Volatility Forecasting: Downside Risk, Jumps and Leverage Effect

Francesco Audrino and Yujia Hu

Volatility Forecasting: Downside Risk, Jumps and Leverage Effect[1]

Francesco Audrino and Yujia Hu

Author's address: Prof. Dr. Francesco Audrino
                       Institute of Mathematics and Statistics
                       Bodanstrasse 6
                       9000 St. Gallen
                       Phone   +41 71 224 2431
                       Fax   +41 71 224 2894
                       Email   francesco.audrino@unisg.ch
                       Website www.mathstat.unisg.ch

                       Yujia Hu
                       Institute of Mathematics and Statistics
                       Bodanstrasse 6
                       9000 St. Gallen
                       Phone   +41 71 224 2475
                       Fax   +41 71 224 2894
                       Email   yujia.hu@unisg.ch
                       Website www.mathstat.unisg.ch

Abstract

We provide new empirical evidence on volatility forecasting in relation to asymmetries present in the dynamics of both return and volatility processes. Leverage and volatility feedback effects among continuous and jump components of the S&P500 price and volatility dynamics are examined using recently developed methodologies to detect jumps and to disentangle their size from continuous return and continuous volatility. Granted that jumps in both return and volatility are important components for generating the two effects, we find jumps in return can improve forecasts of volatility, while jumps in volatility improve volatility forecasts to a lesser extent. Moreover, disentangling jump and continuous variations into signed semivariances further improve the out-of-sample performance of volatility forecasting models, with negative jump semivariance being highly more informative then positive jump semivariance. The model proposed is able to capture many empirical stylized facts while still remaining parsimonious in terms of number of parameters to be estimated.

Keywords

JEL Classification

1 Introduction

Volatility forecasts are crucial for many investment decisions. They are relevant for option pricing, asset allocation, and risk management. Accurate estimates of volatility are essential prerequisites for good forecasts. In light of this, the development of high frequency estimators, based on the notion of increasing sampling frequency, has put the research on this field a step forward. In contrast with model-based estimates, the realized volatility, advocated by Andersen et al. (2001a) and Barndorff-Nielsen and Shephard (2002), among others, consistently estimates the integrated volatility of the return process under the assumption that it follows a continuous sample path. Thus, realized measures represent the foundation of any forecast of volatility. Supporting this, Hansen and Lunde (2006a) suggest that model comparisons be based on the use of realized volatility as a proxy for the latent volatility, and the results in Andersen et al. (2003) indicate that simple autoregressive forecasting models based on realized volatility outperform GARCH related models in an out-of-sample perspective.

Many stylized facts are known for return and volatility dynamics. Those include volatility persistence, the volatility leverage and feedback effects, and jumps that induce skewness and leptokurtosis on the return and volatility distributions. The existence of these effects poses challenges for volatility forecasts by requiring models that can account for those empirical features. In this regard, forecasting models based on non-parametric estimates of volatility are particularly suited as jump information can be extracted directly from the data and volatility persistence, leverage and feedback effects can be modeled parsimoniously. Capturing volatility persistence, the popular heterogeneous autoregressive realized volatility model (HAR-RV) of Corsi (2009), which is based on an approximation of long run volatility components, performs surprisingly well in out-of-sample forecasts.

This paper sheds light on those stylized facts that are useful for volatility forecasting. We consider volatility leverage effects of continuous and jump components of the return/volatility dynamics as key ingredients of the forecasting models we propose. Particularly relevant for the analysis of asymmetries between jump components is the methodology to detect intraday return jumps proposed by Lee and Mykland (2008). Moreover, in the spirit of Corsi (2009), we model the long memory of volatility with HAR components.

A discrepancy exists in the interpretation of the leverage and feedback effects. In option pricing and continuous time models, the leverage and volatility feedback effects materialize in the same way even when the causes differ. They are both commonly interpreted as a negative contemporaneous

correlation between change in log-price and volatility (see Mykland and Zhang, 2009; Bandi and Reno, 2010). In the discrete time literature, a fundamental distinction exists between the two effects and this is inherent to the timing of the correlation. The leverage effect arises when there is negative correlation between volatility and lagged returns that is originally motivated by the effect of the financial leverage. A negative shock in the stock price leads the financial leverage to increase and, in turn, the volatility as well, as it is considered an increasing function of the leverage (see Christie, 1982; Schwert, 1989). In contrast, the volatility feedback effect materializes in the opposite causal direction. The common motivation is that an increase in volatility is associated with an expectation of higher future volatility and, therefore, market participants discount this information, resulting in an immediate drop in stock prices (see French et al., 1987; Campbell and Hentschel, 1992; Bekaert and Wu, 2000; Wu, 2001; Bae et al., 2007, among others). In discrete time models for volatility forecasting, this effect is of minor importance as, ultimately, it has to do with returns. However, persistency in volatility is a channel through which this effect manifests itself, and therefore the discrete time model proposed accounts not only for the leverage effect but also for the volatility feedback effect.

Evidence of the leverage effect with high frequency data has been advanced by Bollerslev et al. (2006) who analyzed sign asymmetries of high frequency returns and their impact on future volatility. This paper presents new evidence of leverage and volatility feedback effects by using the S&P500 index. To the prior literature, it adds consideration of a separate leverage effect for the jump and continuous components of the return and volatility processes. These are analyzed by means of cross-correlation between realized variation and return, with both realized variation and return being disentangled into continuous and jump components. We find that the leverage effect originates from both continuous and jump components. However, the dynamics differ as continuous component correlations tend to persist for a prolonged period, while correlation of jump components is short-lived.

