I_t dummy variables, i.e. $\pi_{i,j} = P(I_t = j | I_{t-1} = i)$, then the approximate likelihood function for the sequence of I_t is equal to:

$$L_1 = \pi_{0,0}^{n_{0,0}} \pi_{0,1}^{n_{0,1}} \pi_{1,0}^{n_{1,0}} \pi_{1,1}^{n_{1,1}} \tag{31}$$

where $n_{i,j}$ is the number of observations with value i followed by value j. The approximate maximized likelihood function is then equal to:

$$\widehat{L}_1 = \left[\frac{n_{0,0}}{n_{0,0}+n_{0,1}}\right]^{n_{0,0}} \left[\frac{n_{0,1}}{n_{0,0}+n_{0,1}}\right]^{n_{0,1}} \left[\frac{n_{1,0}}{n_{1,0}+n_{1,1}}\right]^{n_{1,0}} \left[\frac{n_{1,1}}{n_{1,0}+n_{1,1}}\right]^{n_{1,1}} \tag{32}$$

If the $\{I_t\}$ sequence, i.e. the sequence of VaR violations, is (first-order) independent, then $\pi_{0,0} = 1-\pi$, $\pi_{0,1} = \pi$, $\pi_{1,0} = 1-\pi$ and $\pi_{1,1} = \pi$. This gives the likelihood under the null of first-order independence:

$$L_2 = (1-\pi)^{n_{0,0}+n_{1,0}} \pi^{n_{0,1}+n_{1,1}} \tag{33}$$

which is then estimated as:

$$\widehat{L}_2 = \left[1 - \frac{n_{0,1}+n_{1,1}}{n_{0,0}+n_{1,0}+n_{0,1}+n_{1,1}}\right]^{n_{0,0}+n_{1,0}} \left[\frac{n_{0,1}+n_{1,1}}{n_{0,0}+n_{1,0}+n_{0,1}+n_{1,1}}\right]^{n_{0,1}+n_{1,1}} \tag{34}$$

The LR test statistic for (first-order) independence in the VaR violations is equal to:

$$LR_{ind} = -2(ln(\widehat{L}_2) - ln(\widehat{L}_1)) \sim \chi^2(1). \tag{35}$$

Provided that we condition on the first observation in the test for unconditional coverage, the LR statistic for conditional coverage (i.e. the joint hypothesis of unconditional coverage and independence) is equal to:

$$LR_{cc} = LR_{uc} + LR_{ind} \sim \chi^2(2) \tag{36}$$

where LR_{uc} is the LR statistic for unconditional coverage (computed above for the Kupiec, 1995 test). See additional details and proofs in Christoffersen (1998).

Thirdly we use some of the tests recently suggested by Berkowitz (2001). These tests rely on the density forecast framework such as discussed in Diebold, Gunther, and Tay (1998), Bauwens, Giot, Grammig, and Veredas (2000) and Berkowitz (2001). While we refer the reader to these papers for a full discussion of the density forecast approach, we can summarize the methodology as:

- For all ex-post observed returns $\{r_{t+1}\}$ we first compute the probability integral transformation:

$$x_{t+1} = \int_{-\infty}^{r_{t+1}/\sqrt{g_{t+1}}} f(u)\,du \tag{37}$$

  where $f(u)$ is the ex-ante (i.e. computed at time t) standardized skewed Student density distribution as given by Equation (27). Note that g_{t+1} is also computed at time t using one of the three specifications defined above.

- If the VaR model is correctly specified, then it can be shown that x_t should be IID and distributed uniformly on $[0,1]$. Diebold, Gunther, and Tay (1998) suggest a graphical methodology to check this (the histogram should be flat, the ACF should indicate the absence of autocorrelation,...).

- Berkowitz (2001) advocates a likelihood-ratio testing framework as $z_t = \Phi^{-1}(x_t)$ should be IID $N(0,1)$ when x_t is IID uniform over $[0,1]$, i.e. when the VaR model is correctly specified.[15] In this paper, we estimate the following regression:

$$z_t = \beta_0 + \rho_1 z_{t-1} + u_t \tag{38}$$

  where $u_t \sim N(0,\sigma^2)$. We first test H0: $\rho_1 = 0$, i.e. that the z_t are uncorrelated and then the much more restrictive H0: $\beta_0 = 0$ AND $\rho_1 = 0$ AND $\sigma^2 = 1$, i.e. that the z_t are uncorrelated and that the error term u_t is indeed $N(0,1)$. Last we also look at whether past squared returns are relevant in the extended regression $z_t = \beta_0 + \rho_1 z_{t-1} + \eta_1 r_{t-1}^2 + u_t$ by focusing on the t-statistic of coefficient η_1.

We apply the VaR methodology and tests to our return series for the S&P100 and NASDAQ100 indexes. The methodology is similar to what has been discussed in Section III. Because we first need one-day out-of-sample volatility forecasts based successively on the implied volatility, RiskMetrics and GJR-GARCH models, we estimate the models defined by Equation (1) (volatility forecast based on the VIX/VNX indexes, $h = 1$), Equation (2) (volatility forecast based on the RiskMetrics method, $h = 1$) and Equations (6) and (5) (volatility forecast based on the GJR-GARCH model, $h = 1$) using a skewed Student distribution for the standardized error term. These are out-of-sample volatility and VaR estimates. Indeed the estimation sample is increased as each new observation enters the information set while the forecasts are made for time $t+1$ when the information sample ends at time t. This is exactly what has been done in Section III to get the volatility forecasts for the encompassing regressions. However we fit in this case a skewed Student density distribution to model adequately the left and right quantiles, i.e. VaR levels as in Equations (28), (29) and (30). At the end of this procedure, we thus have one-day ahead (out-of-sample) VaR forecasts for the August 1st, 1989 - June 28th, 2002 time period (S&P100 index) and September 3rd, 1996 - June 28th, 2002 time period (NASDAQ100 index). These VaR forecasts $\{VaR_t\}$ can then be back-tested against the observed returns $\{r_{t+1}\}$.

For the back-testing of the VaR forecasts, we first compute the empirical failure rates and Kupiec (1995) LR tests for the right and left quantiles at 10%, 5% and 1%. Note that the left quantile is relevant for a trader who is long the index (trading losses occur on the left side of the returns density), while the right quantile is relevant for a trader who is short the index (trading losses occur on the right side of the returns density). Secondly we compute the independence and conditional coverage tests detailed above (see Equations (35) and (36)). Finally, we compute the x_t as in Equation (37), then the $z_t = \Phi^{-1}(x_t)$ and estimate Equation (38).

The empirical results for the three models are given in Table XII (S&P100 index) and Table XIII (NASDAQ100 index). For the S&P100 index, the VIX based and GJR-GARCH models deliver accurate VaR results, although they are overly pessimistic for the right tail at 1% ($H_0 : f = \alpha$ is rejected, although $H_0 : f < \alpha$ is not rejected). Regarding the (skewed Student) RiskMetrics model, there are too many VaR violations at 10% and 5%. The P-values for the independence test (rows labelled by I in the table) show that the null hypothesis of first-order independence is not rejected for all models except for the RiskMetrics model with the left tail at 1%. Note that when $\alpha = 1\%$ there are many cases where the test cannot be used because one has $n11 = 0$ (i.e. there are no successive VaR violations). The P-values for the conditional coverage test (rows labelled by CC in the table) show that the null hypothesis of conditional coverage is usually not rejected for the VIX and GJR-GARCH based volatility forecasts.[16] The skewed-Student RiskMetrics model performs rather poorly. When the VIX index is used as the volatility forecast, $z_t = \Phi^{-1}(\int_{-\infty}^{r_t/\sqrt{g_t}} f(u)\,du)$ is uncorrelated, which is not the case for the RiskMetrics and GJR-GARCH volatility forecasts. For all three volatility forecasts, η_1 is significant in the $z_t = \beta_0 + \rho_1 z_{t-1} + \eta_1 r_{t-1}^2 + u_t$ regression which suggests that some volatility clustering has not been correctly taken into account.