The main contribution of this paper is the finding of a superior forecasting performance when the jumps in returns are isolated from the corresponding continuous component. While the previous papers separated jump variation and continuous variation, as proposed by Andersen et al. (2007a), here we disentangle the jump size from the continuous return. With this approach, any effect of return jumps on future volatility has a clear interpretation as a jump leverage effect. Consistent with Corsi and Renò (2010), the model proposed sheds light on the leverage effect also by taking lagged negative returns. Further, in agreement with Patton and Sheppard (2011), we find that the best forecasting model includes "downside risks," which are volatilities generated by negative intraday returns. The realized variation

is in fact also disentangled into continuous signed semivariations and jump signed semivariations, as a way to capture separate dynamics of negative intraday returns with respect to positive intraday returns. The leverage effect in forecasting realized volatility has been considered by both Corsi and Renò (2010) and Patton and Sheppard (2011) in a similar framework. These papers are the most closely related to our study. However, the methodology to assess it differs as we disentangle jumps size from both return and volatility, and we also separate both jump variation and continuous variation into semivariances.

The remainder of the paper is organized as follows. Section 2 presents realized estimators used in the analysis, the continuous time framework on which the estimates are based, and the methodology to disentangle continuous and jump components from both return and volatility processes. Section 3 presents the data used in the empirical analysis and the candidate forecasting models. Section 4 contains evidences for the leverage and volatility feedback effects and presents estimation results and the out-of-sample forecasting performance evaluation. Finally, Section 5 concludes.

2 Theoretical Framework

The underlying framework for the empirical analysis is based on a double jump-diffusion data generating process. This stochastic process features a continuous sample path component and occasional jumps in both return and volatility dynamics. The framework was first laid down by Duffie et al. (2000). The empirical analysis of this class of models can be found in Broadie et al. (2007), Chernov et al. (2003), Eraker et al. (2003), Eraker (2004), and Todorov and Tauchen (2010). Generally, the results support jumps in volatility as well as jumps in returns for speculative prices. Let s_t denote the logarithmic asset price at time t, for $t \in [0, T]$. In stochastic differential equation form, the price and volatility processes are

$$ds_t = m_t dt + \sigma_t dW_t^s + \xi_t^s dJ_t^s, \tag{1}$$

$$d\sigma_t^2 = \kappa \left(\theta - \sigma_t^2 \right) dt + \sigma_t^v dW_t^v + \xi_t^v dJ_t^v, \tag{2}$$

where $\{W_t^s, W_t^v\}$ is a bivariate standard Brownian motion, $\{dJ_t^s, dJ_t^v\}$ is a bivariate count process and $\{\xi_t^s, \xi_t^v\}$ represents the size of the jumps in return and in volatility if a count occurs at time t, with ξ_t^v

5

restricted to be non-negative[1]. The mean process of the squared volatility equation (2) is characterized by a long run level parameter θ and a mean-reverting parameter κ. Moreover the instantaneous volatility of the squared volatility σ_t^v is allowed to be time-varying.

This framework allows the generation of the contemporaneous leverage or volatility feedback effects through the correlation between the continuous components as well as through the correlation between the jump components that may be both in time and in size of the jumps. However, with the discrete time analysis this paper proposes, leverage and volatility feedback effects are interpreted as lag correlation and therefore the assumption of stochastic dependence is not needed.

Following the theory of quadratic variation, the volatility of the price process is estimated with the realized volatility from high frequency data. Below, we present the estimators for volatility and jump components used in the analysis.

2.1 Return Volatility and Jumps

Let $r_{t,i}$ be the discretely sampled ith intraday return for day t. In the presence of jumps in return, the realized variation, $RV_t = \sum_{i=1}^{m} r_{t,i}^2$, introduced by Andersen et al. (2001a,b), captures both continuous and jump components of the quadratic variation:

$$RV_t \xrightarrow{p} \int_{t-1}^{t} \sigma_\tau^2 d\tau + \sum_{t-1 < \tau \leq t} (\xi_\tau^s)^2 \quad t = 1, \dots, T. \tag{3}$$

The bipower variation, introduced by Barndorff-Nielsen and Shephard (2004), is instead a consistent estimator for the continuous component only:

$$BV_t = \mu^{-2} \sum_{i=2}^{m} |r_{t,i}| |r_{t,i-1}| \xrightarrow{m \to \infty} \int_{t-1}^{t} \sigma_\tau^2 d\tau \tag{4}$$

where $\mu = 2^{1/2} \frac{\Gamma(1)}{\Gamma(1/2)} = \sqrt{2/\pi}$, with $\Gamma(\cdot)$ denoting the gamma function.

Andersen et al. (2007b) and Lee and Mykland (2008) propose detecting intraday jumps using the ratio between intraday returns and estimated spot volatility. We follow the methodology laid down by Lee and Mykland (2008) and, in particular, we test the presence of intraday jumps with the statistics

$$\mathcal{L}_{t,i} = \frac{r_{t,i}}{\frac{1}{K-1} \sum_{k=1}^{K-1} |r_{t,i-k}| |r_{t,i-k-1}|}, \tag{5}$$

where the expression contains in the denominator the estimates of the spot volatility as the average

---

[1]It is common in the option pricing literature to assume a compound Poisson process, where jump arrivals have a Poisson distribution with intensity $\{\lambda^s, \lambda^\sigma\}$ and jump sizes in return have a normal distribution. In the present context these assumptions are not needed.

bipower variation over a period with K observations. Lee and Mykland (2008) suggest using $K = 78, 110, 156, 270$, respectively with return sampled at frequencies of 60, 30, 15, and 5 minutes. Under the null of no intraday jump, the test statistics $\mathscr{L}_{t,i}$ follow a normal distribution (with variance μ).

In order to select the rejection region for the test statistics $\mathscr{L}_{t,i}$, Lee and Mykland (2008) propose to look at the asymptotic distribution of maxima of the test statistics. As the sampling frequency tends to zero, under the null of no jumps between time $t, i-1$ and t, i, the absolute value of $\mathscr{L}_{t,i}$ converges to a Gumbel distribution:

$$\frac{\max_{(t,j)} |\mathscr{L}_{t,i}| - C_n}{S_n} \xrightarrow{d} \zeta, \tag{6}$$

where ζ has a standard Gumbel distribution, $C_n = \frac{(2 \log n)^{1/2}}{\mu} - \frac{\log \pi + \log(\log n)}{2\mu(2 \log n)^{1/2}}$, $S_n = \frac{1}{\mu(2 \log n)^{1/2}}$, and n is the number of observations for each period t. We reject the null of no jump at time t, i if

$$\frac{|\mathscr{L}_{t,i}| - C_n}{S_n} > \beta^*, \tag{7}$$

such that $\exp\left(-e^{-\beta^*}\right) = 1 - \alpha$, i.e. $\beta^* = -\log\left(-\log\left(1 - \alpha\right)\right)$, with α being the significance level[2].

The test is able to detect the jump arrival time i_j for each day t, where j denotes the presence of a jump. Moreover the jump size is computed as

$$\kappa_{t,i_j} = (r_{t,i}) \, \mathbb{1}_{\left\{\frac{|\mathscr{L}_{t,i}| - C_n}{S_n} > \beta^*\right\}}. \tag{8}$$

Consequently, the jump size for the day t is

$$jRet_t = \sum_{i_j = i_1, \dots, i_{J_t}} \kappa_{t,i_j}, \quad t = 1, \dots, T, \tag{9}$$

where J_t is the total number of significant jumps for day t, and the jump-adjusted daily return is

$$cRet_t = r_t - jRet_t, \quad t = 1, \dots, T. \tag{10}$$

With this methodology to identify intraday jumps, it is possible to directly estimate the jump variation, i.e. the quadratic variation of return jumps. We follow Andersen et al. (2010) and estimate the quadratic variation due to the continuous and the jump components respectively as

[2]The extant literature has proposed other compelling tests for the presence of jumps. Alternative tests are mainly based on the difference between RV_t and BV_t and follow the asymptotic distribution theory of Barndorff-Nielsen and Shephard (2006). In a robustness check, we obtain the same empirical results regarding forecasting performance of our models with both tests for significant jumps of Barndorff-Nielsen and Shephard (2006) and Corsi et al. (2010), combined with the recursive methodology of Andersen et al. (2010) to identify the size of each intraday jump.

$$CV_t \quad = \quad RV_t - JV_t, \quad t = 1, \ldots, T, \tag{11}$$

$$JV_t \quad = \quad \sum_{j=1}^{J_t} JV_{t,j}, \quad t = 1, \ldots, T, \tag{12}$$

where

$$JV_{t,j} = \mathbb{1}_{\{\kappa_{t,i} \neq 0\}} \left(\kappa_{t,i_j}^2 - \frac{1}{m - J_t} \sum_{k \in \{1, \ldots m\} \setminus \{i_j, \ldots, i_{J_t}\}} r_{t,k}^2 \right), \quad j = 1, \ldots, J_t, \quad t = 1, \ldots, T, \tag{13}$$

is the contribution to the quadratic variation of each intraday jump κ_{t,i_j}. We find that the daily jump variation identified with this methodology is highly correlated with more traditional methods to identify jump variation that primarily takes the positive differences between RV_t and BV_t.

2.2 Downside Continuous and Jump Variation

To capture the sign asymmetry of the volatility process, the continuous variation and jump variation are decomposed using signed intraday returns. Barndorff-Nielsen et al. (2010) introduced a new estimator that captures the quadratic variation due to signed returns, termed realized semivariance. In a similar way, the continuous variation and jumps variation can both be decomposed into signed semivariations by using the test of Lee and Mykland (2008). This represents the main advantage of the test, as it is able to identify the sign and the timing of intraday jumps. If jump variation is instead identified by means of significant $(RV_t - BV_t)$, as normally done in the previous literature, no-information is available about the sign or timing of intraday jumps and therefore we are not able to disentangle jump or continuous semivariations.

The realized semivariances of Barndorff-Nielsen et al. (2010) are defined as

$$RS_t^+ \quad = \quad \sum_{i=1}^{m} r_{t,i}^2 \mathbb{1}_{\{r_{t,i} > 0\}} \tag{14}$$

$$RS_t^- \quad = \quad \sum_{i=1}^{m} r_{t,i}^2 \mathbb{1}_{\{r_{t,i} < 0\}}. \tag{15}$$

Similarly, the jump semivariation is generated by signed intraday jumps and it is defined as follows:

$$JSV_t^+ \quad = \quad \sum_{j=1}^{J_t} JSV_{t,j}^+, \tag{16}$$

$$JSV_t^- \quad = \quad \sum_{j=1}^{J_t} JSV_{t,j}^-, \tag{17}$$

where

$$JSV_{t,j}^{+} = \mathbb{1}_{\{\kappa_{t,i}>0\}} \left(\kappa_{t,i_j}^2 - \frac{1}{m-J_t} \sum_{k\in\{1,...m\}\backslash\{i_j,...,i_{jt}\}} r_{t,k}^2 \right), \tag{18}$$

$$JSV_{t,j}^{-} = \mathbb{1}_{\{\kappa_{t,i}<0\}} \left(\kappa_{t,i_j}^2 - \frac{1}{m-J_t} \sum_{k\in\{1,...m\}\backslash\{i_j,...,i_{jt}\}} r_{t,k}^2 \right), \tag{19}$$

for $j=1,\ldots,J_t$ and $t=1,\ldots,T$. Consequently, the continuous semivariation is given by

$$CSV_t^{+} = RS_t^{+} - JSV_t^{+}, \tag{20}$$

$$CSV_t^{-} = RS_t^{-} - JSV_t^{-}. \tag{21}$$

The decomposition in semivariances is complete as $RV_t = RS_t^{+} + RS_t^{-}$, $JV_t = JSV_t^{+} + JSV_t^{-}$ and $CV_t = CSV_t^{+} + CSV_t^{-}$. To study the volatility feedback effect, a particular focus is on the downside continuous variation and jump variation captured by CSV^{-} and JSV^{-}, respectively, as Patton and Sheppard (2011) show that negative semivariances are more informative then positive semivariances for forecasting future volatility. These quantities are used in the empirical analysis of the following sections.