Regarding the NASDAQ100 index, the VXN based and RiskMetrics models deliver (qualitatively) the same results as all empirical failure rates are equal to the theoretical values, except when $\alpha = 10\%$ in the left tail where the models are too optimistic. The P-values for the independence and conditional coverage tests (rows labelled by I and CC) show that the corresponding null hypotheses are not rejected, except when $\alpha = 10\%$ in the left tail. When $\alpha = 1\%$, $n11$ is again equal to 0 and the tests cannot be used. In both cases, z_t is not autocorrelated and lagged squared returns are not significant in the $z_t = \beta_0 + \rho_1 z_{t-1} + \eta_1 r_{t-1}^2 + u_t$ regression. For the GJR-GARCH based method, there are some large discrepancies on the right side of the density distribution as all empirical failure rates are statistically different from the theoretical levels. While the null hypothesis of independence is not rejected, the null of conditional coverage is almost always rejected. The z_t are not correlated and η_1 is not significant either.

Quite surprisingly and although the skewed Student density distribution is very flexible, the null hypothesis that H0: $\beta_0 = 0$ AND $\rho_1 = 0$ AND $\sigma^2 = 1$ is rejected for all models and for both indexes (except marginally for the S&P100 and GJR-GARCH model).[17] Regarding the volatility forecasts made from the VIX/VXN indexes, these tests thus show that they are meaningful inputs in VaR models as the number of VaR violations is correctly modelled in most cases, the null hypotheses of independence and conditional coverage are usually not rejected and the probability integral transformations are not correlated.

V. Conclusion

In this paper we looked at the information content of three competing volatility forecasts with respect to ex-post realized volatility at a 5-, 10- and 22-day horizon. We considered the volatility forecast based on the VIX/VXN implied volatility indexes and two measures (RiskMetrics and GJR-GARCH) of volatility based on historical returns. Next to encompassing regressions (OLS and Instrumental Variable regressions) and tests of efficiency and absence of bias, we also looked at the information content of the volatility forecasts when market risk must be quantified. Our empirical application dealt with the S&P100 index (for which the VIX index gives the 22-day implied volatility embedded in option contracts) and with the NASDAQ100 index (VXN index is the associated implied volatility index). Thus the focus of the paper was on the added value of relatively simple and easy-to-use volatility forecasts as given by the levels of implied volatility indexes (which are computed and widely made available by the options and futures exchanges, the CBOE in our case).

Our results support the hypothesis according to which implied volatility (as given by the VIX/VXN indexes) has a high information content and provides meaningful inputs into market risk models of the VaR type. In the out-of-sample study and while (almost) no volatility forecast is both unbiased and efficient, we find that VIX/VXN based volatility forecasts are the closest to being unbiased, have the highest information content (with a statistical fit that increases with the time horizon) and are meaningful inputs in VaR models as the number of VaR violations is correctly modelled in most cases, the null hypotheses of independence and conditional coverage are usually not rejected and the probability integral transformations are not correlated. Moreover, the in-sample study shows that GARCH-type forecasts have little incremental information over that contained in the implied volatility indexes.

Our results, along with recent results given by Blair, Poon, and Taylor (2001), Claessen and Mittnik (2002) or Giot (2003), show that implied volatility indexes do provide accurate and meaningful information as to future volatility forecasts. Moreover, they compete very favorably with volatility forecasts based on historical returns and are simple to use, both in a volatility forecasting and market risk evaluation framework. Therefore it can be argued that exchanges, in the U.S. and in other countries, should design implied volatility indexes (along the lines of the VIX or VXN indexes at the CBOE) and make them available to the investing community. Indeed, with the growing importance of index related funds (for example passive investment funds that track the S&P500 index), there is a growing need for valuable information regarding expected volatility and market risk for the underlying indexes. While more sophisticated econometric models which deal with intraday data and realized volatility computed from squared intraday returns are also noteworthy (see Andersen, Bollerslev, Diebold, and Ebens, 2001, Giot and Laurent, 2001, Andersen, Bollerslev, Diebold, and Labys, 2002, Martens and Zein, 2002 or Neely, 2002), they are more complicated to use and estimate. Therefore, volatility forecasts based on implied volatility indexes computed by the exchanges do provide an interesting alternative and deliver meaningful volatility/market risk information without the need for complex econometric models.

References

Alexander, C., 2001, *Market models*. (Wiley).

Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens, 2001, The distribution of stock return volatility, *Journal of Financial Economics* 61, 43–76.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys, 2002, Modelling and forecasting realized volatility, Mimeo.

Bauwens, L., P. Giot, J. Grammig, and D. Veredas, 2000, A Comparison of Financial Duration Models Via Density Forecasts, CORE DP 2060.

Berkowitz, J., 2001, Testing density forecasts, with applications to risk management, *Journal of Business and Economic Statistics* 19, 465–474.

Berkowitz, J., and J. O'Brien, 2002, How accurate are Value-at-Risk models at commercial banks?, *Journal of Finance* 57, 1093–1112.

Black, F., 1976, Studies of Stock Market Volatility Changes, *Proceedings of the American Statistical Association, Business and Economic Statistics Section* pp. 177–181.

Black, F., and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–659.

Blair, B.J., S. Poon, and S.J. Taylor, 2001, Forecasting S&P100 volatility: the incremental information content of implied volatilities and high-frequency index returns, *Journal of Econometrics* 105, 5–26.

Bollerslev, T., 1986, Generalized autoregressive condtional heteroskedasticity, *Journal of Econometrics* 31, 307–327.

Bollerslev, T., R.Y. Chou, and K.F. Kroner, 1992, ARCH modeling in finance: a review of the theory and empirical evidence, *Journal of Econometrics* 52, 5–59.

Canina, L., and S. Figlewski, 1993, The information content of implied volatility, *Review of Financial Studies* 6, 659–681.

Christensen, B.J., and N.R. Prabhala, 1998, The relation between implied and realized volatility, *Journal of Financial Economics* 50, 125–150.

Christoffersen, P.F., 1998, Evaluating interval forecasts, *International Economic Review* 39, 841–862.

Claessen, H., and S. Mittnik, 2002, Forecasting stock market volatility and the informational efficiency of the DAX-index options market, CFS Working Paper 2002/04.

Day, T.E., and C.M. Lewis, 1992, Stock market volatility and the information content of stock index options, *Journal of Econometrics* 52, 267–287.

Diebold, F. X., T. A. Gunther, and A.S. Tay, 1998, Evaluating Density Forecasts, with Applications to Financial Risk Management, *International Economic Review* 39, 863–883.

Ding, Z., C. W. J. Granger, and R.F. Engle, 1993, A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance* 1, 83–106.

Ederington, L., and W. Guan, 2002, Is implied volatility an informationally efficient and effective predictor of future volatility?, *Journal of Risk* 4, 3.

Engle, R.F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* 50, 987–1007.

Engle, R.F., 1995, *ARCH selected readings*. (Oxford University Press Oxford).

Figlewski, S., 1997, Forecasting volatility, *Financial Markets, Institutions and Instruments* 6, 2–87.

Fleming, J., Ostdiek B., and R.E. Whaley, 1995, Predicting stock market volatility: a new measure, *Journal of Futures Markets* 15, 265–302.

French, K., G. Schwert, and R. Stambaugh, 1987, Expected Stock Returns and Volatility, *Journal of Financial Economics* 19, 3–29.

Giot, P., 2002, Implied volatility indices as leading indicators of stock index returns?, CORE DP 2002/50.

Giot, P., 2003, The information content of implied volatility in agricultural commodity markets, *Forthcoming in the Journal of Futures Markets*.

Giot, P., and S. Laurent, 2001, Modelling daily Value-at-Risk using realized volatility and ARCH type models, Maastricht University METEOR RM/01/026.

Giot, P., and S. Laurent, 2002, Value-at-Risk for long and short trading positions, *Forthcoming in the Journal of Applied Econometrics*.

Glosten, L.R., R. Jagannathan, and D.E. Runkle, 1993, On the relation between expected excess return on stocks, *Journal of Finance* 48, 1779–1801.

Hull, J.C., 2000, *Options, futures and other derivatives*. (Prentice-Hall).

Jorion, P., 1995, Predicting volatility in the Foreign Exchange market, *Journal of Finance* 50, 507–528.

Jorion, P., 2000, *Value-at-Risk*. (McGraw-Hill).

Kupiec, P., 1995, Techniques for verifying the accuracy of risk measurement models, *Journal of Derivatives* 2, 173–84.

Lambert, P., and S. Laurent, 2001, Modelling financial time series using GARCH-type models and a skewed Student density, Mimeo, Université de Liège.

Martens, M., and J. Zein, 2002, Predicting financial volatility: high-frequency time-series forecasts vis-a-vis implied volatility, Mimeo, Erasmus University Rotterdam.

Neely, C.J., 2002, Forecasting foreign exchange volatility: is implied volatility the best we can do?, Federal Reserve Bank of St.Louis, WP 2002-017.

Nelson, D.B., 1991, Conditional heteroskedasticity in asset returns: a new approach, *Econometrica* 59, 349–370.

Pagan, A., and G. Schwert, 1990, Alternative models for conditional stock volatility, *Journal of Econometrics* 45, 267–290.

Poon, S., and C. Granger, 2003, Forecasting volatility in financial markets: a review, *Forthcoming in the Journal of Economic Literature*.

Saunders, A., 2000, *Financial institutions management*. (McGraw-Hill).

Whaley, R.E., 1993, Derivatives on market volatility: hedging tools long overdue, *Journal of Derivatives* 1, 71–84.

Whaley, R.E., 2000, The investor fear gauge, Mimeo, The Fuqua School of Business, Duke University.

Xu, X., and S.J. Taylor, 1995, Conditional volatility and the informational efficiency of the PHLX currency options market, *Journal of Banking and Finance* 19, 803–821.

Table I

MAE, RMSE and Proportion of variance (S&P100 index)

	Implied volatility			RiskMetrics			GJR-GARCH		
Overlapping data									
	MAE	RMSE	P	MAE	RMSE	P	MAE	RMSE	P
5-day	0.72	0.96	0.33	0.70	1.00	0.27	0.75	1.01	0.26
10-day	0.87	1.15	0.42	0.85	1.23	0.32	0.95	1.27	0.28
22-day	1.18	1.52	0.45	1.19	1.68	0.33	1.40	1.78	0.25
Non-overlapping data									
	Implied volatility			RiskMetrics			GJR-GARCH		
	MAE	RMSE	P	MAE	RMSE	P	MAE	RMSE	P
5-day	0.72	0.98	0.31	0.69	1.01	0.26	0.75	1.03	0.24
10-day	0.87	1.17	0.39	0.83	1.23	0.32	0.92	1.27	0.28
22-day	1.13	1.50	0.44	1.17	1.67	0.32	1.40	1.80	0.20

Mean absolute error (MAE), root mean square error (RMSE) and proportion of variance (P) for the implied volatility/RiskMetrics/GJR-GARCH volatility forecasts vs the realized volatility at the 5-, 10- and 22-day horizon. The top panel presents results for overlapping realized volatility data (see text of paper for details), while the bottom panel shows results for non-overlapping realized volatility data. Results are for the S&P100 index and for the August 1st, 1989 - June 28th, 2002 time period.

Table II

MAE, RMSE and Proportion of variance (NASDAQ100 index)

	Overlapping data								
	Implied volatility			RiskMetrics			GJR-GARCH		
	MAE	RMSE	P	MAE	RMSE	P	MAE	RMSE	P
5-day	1.35	1.90	0.51	1.54	2.17	0.37	1.40	2.03	0.44
10-day	1.59	2.28	0.58	1.91	2.73	0.40	1.82	2.62	0.45
22-day	2.07	2.87	0.64	2.57	3.62	0.42	2.58	3.70	0.40
	Non-overlapping data								
	Implied volatility			RiskMetrics			GJR-GARCH		
	MAE	RMSE	P	MAE	RMSE	P	MAE	RMSE	P
5-day	1.33	1.85	0.53	1.53	2.10	0.40	1.36	1.94	0.49
10-day	1.65	2.37	0.55	1.92	2.75	0.39	1.84	2.69	0.42
22-day	2.13	2.98	0.62	2.76	4.13	0.27	2.91	4.33	0.20

Mean absolute error (MAE), root mean square error (RMSE) and proportion of variance (P) for the implied volatility/RiskMetrics/GJR-GARCH volatility forecasts vs the realized volatility at the 5-, 10- and 22-day horizon. The top panel presents results for overlapping realized volatility data (see text of paper for details), while the bottom panel shows results for non-overlapping realized volatility data. Results are for the NASDAQ100 index and for the September 3rd, 1996 - June 28th, 2002 time period.

Table III

Realized volatility: basic regressions (S&P100 index)

5-day volatility forecasts				
Model	β_0	β_1	R^2	DW
VIX	-0.290 (0.118)	0.972 (0.047)	0.399	1.858
RM	0.443 (0.100)	0.738 (0.043)	0.315	1.905
GJR	-0.145 (0.141)	0.955 (0.059)	0.287	2.021
10-day volatility forecasts				
Model	β_0	β_1	R^2	DW
VIX	-0.181 (0.197)	0.930 (0.055)	0.468	1.913
RM	0.748 (0.174)	0.732 (0.053)	0.374	1.989
GJR	-0.347 (0.281)	1.024 (0.084)	0.317	1.967
22-day volatility forecasts				
Model	β_0	β_1	R^2	DW
VIX	0.019 (0.394)	0.906 (0.076)	0.497	2.089
RM	1.390 (0.347)	0.692 (0.071)	0.394	2.143
GJR	-0.046 (0.724)	0.946 (0.147)	0.222	1.565

Results for the regression of the realized volatility at the 5-, 10- and 22-day horizon vs the implied volatility (VIX)/RiskMetrics (RM)/GJR-GARCH (GJR) volatility forecasts, i.e. $RV_{h,t} = \beta_0 + \beta_1 X_t + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and X_t is successively the VIX/RM/GJR volatility forecasts at the h-day horizon. Results are for the S&P100 index (non-overlapping realized volatility data) and for the August 1st, 1989 - June 28th, 2002 time period.

Table IV

Realized volatility: basic regressions (NASDAQ100 index)

5-day volatility forecasts				
Model	β_0	β_1	R^2	DW
VXN	-1,200 (0,356)	1,209 (0,064)	0,552	1,719
RM	0,763 (0,328)	0,822 (0,056)	0,423	1,658
GJR	0,363 (0,308)	0,986 (0,058)	0,497	2,020
10-day volatility forecasts				
Model	β_0	β_1	R^2	DW
VXN	-0,964 (0,659)	1,132 (0,084)	0,559	2,051
RM	1,705 (0,592)	0,762 (0,072)	0,439	1,972
GJR	1,166 (0,607)	0,937 (0,083)	0,467	2,289
22-day volatility forecasts				
Model	β_0	β_1	R^2	DW
VXN	-0,964 (1,245)	1,096 (0,105)	0,628	2,332
RM	4,202 (1,190)	0,639 (0,097)	0,404	2,227
GJR	4,081 (1,338)	0,745 (0,127)	0,351	2,162

Results for the regression of the realized volatility at the 5-, 10- and 22-day horizon vs the implied volatility (VXN)/RiskMetrics (RM)/GJR-GARCH (GJR) volatility forecasts, i.e. $RV_{h,t} = \beta_0 + \beta_1 X_t + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and X_t is successively the VXN/RM/GJR volatility forecasts at the h-day horizon. Results are for the NASDAQ100 index (non-overlapping realized volatility data) and for the September 3rd, 1996 - June 28th, 2002 time period.