2.3 Volatility Jumps

The estimation of volatility jumps is a prominent topic of research (see Todorov and Tauchen, 2010). Parametric methods can be employed to estimate the jump intensity parameter and the average jump size for the process (2) given a distributional assumption on the jump process. However, we consider an alternative methodology to estimate the size of (continuous) volatility jumps without imposing assumptions on its distribution. As is standard practice in the realized volatility literature, the volatility estimate (continuous variation, eq. 11) is taken as observed.

We consider an auxiliary AR(1) model on the level of continuous variation with GARCH(1,1) dynamics. This choice is motivated empirically and theoretically. The AR-GARCH(1,1) on the changes in continuous volatility is a natural discrete approximation of the volatility process in (2) without the jump component, and, moreover, it fits very well to the data. It is specified as

$$(\Delta CV)_t = c + \phi \cdot CV_{t-1} + \sigma_t^v \cdot e_t, \tag{22}$$

$$e_t \sim N(0,1),$$

$$(\sigma_t^v)^2 = \omega + \alpha \cdot e_{t-1}^2 + \beta \cdot (\sigma_{t-1}^v)^2, \tag{23}$$

with stationarity constraint $|\phi| < 1$, $(\alpha + \beta) < 1$, and non-negativity constraints on $(\sigma_t^\nu)^2$: $\omega > 0, \{\alpha, \beta\} \geq 0$. The jump in volatility is estimated as

$$VolJ_t = \mathbb{1}_{\{e_t > \Phi_{1-\alpha}\}} \cdot e_t \tag{24}$$

Intuitively, in the absence of volatility jumps, the auxiliary model approximates the true volatility dynamics. However, in the presence of large jumps, the model is not able to fit the data and large residuals from the model fit represent contributions of volatility jumps. The jump is in the continuous variation; this combination of terms sounds somehow antithetical. The continuous variation captures the volatility without jumps in returns. However, it may well be that it presents jumps itself, whether or not there are jumps in returns. Jumps in continuous variation are therefore labeled volatility jumps. The continuous part of the continuous variation is labeled adjusted continuous variation and it is given by

$$adjCV_t = CV_t - VolJ_t. \tag{25}$$

To maintain positivity of the adjusted continuous variation, a restriction is made on the volatility jumps. It may happen that the estimated volatility jumps, $VolJ_t$, are higher than the total continuous variation CV_t. In such cases, the volatility jumps are set equal to the continuous variation such that $adjCV_t$ equals zero. The interpretation of this is that the volatility on certain specific days is entirely driven by jumps.

3   Data and Forecasting Models

3.1   Data and Summary Statistics

The data used in the analysis consists of tick-by-tick prices for the Standard & Poor (S&P500) Future. The sample period covers almost the entire history of this security: from April 28, 1982 to August 6, 2010. The S&P500 Future is traded "open outcry" on regular market opening hours of the Chicago Mercantile Exchange (from 8:30 AM to 15:15) and it has also been used by Bollerslev et al. (2009) and Corsi (2009), among others, as an index for the composite market.

The data cleaning procedure follows Hansen and Lunde (2006b). In particular, as trades before 9:00 AM for the first several years from the introduction of this security are often not very active, all the trades outside 9:00 AM and 15:15 are discarded. Transactions with zero volume are also discarded. Moreover, days with less than five consecutive trading hours are also discarded. Prices are then sampled

every five minutes, starting from 9:00 AM, in order to smooth the impact of market microstructure noise. Furthermore, whenever there is no observed price attached to a specific time stamp, the "previous tick" method is used to replace the missing price.

[Figure 1 about here.]

Figure 1 represents the time series of returns, log realized variation and jumps in return and volatility. The highest levels of realized volatility were reached on "black Monday", October 19, 1987, and the following day October 20, where returns were respectively about -26% and -9.6%. The test (7) detects jumps during both days. However for October 19, 1987, the intraday price changes were less drastic than on the following day, resulting in a lower jump variation. In October 20, 1987, jump variation was at the highest level. During that day negative jumps were combined with positive jumps, while on October 19, sequences of negative intraday returns dominated. The jumps test are carried out through the analysis with $\alpha = 0.01$ significance level.

Jumps in volatility are also identified for both October 19 and 20, 1987, with the volatility jump for October 19 being at the highest level. Accordingly, the market crash of October 19, 1987 is explained with jumps in returns and even more with a jump in volatility. This underscores the importance of allowing for volatility jumps in order to fully capture the dynamics of extreme events. Return jumps alone, identified with the testing methodology, are not able to fully account for extreme market events. Summary statistics for the intensity of jumps in return and in volatility are reported in Table 1.

[Table 1 about here.]

On average, there is less than one return jump per day. With significance level $\alpha = 0.01$ and five minutes sampling, the number of days with at least one intraday jump represents 19.29% of the sample period. The intraday jump intensity ranges from 0 to 10 jumps (maximum corresponding to October 19, 1987). Conditional on the presence of jumps there are on average 1.43 intraday return jumps each day. Regarding jumps in volatility, we find an intensity of about 6.65% of the trading days. The analysis for volatility jumps is also carried out at significance level $\alpha = 0.01$ in eq. (24). As with jumps in return, other significance levels are experimented. The chosen cutoff level $\alpha = 0.01$ is empirically motivated by the Quantile-Quantile Plot showed in Figure 2.

[Figure 2 about here.]

The empirical distribution with this cutoff level best approximates that of a standard normal distribution. Without a threshold, the residuals from the auxiliary GARCH model fitting deviate substantially from the assumed standard normal distribution. Only large departures from the standard normal distribution are categorized as volatility jumps.

The contribution of return jumps to total return levels, the relative contribution of jump variation to the total variation and the contribution of volatility jumps to the continuous volatility are reported in Figure 3.

[Figure 3 about here.]