Table V
Realized volatility: encompassing regressions (S&P100 index)

Model	β_0	β_1	β_2	R^2	DW	UE	U
5-day volatility forecasts							
VIX	-0.253 (0.120)	0.898 (0.065)	0.070 (0.043)	0.402	1.980	0	0
RM	0.459 (0.101)	0.663 (0.072)	0.072 (0.055)	0.316	2.001	0	0
GJR	-0.072 (0.172)	0.879 (0.118)	0.049 (0.066)	0.288	2.070	0	0.009
10-day volatility forecasts							
VIX	-0.139 (0.201)	0.856 (0.087)	0.071 (0.064)	0.471	2.016	0	0.005
RM	0.754 (0.177)	0.714 (0.124)	0.017 (0.104)	0.374	2.004	0	0
GJR	0.259 (0.360)	0.594 (0.182)	0.267 (0.101)	0.331	2.202	0	0.001
22-day volatility forecasts							
VIX	-0.014 (0.415)	0.962 (0.156)	-0.051 (0.122)	0.499	1.928	0.003	0.930
RM	1.359 (0.361)	0.848 (0.320)	-0.147 (0.292)	0.396	1.990	0	0
GJR	1.919 (0.746)	-0.039 (0.222)	0.622 (0.111)	0.367	2.170	0	0

Results for the encompassing regression of the realized volatility at the 5-, 10- and 22-day horizon vs the implied volatility (VIX)/RiskMetrics (RM)/GJR-GARCH (GJR) volatility forecasts and lagged realized volatility, i.e. $RV_{h,t} = \beta_0 + \beta_1 X_t + \beta_2 RV_{h,t-1} + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and X_t is successively the VIX/RM/GJR volatility forecasts at the h-day horizon. UE gives the P-value for the Fisher test of unbiasedness and efficiency, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$ AND $\beta_2 = 0$, while U gives the P-value for the Fisher test of unbiasedness, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$. Results are for the S&P100 index (non-overlapping realized volatility data) and for the August 1st, 1989 - June 28th, 2002 time period.

Table VI

Realized volatility: encompassing regressions (NASDAQ100 index)

5-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VXN	-1.083 (0.371)	1.101 (0.103)	0.087 (0.064)	0.555	1.858	0.004	0.002
RM	0.881 (0.328)	0.567 (0.101)	0.243 (0.080)	0.440	1.975	0.000	0.000
GJR	0.162 (0.328)	1.220 (0.145)	-0.183 (0.104)	0.502	1.872	0.020	0.042
10-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VXN	-1.087 (0.704)	1.216 (0.157)	-0.067 (0.103)	0.556	1.956	0.428	0.291
RM	1.771 (0.604)	0.728 (0.176)	0.028 (0.153)	0.436	2.003	0.011	0.015
GJR	0.700 (0.655)	1.433 (0.260)	-0.386 (0.190)	0.479	2.082	0.001	0.006
22-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VXN	-2.101 (1.279)	1.654 (0.219)	-0.456 (0.157)	0.666	1.888	0.032	0.013
RM	4.388 (1.224)	0.778 (0.395)	-0.152 (0.390)	0.396	2.179	0.005	0.003
GJR	4.277 (1.345)	0.285 (0.310)	0.388 (0.245)	0.367	2.317	0.001	0.005

Results for the encompassing regression of the realized volatility at the 5-, 10- and 22-day horizon vs the implied volatility (VXN)/RiskMetrics (RM)/GJR-GARCH (GJR) volatility forecasts and lagged realized volatility, i.e. $RV_{h,t} = \beta_0 + \beta_1 X_t + \beta_2 RV_{h,t-1} + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and X_t is successively the VXN/RM/GJR volatility forecasts at the h-day horizon. UE gives the P-value for the Fisher test of unbiasedness and efficiency, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$ AND $\beta_2 = 0$, while U gives the P-value for the Fisher test of unbiasedness, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$. Results are for the S&P100 index (non-overlapping realized volatility data) and for the September 3rd, 1996 - June 28th, 2002 time period.

Table VII

Realized volatility: encompassing regressions (S&P100 index)

5-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VIX	0.008 (0.140)	0.811 (0.086)	-0.020 (0.055)	0.256	2.000	0	0.001
RM	0.567 (0.124)	0.617 (0.105)	-0.029 (0.071)	0.156	2.013	0	0
GJR	0.028 (0.204)	0.771 (0.132)	-0.036 (0.072)	0.154	2.044	0	0
10-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VIX	0.239 (0.218)	0.751 (0.106)	-0.012 (0.083)	0.336	2.095	0	0.043
RM	0.792 (0.205)	0.774 (0.172)	-0.147 (0.136)	0.237	2.016	0	0
GJR	0.156 (0.362)	0.640 (0.170)	0.095 (0.102)	0.213	2.124	0	0.001
22-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VIX	0.042 (0.410)	1.017 (0.167)	-0.197 (0.138)	0.473	1.545	0	0.966
RM	1.226 (0.387)	1.123 (0.385)	-0.494 (0.347)	0.306	1.733	0.001	0.001
GJR	-0.043 (0.845)	0.575 (0.243)	0.255 (0.134)	0.283	1.780	0	0

Results for the encompassing regression of the realized volatility at the 5-, 10- and 22-day horizon vs the implied volatility (VIX)/RiskMetrics (RM)/GJR-GARCH (GJR) volatility forecasts and lagged realized volatility, i.e. $RV_{h,t} = \beta_0 + \beta_1 X_t + \beta_2 RV_{h,t-1} + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and X_t is successively the VIX/RM/GJR volatility forecasts at the h-day horizon. UE gives the P-value for the Fisher test of unbiasedness and efficiency, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$ AND $\beta_2 = 0$, while U gives the P-value for the Fisher test of unbiasedness, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$. Results are for the S&P100 index (non-overlapping realized volatility data) and for the August 1st, 1989 - December 31th, 1996 time period.

Table VIII

Realized volatility: encompassing regressions (S&P100 index)

5-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VIX	-0.073 (0.390)	0.820 (0.150)	0.113 (0.067)	0.191	1.967	0	0.006
RM	1.346 (0.274)	0.314 (0.132)	0.176 (0.083)	0.120	2.005	0	0
GJR	0.776 (0.285)	0.874 (0.188)	-0.132 (0.112)	0.168	1.961	0.003	0.001
10-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VIX	0.513 (0.668)	0.711 (0.192)	0.099 (0.100)	0.192	1.933	0.015	0.178
RM	2.397 (0.493)	0.165 (0.218)	0.230 (0.157)	0.114	2.010	0	0
GJR	0.798 (0.593)	1.281 (0.340)	-0.385 (0.205)	0.195	1.991	0.002	0.001
22-day volatility forecasts							
Model	β_0	β_1	β_2	R^2	DW	UE	U
VIX	2.950 (1.303)	0.457 (0.312)	0.036 (0.196)	0.101	1.925	0.070	0.085
RM	4.597 (0.968)	-0.107 (0.507)	0.346 (0.431)	0.068	2.013	0	0
GJR	3.695 (1.086)	0.393 (0.374)	0.048 (0.237)	0.085	2.003	0.001	0.003

Results for the encompassing regression of the realized volatility at the 5-, 10- and 22-day horizon vs the implied volatility (VIX)/RiskMetrics (RM)/GJR-GARCH (GJR) volatility forecasts and lagged realized volatility, i.e. $RV_{h,t} = \beta_0 + \beta_1 X_t + \beta_2 RV_{h,t-1} + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and X_t is successively the VIX/RM/GJR volatility forecasts at the h-day horizon. UE gives the P-value for the Fisher test of unbiasedness and efficiency, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$ AND $\beta_2 = 0$, while U gives the P-value for the Fisher test of unbiasedness, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$. Results are for the S&P100 index (non-overlapping realized volatility data) and for the January 2nd, 1997 - June 28th, 2002 time period.

Table IX
Realized volatility: IV regressions for VIX forecasts (S&P100 index)

Basic regression

Horizon	β_0	β_1	R^2	DW
5-day	-0.258 (0.134)	0.959 (0.054)	0.331	1.682
10-day	-0.231 (0.246)	0.945 (0.070)	0.364	1.575
22-day	-0.123 (0.567)	0.938 (0.111)	0.331	1.435

Encompassing regression

Horizon	β_0	β_1	β_2	R^2	DW	UE	U
5-day	-0.221 (0.137)	0.875 (0.078)	0.080 (0.048)	0.352	2.025	0.000	0.000
10-day	-0.198 (0.258)	0.890 (0.125)	0.051 (0.084)	0.403	2.030	0.000	0.020
22-day	-0.307 (0.715)	1.117 (0.340)	-0.157 (0.241)	0.411	1.971	0.007	0.908

Results for the Instrumental Variables regression of the realized volatility at the 5-, 10- and 22-day horizon vs the implied volatility forecasts and lagged realized volatility, i.e. (a) $RV_{h,t} = \beta_0 + \beta_1 INS_{h,t} + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and $INS_t = \widehat{\sigma}_{imp,h,t}$ is the forecasted implied volatility from the previous regression $\sigma_{imp,h,t} = \beta_0' + \beta_1' \sigma_{imp,h,t-1} + e_t$ and (b) $RV_{h,t} = \beta_0 + \beta_1 INS_{h,t} + \beta_2 RV_{h,t-1} + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and $INS_t = \widehat{\sigma}_{imp,h,t}$ is the forecasted implied volatility from the previous regression $\sigma_{imp,h,t} = \beta_0' + \beta_1' \sigma_{imp,h,t-1} + \beta_2' RV_{h,t-1} + e_t$. UE gives the P-value for the Fisher test of unbiasedness and efficiency in the encompassing regression, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$ AND $\beta_2 = 0$, while U gives the P-value for the Fisher test of unbiasedness, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$. Results are for the S&P100 index (non-overlapping realized volatility data) and for the August 1st, 1989 - June 28th, 2002 time period.

Table X
Realized volatility: IV regressions for VXN forecasts (NASDAQ100 index)

Basic regression

Horizon	β_0	β_1	R^2	DW
5-day	-0.860 (0.424)	1.144 (0.076)	0.436	1.575
10-day	-0.802 (0.813)	1.111 (0.104)	0.444	1.660
22-day	-0.044 (1.848)	1.017 (0.157)	0.398	1.702

Encompassing regression

Horizon	β_0	β_1	β_2	R^2	DW	UE	U
5-day	-0.584 (0.444)	0.923 (0.130)	0.173 (0.076)	0.471	2.003	0.013	0.006
10-day	-0.859 (0.894)	1.148 (0.218)	-0.030 (0.135)	0.471	2.032	0.750	0.619
22-day	-0.966 (2.449)	1.381 (0.517)	-0.282 (0.341)	0.424	2.021	0.843	0.668

Results for the Instrumental Variables regression of the realized volatility at the 5-, 10- and 22-day horizon vs the implied volatility forecasts and lagged realized volatility, i.e. (a) $RV_{h,t} = \beta_0 + \beta_1 INS_{h,t} + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and $INS_t = \widehat{\sigma}_{imp,h,t}$ is the forecasted implied volatility from the previous regression $\sigma_{imp,h,t} = \beta_0' + \beta_1' \sigma_{imp,h,t-1} + e_t$ and (b) $RV_{h,t} = \beta_0 + \beta_1 INS_{h,t} + \beta_2 RV_{h,t-1} + e_t$ where $RV_{h,t}$ is the realized volatility at the h-day horizon and $INS_t = \widehat{\sigma}_{imp,h,t}$ is the forecasted implied volatility from the previous regression $\sigma_{imp,h,t} = \beta_0' + \beta_1' \sigma_{imp,h,t-1} + \beta_2' RV_{h,t-1} + e_t$. UE gives the P-value for the Fisher test of unbiasedness and efficiency in the encompassing regression, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$ AND $\beta_2 = 0$, while U gives the P-value for the Fisher test of unbiasedness, i.e. H0: $\beta_0 = 0$ AND $\beta_1 = 1$. Results are for the NASDAQ100 index (non-overlapping realized volatility data) and for the September 3rd, 1996 - June 28th, 2002 time period.

Table XI
In-sample Log-likelihoods

	S&P100 index	NASDAQ100 index
GARCH(1,1)	-4299.248	-3305.249
GARCH(1,1) with $\sigma^2_{imp,t-1}$	-4250.593	-3271.851
$\sigma^2_{imp,t-1}$ only	-4263.357	-3274.846

In-sample Log-likelihoods after ML estimation for the GARCH(1,1) model without and with the lagged implied volatility, and with the lagged implied volatility only (i.e. no GARCH effects in this third specification). The estimation period is August 1st, 1989 - June 28th, 2002 (S&P100 index) and September 3rd, 1996 - June 28th, 2002 (NASDAQ100 index).

Table XII
Out-of-sample VaR results for the S&P100 index

Lagged implied volatility only

	LQ=10%	LQ=5%	LQ=1%	RQ=10%	RQ=5%	RQ=1%
$\widehat{f}$	10.98	5.57	0.80	10.61	5.10	**0.43**
I	0.73	0.76	-	0.94	0.23	-
CC	0.19	0.33	-	0.51	0.46	-

H0: $\rho_1 = 0$; P-value = 0.192
H0: $\beta_0 = 0$ AND $\rho_1 = 0$ AND $\sigma^2 = 1$; P-value = 0.036
H0: $\eta_1 = 0$; t-stat = 3.517

RiskMetrics

	LQ=10%	LQ=5%	LQ=1%	RQ=10%	RQ=5%	RQ=1%
$\widehat{f}$	**11.90**	**6.73**	1.35	**11.25**	**6.06**	0.80
I	0.19	0.25	0.02	0.75	0.98	-
CC	0	0	0.01	0.06	0.03	-

H0: $\rho_1 = 0$; P-value = 0.032
H0: $\beta_0 = 0$ AND $\rho_1 = 0$ AND $\sigma^2 = 1$; P-value = 0
H0: $\eta_1 = 0$; t-stat = 3.297

GJR-GARCH

	LQ=10%	LQ=5%	LQ=1%	RQ=10%	RQ=5%	RQ=1%
$\widehat{f}$	10.61	5.20	0.89	10.30	4.83	**0.37**
I	0.13	0.51	-	0.78	0.82	-
CC	0.17	0.71	-	0.81	0.88	-

H0: $\rho_1 = 0$; P-value = 0.037
H0: $\beta_0 = 0$ AND $\rho_1 = 0$ AND $\sigma^2 = 1$; P-value = 0.056
H0: $\eta_1 = 0$; t-stat = 3.580

Summary results for the back-testing of the out-of-sample VaR models on the S&P100 index, August 1st, 1989 - June 28th, 2002 time period. The first volatility specification (top panel) is $g_t = \omega + \eta\,\sigma^2_{imp,t-1}$, the second volatility specification (middle panel) is $g_t = 0.06 r^2_{t-1} + 0.94 g_{t-1}$ while the third volatility specification (bottom panel) is $g_t = \omega + \alpha_1 r^2_{t-1} + \alpha_n r^2_{t-1} d_{t-1} + \delta_1 g_{t-1}$, all three with a skewed Student density distribution for the error term. We first give the empirical failure rates for the left (LQ) and right (RQ) quantiles at 10, 5 and 1%. A bold figure indicates that the empirical failure rate is significantly different (LR test or unconditional coverage) from the theoretical value. Secondly, *I* and *CC* give the P-values for the independence and conditional coverage tests respectively. Thirdly, the statistical tests (density forecast framework) $H0 : \ldots$ are for $z_t = \beta_0 + \rho_1 z_{t-1} + \eta_1 r^2_{t-1} + u_t$, $u_t \sim N(0, \sigma^2)$, where z_t is equal to $\Phi^{-1}(\int_{-\infty}^{r_t/\sqrt{g_t}} f(u)\,du)$.