The contribution of negative jumps to negative returns appears to be higher, on average, than the contribution of positive jumps to the corresponding signed return. This jump sign asymmetry may account for the negative skewness observed in financial markets. Based on rolling averages of 3 months and 1 year windows, the contribution of the negative jumps to corresponding signed returns ranges from approximately 0.1% to 4.5%, on 3 months basis, and from 1% to 3.5%, on a yearly basis, while that of positive jumps ranges from 1% to 2.5%, on a yearly basis. Jump variation accounts for approximately 3% to 7% on a yearly basis. This statistic is in line with the ones reported by the previous literature despite the use of a different methodology to test for the presence of jumps. Corsi and Renò (2010), for example, report an average percentage over almost 28 years of about 6% for the S&P500 index, and Andersen et al. (2010) report an average percentage over five years that ranges from 2.1% to 5.8% for different stocks. By contrast, volatility jumps make a minor contribution to the continuous variation as they are more sporadic than return jumps. They account for about 0% to 8% of the continuous variation on a yearly basis, and for the entire sample period from 1982 to 2010 they represent 3.4% of continuous variation.

3.2 Forecasting with Jumps, Leverage effect, and Volatility Persistency

Given the measures of volatility and jump magnitude introduced previously, it is of interest to study their additional forecasting power for future volatility. The models proposed are based on the Heterogeneous Autoregressive (HAR) framework of Corsi (2009) and extend it by considering semivariances, jumps in return and in volatility, and the leverage effect due to continuous return and return jumps.

The realized volatility features long memory. The HAR model, although it does not belong formally to the class of long-memory models, represents a parsimonious approximation, which is able to closely

mimic such a stylized fact. In more detail, the benchmark model is

$$\log{(RV)}_{t,t+h} = \alpha + \phi^d \cdot \log{(RV)}_t + \phi^w \log{(RV)}_{t-5,t} + \phi^m \log{(RV)}_{t-22,t} + error_t, \qquad (26)$$

where $\log{(RV)}_{t,t+h}$ is the average log realized variation between time t and $t+h$. The variables $\log{(RV)}_{t-5,t} = \frac{1}{5}\sum_{i=t-4}^{t}\log{(RV_i)}$ and $\log{(RV)}_{t-22,t} = \frac{1}{22}\sum_{i=t-21}^{t}\log{(RV_i)}$ capture long memory features of the volatility process. The coefficients of the model may be interpreted as the reaction of heterogeneous agents who forecast with different time horizons: daily, weekly and monthly. The representation of the HAR forecasting model based on realized variation instead of log realized variation,

$$RV_{t,t+h} = \alpha + \phi^d RV_t + \phi^w RV_{(t-5,t]} + \phi^m_{(t-22,t]} + error_t, \qquad (27)$$

is also considered. However, the last produces inferior forecasts and therefore results are summarized and discussed for the log-log model.

Several model extensions based on the HAR framework have been proposed in the recent literature: see, among others, Andersen et al. (2007a), Corsi and Renò (2010), and Patton and Sheppard (2011). Andersen et al. (2007a) included jump variation in forecasting volatility dynamics. Corsi and Renò (2010) studied the impact of the leverage effect on volatility by including past signed returns. Finally, Patton and Sheppard (2011) and Chen and Ghysels (2010) considered the use of realized semivariances to forecast volatility. None of the previous studies, however, used the return jumps themselves and jumps in volatility to generalize the HAR model.

Jumps in return arguably have an impact on future volatility. Contrary to the effect of jump variations, this is interpreted as a leverage effect, which can manifest itself differently whether it comes about through continuous returns or jumps. In fact, although both components together generate the final leverage effect, their dynamics differ, as jumps have a very short impact on future volatility while continuous returns tend to have a persistent impact on future volatility.

Disentangling jump variation and continuous variation systematically improves volatility forecasts, and disentangling downside semivariation and upside semivariation also improves forecasts. We combine the ideas of Andersen et al. (2007a) and Patton and Sheppard (2011) by considering a complete decomposition into signed continuous and jump variation and this leads to even better forecasts. Negative jump semivariation has in fact a completely different impact on future volatility than positive jump semivariation. Lastly, jumps in volatility may also improve future volatility forecasts through the effect

of volatility trading strategies. Taking into account all these effects, the candidate model we propose is

$$
\begin{aligned}
\log\left(RV\right)_{t,t+h} = {} & \alpha + \beta^{+} \cdot \log\left(CSV^{+}\right)_{t} + \beta^{-} \cdot \log\left(CSV^{-}\right)_{t} + \gamma^{+} \cdot \log\left(1+JSV^{+}\right)_{t} + \gamma^{-} \cdot \log\left(1+JSV^{-}\right)_{t} \\
& + \theta \cdot \log\left(1+VolJ\right)_{t} + \delta \cdot cRet_{t}^{-} + \delta_{w} \cdot cRet_{(t-5,t]}^{-} \\
& + \vartheta_{w} \cdot \log\left(RV\right)_{(t-5,t]} + \vartheta_{m} \cdot \log\left(RV\right)_{(t-22,t]} + error_{t},
\end{aligned} \tag{28}
$$

where past negative returns and jumps are defined by $cRet_{t}^{-} = \mathbb{1}_{\{cRet_{t}<0\}} \cdot cRet_{t}$, and $cRet_{(t-5,t]}^{-} = \frac{1}{5}\sum_{i=t-4}^{t} \cdot cRet_{i}^{-}$. Note that the realized jump semivariances already capture the effect of return jumps, since they are derived from them. The interpretation of their effects is that of jump risks that capture jump leverage effects. Therefore, return jumps are not included in this model specification.

Given that one of the goals of this study is volatility forecasting, one can argue that such a model may be overparameterized and therefore may lead to poor out-of-sample results. As we will show in the next sections this is not (entirely) the case. Nevertheless, to overcome this problem we consider a simplified version of model (28) that differs from the previous one by discarding jump semivariations, the weekly persistency in the leverage effect, and the jumps in volatility[3]. Moreover the impact of jump semivariations on future volatility are being replaced by that of return jumps which is to be interpreted as a leverage effect due to jumps:

$$
\begin{aligned}
\log\left(RV\right)_{t,t+h} = {} & \alpha + \beta^{+} \cdot \log\left(CSV^{+}\right)_{t} + \beta^{-} \cdot \log\left(CSV^{-}\right)_{t} \\
& + \delta \cdot cRet_{t}^{-} + \varphi \cdot jRet_{t} + \vartheta_{w} \cdot \log\left(RV\right)_{(t-5,t]} + \vartheta_{m} \cdot \log\left(RV\right)_{(t-22,t]} + error_{t}.
\end{aligned} \tag{29}
$$

This model is very simple, consisting only of seven parameters to be estimated, e.g., three parameters less than the LHAR-CJ model proposed by Corsi and Renò (2010). Still, the model focuses on the most relevant effects for volatility forecasting: downside risk, leverage effect, and the HAR structure capturing long-memory.