Table XIII
Out-of-sample VaR results for the NASDAQ100 index

Lagged implied volatility only

	LQ=10%	LQ=5%	LQ=1%	RQ=10%	RQ=5%	RQ=1%
$\widehat{f}$	**12.22**	4.98	0.55	9.83	5.53	1.23
I	0.47	0.33	-	0.81	0.24	-
CC	0.02	0.62	-	0.95	0.33	-

H0: $\rho_1 = 0$; P-value = 0.519
H0: $\beta_0 = 0$ AND $\rho_1 = 0$ AND $\sigma^2 = 1$; P-value = 0
H0: $\eta_1 = 0$; t-stat = 1.001

RiskMetrics

	LQ=10%	LQ=5%	LQ=1%	RQ=10%	RQ=5%	RQ=1%
$\widehat{f}$	**12.90**	5.94	0.68	10.44	5.46	1.09
I	0.30	0.57	-	0.57	0.85	-
CC	0	0.23	-	0.72	0.71	-

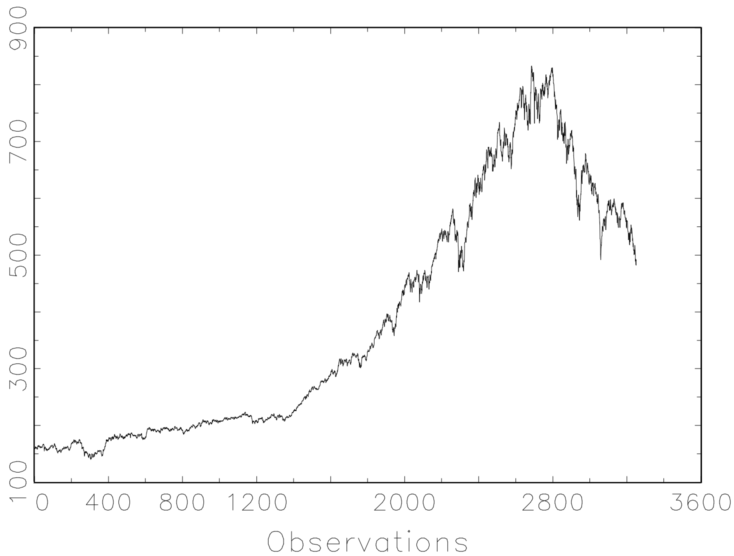
H0: $\rho_1 = 0$; P-value = 1
H0: $\beta_0 = 0$ AND $\rho_1 = 0$ AND $\sigma^2 = 1$; P-value = 0
H0: $\eta_1 = 0$; t-stat = 0.912

GJR-GARCH

	LQ=10%	LQ=5%	LQ=1%	RQ=10%	RQ=5%	RQ=1%
$\widehat{f}$	**12.76**	6.08	1.02	**12.22**	**6.96**	**1.64**
I	0.83	0.79	-	0.33	0.72	-
CC	0	0.18	-	0.01	0	-

H0: $\rho_1 = 0$; P-value = 0.742
H0: $\beta_0 = 0$ AND $\rho_1 = 0$ AND $\sigma^2 = 1$; P-value = 0
H0: $\eta_1 = 0$; t-stat = 0.641

Summary results for the back-testing of the out-of-sample VaR models on the NASDAQ100 index, September 3rd, 1996 - June 28th, 2002 time period. The first volatility specification (top panel) is $g_t = \omega + \eta\,\sigma^2_{imp,t-1}$, the second volatility specification (middle panel) is $g_t = 0.06r^2_{t-1} + 0.94g_{t-1}$ while the third volatility specification (bottom panel) is $g_t = \omega + \alpha_1 r^2_{t-1} + \alpha_n r^2_{t-1} d_{t-1} + \delta_1 g_{t-1}$, all three with a skewed Student density distribution for the error term. We first give the empirical failure rates for the left (LQ) and right (RQ) quantiles at 10, 5 and 1%. A bold figure indicates that the empirical failure rate is significantly different (LR test or unconditional coverage) from the theoretical value. Secondly, *I* and *CC* give the P-values for the independence and conditional coverage tests respectively. Thirdly, the statistical tests (density forecast framework) $H0 : \ldots$ are for $z_t = \beta_0 + \rho_1 z_{t-1} + \eta_1 r^2_{t-1} + u_t$, $u_t \sim N(0,\sigma^2)$, where z_t is equal to $\Phi^{-1}(\int_{-\infty}^{r_t/\sqrt{g_t}} f(u)\,du)$.



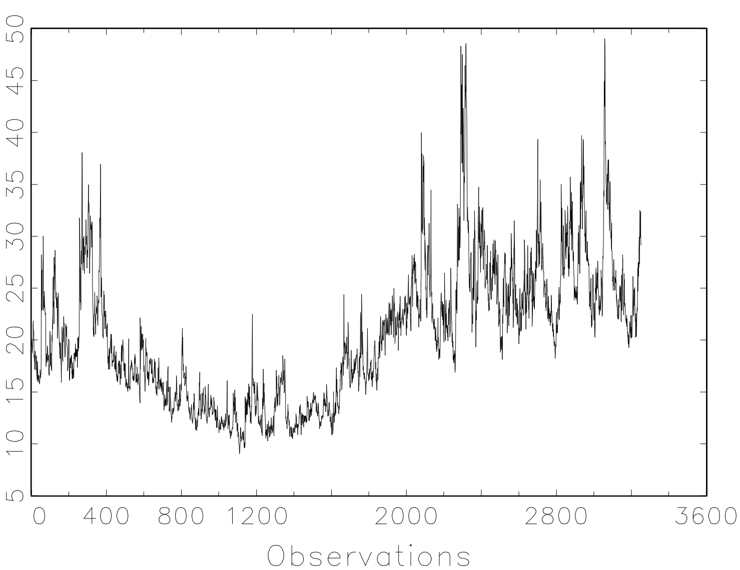


Figure 1. S&P100 index and VIX index. This figure shows the level of the S&P100 index (top of figure) and implied volatility index VIX (bottom of figure) over the August 1st, 1989 - June 28th, 2002 time period.

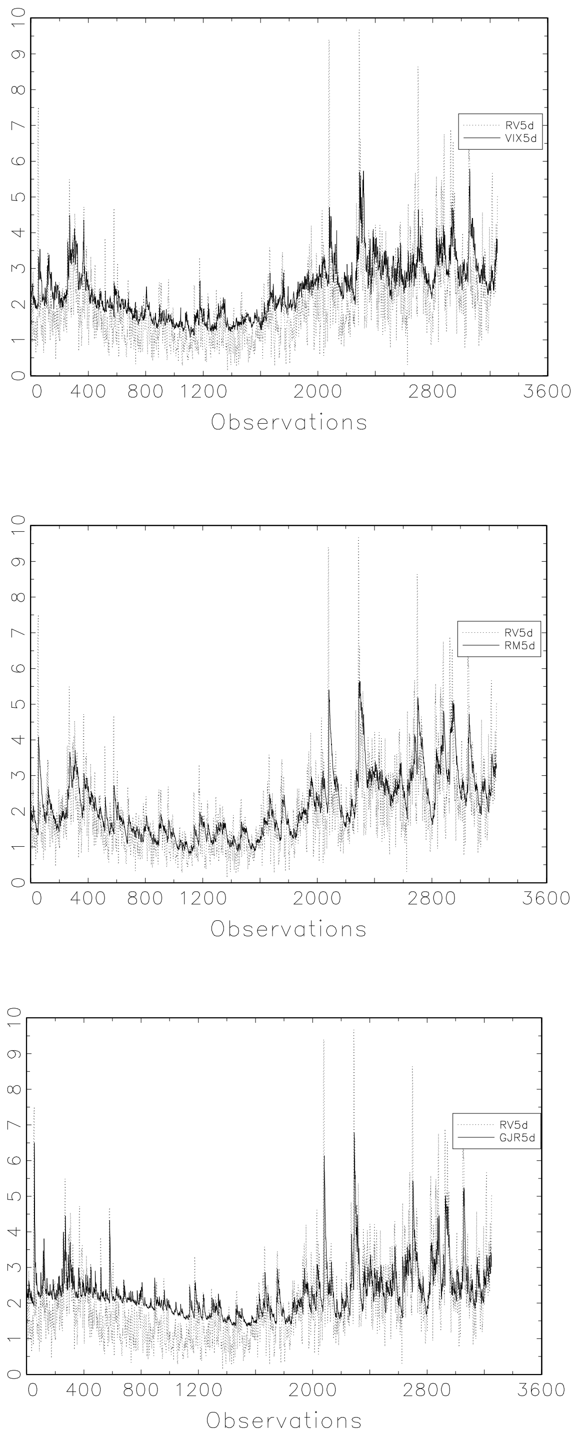


Figure 2. Realized volatility and volatility forecasts at the 5-day time horizon (S&P100 index). This figure shows the 5-day realized volatility for the S&P100 index and the 5-day volatility forecasts based on the implied volatility index (VIX5d), RiskMetrics (RM5d) and GJR-GARCH models (GJR5d), August 1st, 1989 - June 28th, 2002 time period.

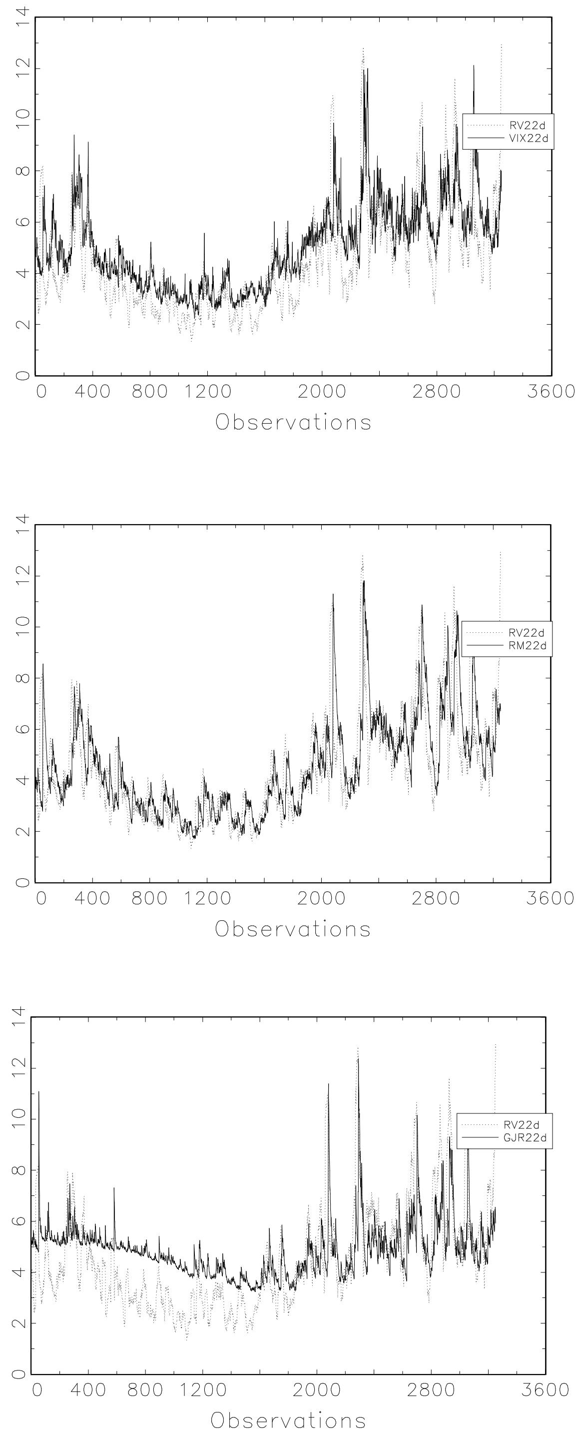
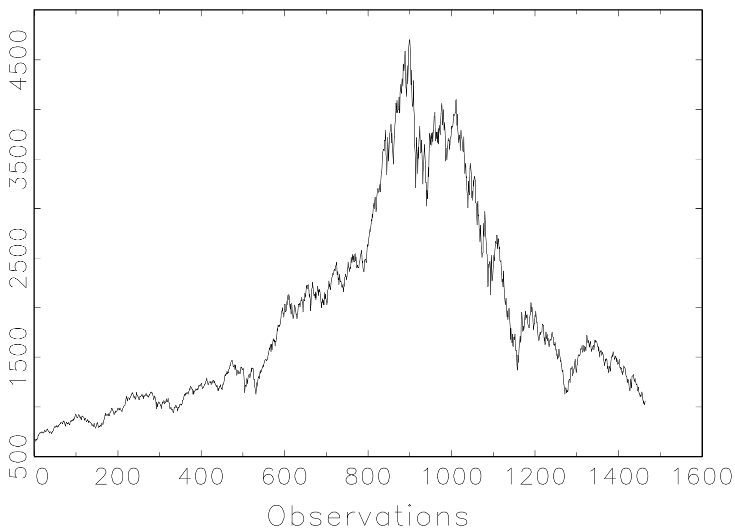


Figure 3. Realized volatility and volatility forecasts at the 22-day time horizon (S&P100 index). This figure shows the 22-day realized volatility for the S&P100 index and the 22-day volatility forecasts based on the implied volatility index (VIX22d), RiskMetrics (RM22d) and GJR-GARCH models (GJR22d), August 1st, 1989 - June 28th, 2002 time period.



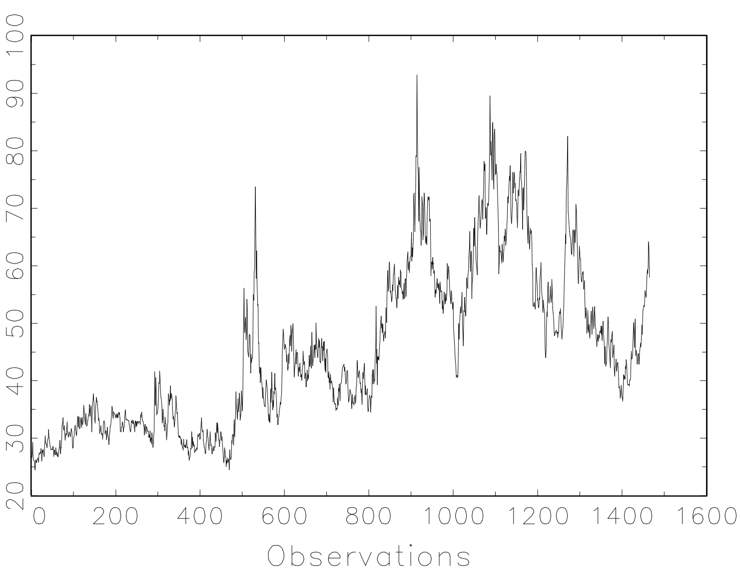


Figure 4. NASDAQ100 index and VXN index. This figure shows the level of the NASDAQ100 index (top of figure) and implied volatility index VXN (bottom of figure) over the September 3rd, 1996 - June 5th, 2002 time period.