The models are estimated by OLS with Newey and West (1987) covariance matrix correction to account for serial correlation. The bandwidth used is $2(h-1)$, where h is the forecasting horizon.

---

[3]We tried to discard these effects one-by-one with an approach similar to backward stepwise subset selection and verified that the results both in-sample and out-of-sample were quantitatively similar, yielding the specification (29).

4 Empirical Evidence

4.1 Leverage Effect, Volatility Feedback Effect, and Persistency

Leverage and volatility feedback effects are commonly studied by mean of correlations. Bollerslev et al. (2006) provide exhaustive evidence of the leverage effect at high-frequency by studying cross-correlations among return and volatility series for a horizon spanning several days. Exploiting the methodology we use to disentangle the continuous and jump components of return and volatility dynamics, we present new evidence of leverage and volatility feedback effects arising from both continuous and jump components. Moreover, we use signed intraday returns to capture sign asymmetries. Figure 4 reports cross-correlations among different components of the return and quadratic variation as evidence of leverage and volatility feedback effects, and persistency in volatility.

[Figure 4 about here.]

As in Bollerslev et al. (2006), the leverage effect (negative correlation between RV_{t+h} and Ret_t) is significant for prolonged days, while there is no evidence of volatility feedback effect (negative correlation between RV_{t-h} and Ret_t). The sign and size asymmetry of the leverage effect becomes evident as negative returns generate all the negative correlation of return and lagged volatility (see top right plot, figure 4) and the magnitude of this effect is higher than the positive correlation between positive returns and realized variation. Regarding volatility feedback effect, we observe a negative correlation between realized variation and future negative returns. However, this correlation is rather small and insufficient to generate the negative size asymmetry, as realized variation has an even higher magnitude of correlation with positive future returns.

Jumps in return are highly responsible for the leverage effect. The negative correlation between jumps and future volatility is very high in magnitude (see second row of figure 4) but this effect is short lived, lasting only one day period. The plots on the third and fourth rows of figure 4 reports the cross correlation of realized variation with signed continuous and jump semivariations. All those components have an impact on future volatility and they represent different sources of risk.

The persistence in volatility is also examined. It is well known that volatility is autocorrelated for a prolonged period of time. Evidence is found in Ding et al. (1993) for the S&P500 stock index. Models to capture the persistence in conditional volatility have been proposed in their early stage by Engle and Bollerslev (1986) and Baillie et al. (1996). This autocorrelation is one of the drivers of the volatility

feedback effect. As mentioned above, when an increase in volatility is associated with an expectation of higher future volatility, market participants may discount this information, resulting in an immediate drop in stock prices. We find that the persistence in volatility is generated by the continuous component of the return variation. Clearly, jump variation has a confounding effect on realized variation as well, but this effect is also short lived (The cross-correlations of realized variation with continuous variation and with jump variation are reported in the fifth row of figure 4.)

Finally, both continuous return and jumps in return are negatively and contemporaneously correlated with volatility jumps. Moreover, they have a small and short-lived impact on future volatility jumps. On the other hand, volatility jumps also have a short-lived impact only on future return jumps and continuous returns (Cross-correlations of continuous returns and return jumps with jumps in volatility, depicting asymmetries for the jump components of the volatility, are represented in the last row of Figure 4).

4.2   In-Sample Analysis

Estimation results of the proposed forecasting models are discussed in this section. The in-sample evaluations are based on different horizons. The regression results of the candidate models are reported in Table 2, for horizons of 1, 5, 15 and 22 days.

[Table 2 about here.]

In order to check the stability of the coefficients, regressions are run for different overlapping subsamples. As the full sample period is relatively long, consisting of more than 28 years, it is likely that structural breaks have occurred. This seems to be the case, as some coefficients estimated by including data for the initial 8 years, from 1982 to 1989 (the forecasts are relative to the period 1990-1997), appear to differ slightly from those estimated by using data starting from 1990. Conversely, the coefficients associated with the model for the last 20 years are relatively stable. Figure 5 reports the estimated coefficients of one period forecast of model (28) with rolling subsamples of 2000 observations (corresponding to almost 8 years).

[Figure 5 about here.]

With an in-deep inspection, the difference of the coefficient estimates for the initial subsample from those of the remaining subsample is mainly caused by the market crash of October 1987. The sample

16

period is therefore reduced and the results of the in-sample analysis contained in table 2 are relative to the subperiod from January 4, 1988 to August 6, 2010[4]. By contrast, the out-of-sample evaluations of the next sections, as they are performed using rolling forecasts, will be based on the whole sample period. Volatility jumps have a marginal power to forecast the one period ahead volatility only with the initial subsamples, which contain the market crash data, while for the remaining subsamples they do not add further value. This indicates that volatility jumps are useful for forecasting during extremely agitated periods.

To check the stability of the estimates for different forecasting horizons, figure 6 reports the estimated coefficients for forecasting horizons ranging from 1 day to 30 days. The estimated parameters appear to be well-behaved. The 95% forecast confidence intervals (adjusted with Newey-West serial correlation consistent standard errors), are relatively large for the coefficients associated with jump semivariations.

[Figure 6 about here.]