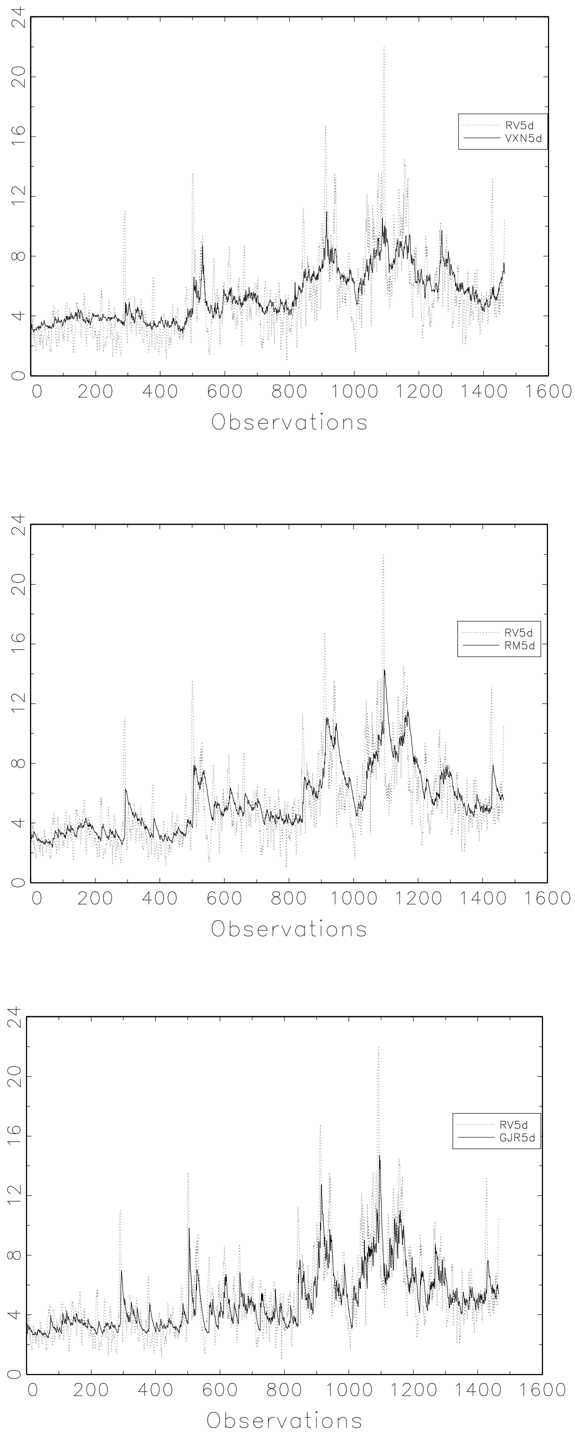


Figure 5. Realized volatility and volatility forecasts at the 5-day time horizon (NASDAQ100 index). This figure shows the 5-day realized volatility for the NASDAQ100 index and the 5-day volatility forecasts based on the implied volatility index (VXN5d), RiskMetrics (RM5d) and GJR-GARCH models (GJR5d), September 3rd, 1996 - June 5th, 2002 time period.

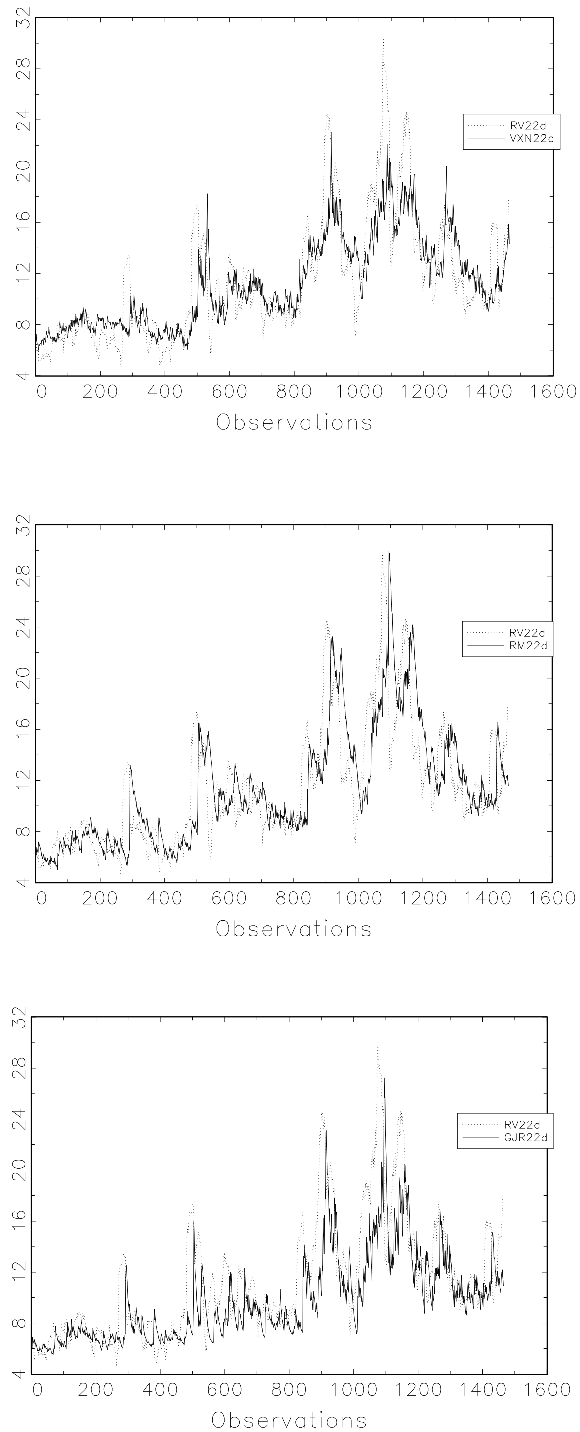


Figure 6. Realized volatility and volatility forecasts at the 22-day time horizon (NASDAQ100 index). This figure shows the 22-day realized volatility for the NASDAQ100 index and the 22-day volatility forecasts based on the implied volatility index (VXN22d), RiskMetrics (RM22d) and GJR-GARCH models (GJR22d), September 3rd, 1996 - June 28th, 2002 time period.

Notes

[1]See Poon and Granger (2003) for a recent and thorough review.

[2]See Alexander (2001) for a recent review and discussion of these models.

[3]See the CBOE website at http://www.cboe.com.

[4]The RiskMetrics specification for modelling volatility was put forward by J.P. Morgan in 1994 as a straightforward measure of volatility. It is widely used in the RiskMetrics specification of Value-at-Risk models. See Alexander (2001) for further information.

[5]In practice, daily returns can exhibit some weak autocorrelation. This usually involves fitting an AR, MA or ARMA structure on the r_t prior to the ARCH-type analysis, along with a constant. The rest of the paper thus proceeds with r_t, which are the demeaned and uncorrelated returns (or residuals from the ARMA analysis).

[6]Standard practice suggests drawing the ε_t from the normal distribution or the $t(0,1,\nu)$ distribution (to allow for fat tails in the distribution of returns). We choose the $t(0,1,\nu)$ distribution for the standardized residuals in the first part of this paper.

[7]We however use t (instead of $t+h$) in $RV_{h,t}$ as our goal (see below) is to compare the volatility forecasts made at time t with the observed realized volatility over the $[t+1, t+h]$ time period.

[8]See Alexander (2001) for a discussion of the estimation of GARCH-type models.

[9]See also Giot (2002).

[10]See also Pagan and Schwert (1990) or Christensen and Prabhala (1998).

[11]Of course, index returns display a large amount of heteroscedasticity, but we deal with volatility-based variables in our regressions and not index returns. Nevertheless we also per-

formed the tests of unbiasedness and efficiency using a Wald test (instead of the Fisher test) and we got the same results.

[12]Note however that heteroscedasticity tests on the residuals of the previous regressions, i.e. on the global August 1st, 1989 -June 28th, 2002 time period, indicate that the null hypothesis of no-heteroscedasticity is almost never rejected. Heteroscedasticity tests on the residuals of the separate regressions indicate that the null hypothesis of no-heteroscedasticity is never rejected.

[13]Combined with a GARCH-type volatility specification, Giot and Laurent (2002) show that the resulting VaR model provides excellent out-of-sample performance when back-tested against other models.

[14]The asymmetry coefficient $\xi > 0$ is defined such that the ratio of probability masses above and below the mean is $\frac{\Pr(\varepsilon \geq 0|\xi)}{\Pr(\varepsilon < 0|\xi)} = \xi^2$.

[15]$\Phi^{-1}(.)$ is the inverse of the standard normal distribution function.

[16]Once again, the test cannot be used when $\alpha = 1\%$ because $n11 = 0$.

[17]Histograms for the z_t (not reported in this paper) tell the same story.