The following results from the estimation are worth remarking on. "Bad volatilities" have a significant impact on future volatility. Both continuous and jump downside semivariations increase future realized variation. However, when "good volatilities" are disentangled into upside continuous variation and upside jump variation, a striking difference emerges. The effect of upside continuous semivariation on future volatility is still positive and statistically significant, but that of the upside jump semivariation is on average negative and negligible. This may well answer the critique advanced by Corsi et al. (2010) as they argue that influential studies such as Andersen et al. (2007a) and Forsberg and Ghysels (2007), among others, find a negative or null impact of jumps (more appropriately jump variation) on future volatility, while economic theory suggests the opposite. In fact, one needs to distinguish between downside and upside jumps. Future volatility is indeed increasing with downside jump variation as jump variations are likely associated with an increase in uncertainty on fundamental values. However the effect of upside jump variation does not necessarily increase future volatility. This would be consistent with the economic model of Veronesi (1999). Intuitively, a positive jump is associated with the occurrence of good news as well as the expectation of higher future returns. The effect of the latter is able to offset the effect of an increasing uncertainty. The future volatility given the occurrence of a

---

[4]Although the estimated coefficients obtained by using the full sample data, including the market crash of 1987, do not differ substantially from the ones reported in table 2, there may be a minor loss in term of consistency of the estimates by using the full sample data. However, the signs associated with all significant coefficients are the same for the various subsamples.

positive jump is therefore lower than the one associated with the occurrence of a negative jump.

Consistent with the leverage effect, both continuous and jump components of the return have a significant impact on future volatility. As mentioned previously, the persistency in leverage effect is captured mainly by the continuous component. In fact, as the forecasting horizon increases, the leverage effect generated by past jumps becomes less statistically significant. On the contrary, the coefficient associated with the negative continuous return over the past day and over the past week remains statistically significant even for longer forecasting horizons.

Concerning long memory features, the persistence parameter associated with the one week and one month realized volatility is highly statistically significant for all horizons. Ultimately, jumps in volatility (continuous variation) do not appear to be statistically significant. Given that volatility jumps are present during extremely agitated periods, their null forecasting performance on the short horizon is due to the fact that only a few volatility jumps are identified for the sample period under investigation. There are in fact volatility jumps for only 3.6% of the sample days.

Finally, by examining the (in-sample) performance of the simplified model (29) in comparison with the one of the full model, the losses in accuracy that occur for the simplified model are negligible based on the adjusted R^2.

4.3 Out-of-Sample Forecasting Performance

The methodology used to assess the forecasting performance of the proposed models is presented in this section. The analysis is based on the out-of-sample predictive accuracy of these models in comparison to the HAR-RV model (26). The predicted variable is the cumulative average log realized variation between time t and $t + h$. For predictions of horizons $h > 1$, a "direct method" is employed, that is, the model specifies only the relation between $\log (RV)_{t,t+h}$ and the regressors at time t. The evaluation of the performance is done recursively with rolling windows. The forecasting horizons considered are $h = \{1, 5, 15, 22\}$.

The recursive procedure is applied as follows: The forecasts are generated by using an in-sample estimation window of 2000 observations, corresponding to about eight years, starting from April 28, 1982. For $h = 1$, the performance is evaluated on 5040 out-of-sample data points, corresponding to about 20 years. The forecasting performance is based on the MSE function of the log realized variation forecasts and the negative QLIKE loss function:

$$MSE = \left(\ln\widehat{(RV)}_{t,t+h|\mathscr{F}_t} - \ln(RV)_{t,t+h} \right)^2, \tag{30}$$

$$QLIKE = \ln\widehat{(RV)}_{t,t+h|\mathscr{F}_t} + \frac{RV_{t,t+h}}{\widehat{RV}_{t,t+h|\mathscr{F}_t}}. \tag{31}$$

The MSE is a symmetric loss function while QLIKE is asymmetric. Patton (2011) shows that the QLIKE loss function is robust to noise in the volatility proxy, as volatility forecasts represent a case where the true values are not observable. Moreover, this function has certain optimal properties in that it is less sensitive to large observations by more heavily penalizing under-predictions then over-predictions. It is therefore more suited to yield model rankings in the presence of imperfect proxies.

To test for the superior forecasting performance of the proposed models over the benchmark HAR-RV model, the Diebold and Mariano (1995) test is employed. The asymptotic variance of the loss differential (of each model and the HAR-RV model) is estimated with Newey-West HAC consistent sample variance as suggested by Giacomini and White (2006). Clark and McCracken (2001) show that in the presence of nested models, the distribution of the Diebold-Mariano test statistics for mean square error losses, under the null, can be nonstandard. Therefore the test of forecasting performance for mean square error losses is based on the Clark and West (2007) test statistic, which appropriately corrects the loss differential. Table 3 reports the test statistics under the null of equal forecasting performance for each model pair. A positive loss differential represents a superior average forecasting performance of the proposed model with respect to the HAR-RV.

[Table 3 about here.]

At the 99% confidence level, the proposed models outperform the base model for forecasting horizons of one day and five days. For longer forecasting horizons, overall the models also perform well, with the reduced model (eq. 29) achieving surprisingly high performance. Based on the QLIKE loss function, the outperformance of the full model (28) compared with the HAR-RV model is only weakly statistically significant. This is a warning signal that model (28) is probably overparameterized, although as shown in the next section, outperforming the benchmark HAR-RV model in forecasting long-run volatility seems to be a difficult task.

4.4 Model Confidence Set

The HAR-RV model in logarithmic form generally has a good out-of-sample forecasting performance. In an attempt to raise the bar, other reference models recently proposed in the literature are considered. First, we test the performance of the HAR model for the realized variation, which yielded the lowest forecasting power among all the candidate models when evaluated on the loss functions (30) and (31). Thus, results for that model are not reported. Then, we investigate the accuracy of the volatility forecasts for several models involving two model specifications of Andersen et al. (2007a), named HAR-RV-J and HAR-RV-CJ (eq. 13, p. 709, and eq. 28, p. 715), which explicitly take into account continuous and jump variation; two models proposed by Patton and Sheppard (2011) based on semivariances, one with complete decomposition into semivariances and the other with a downside semivariance only for the daily component (eq. 19, p. 16, and eq. 17, p. 13, without upside semivariance); and two models of Corsi and Renò (2010) that add the leverage effect, named LHAR-CJ and LHAR-CJ+ (eq. 2.4, p. 8, and table 3, p. 16). Finally, simple and commonly used models are also taken into account. These are an autoregressive model on the daily component only and an exponential smoothing model on log realized variation. The last is popularly used by risk practitioners (see Taylor, 2004) and is specified as

$$\log \widehat{(RV)}_{t,t+h} = \alpha \cdot \log (RV)_{t-h,t} + (1 - \alpha) \cdot \log \widehat{(RV)}_{t-h,t} \tag{32}$$

with parameter α optimized for each rolling window following

$$\arg \min_{\alpha} \sum_{t} \left[\log (RV)_{t-h,t} - \log \widehat{(RV)}_{t-h,t} \right].$$

All the models, except the HAR-RV, are evaluated on log realized variation. They are also estimated on the same rolling window and evaluated on the same out-of-sample data points. It is also worth mentioning that jumps detection and jump variation estimation are executed in the same way as described in this paper.[5]

In order to evaluate forecasts of those models, the model confidence set (MCS) methodology of Hansen et al. (2003, 2011) is the most well-suited. The comparison is done among a set of models, as pairwise comparisons would not be appropriate. The methodology allows the models to be ranked based

---

[5]In their original work, Andersen et al. (2007a) do not apply a test for significant jump detection but jumps are identified as positive differences between *RV* and *BV*. Corsi and Renò (2010) apply a test based on the difference between *RV* and a threshold estimator for the continuous variation. Numerically, the results may differ slightly with different jump identification procedures. However, the performance of our proposed models is robust even compared to other testing procedures for jumps.

on a given loss function and it gives indication of whether the forecast performances are significantly different. Out of the surviving models in the confidence set, the interpretation is that they have equal predictive ability and they yield the best forecasts given a confidence level.

The model confidence set approach allows the user to specify different criteria to establish equal forecasting performance for the model set and subsets. We use both the "range" statistic, and the "semi-quadratic" statistic:

$$T_R = \max_{i,j \in M_{set}} \frac{|\bar{d}_{ij}|}{\sqrt{\widehat{var}(\bar{d}_{ij})}},$$ (33)

$$T_{SQ} = \sum_{i,j \in M_{set}} \frac{\bar{d}_{ij}^2}{\widehat{var}(\bar{d}_{ij})},$$ (34)

where $\bar{d}_{ij}$ is the mean loss differential between each pair combination of models.

Table 4 reports the model confidence set and the selected models at the 5% and 15% significance levels. The model confidence set p-value is obtained through a block bootstrap procedure. An autoregressive process is estimated for each d_{ij}, the loss differential between each model i and j, and the lag length for it is determined by Akaike information criteria as suggested by Hansen et al. (2003). The block length for the bootstrap procedure is then fixed as the maximum lag length among d_{ij} and it varies between 9 and 30 depending on the forecasting horizon and the loss function used. 5000 bootstrap repetitions are used to compute the test statistics. The model confidence set p-value obtained by using the semi-quadratic test statistics is less conservative then the one obtained by using the range statistics. The surviving model set after using the semi-quadratic test statistic is therefore larger than the set obtained by using the range test statistic.

[Table 4 about here.]

Overall, the best forecasting models for all different horizons are the candidate models (28) and (29) and the two LHAR-CJ models proposed by Corsi and Renò (2010), which take into account past negative return without disentangling return jumps size. The reduced model of eq. (29) achieved equal forecasting performance, compared with the LHAR-CJ+ model, with a lower model complexity (4 variables less). Jumps in return indeed substantially improve the forecasting performance, while jumps in continuous variation (called volatility jumps) are infrequent and have a minor effect on future realized

21

variation. Jumps in quadratic variation are in fact mainly due to return jumps. For the one period ahead forecast, the full model (28), which includes volatility jumps, does not survive based on the MSE loss function. However, it does survive based on the QLIKE loss function. This is due to the fact that by adding volatility jumps, the model tends to produce marginally less conservative forecasts of volatility. The QLIKE loss, since it less heavily penalizes over-predictions then under-predictions, still selects this model. Not surprisingly, the best forecasting models are the ones that include the leverage effect.

Finally, the following considerations can be pointed out for the simplest models: For long horizons, the simple HAR-RV model is hard to beat. As has already been shown previously in the literature, the autoregressive model that takes only the daily component into account is not selected for any fore-casting horizon and loss function used, pointing to the importance of correctly modeling long-memory. Moreover, the exponential smoothing model, often used in practice, is clearly outperformed.

5   Conclusion

This paper analyzed the performance of volatility forecasting models that take into account downside risk, jumps, and the leverage effect. The volatility forecasting model proposed consists of the following ingredients: First, the size of return jumps is identified based on the test of Lee and Mykland (2008). Second, jump variation and continuous variation are disentangled based on the methodology proposed by Andersen et al. (2010). Third, the size of jumps in volatility is also identified with inference based on an auxiliary AR-GARCH model. Finally, signed jump and continuous semivariations are computed using signed intraday returns. The best candidate model for forecasting realized volatilities must simul-taneously take into account return jumps, "good" and "bad" risks, leverage effect, and strong volatility persistence (i.e. long memory).

The model is motivated by overwhelming evidence of asymmetries in financial time series. We show that correlation asymmetries are present for both continuous and jump components among return and volatility processes. Moreover, asymmetries exist not only in size but also in sign, justifying the use of semivariances in forecasting volatility.

The forecasting model is very simple to implement as based on the parsimonious HAR framework. The gain over the base HAR-RV in terms of out-of-sample forecasting power is substantial and this is especially true for short and mid forecasting horizons. Therefore, volatility forecasts with return and jump asymmetries are warranted.

References

Andersen T, Bollerslev T, Diebold F. 2007a. Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics* **89**: 701–720.

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

Andersen T, Bollerslev T, Diebold F, Labys P. 2001b. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* **96**: 42–55.

Andersen T, Bollerslev T, Diebold F, Labys P. 2003. Modeling and forecasting realized volatility. *Econometrica* **71**: 579–625.

Andersen T, Bollerslev T, Dobrev D. 2007b. No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and iid noise: Theory and testable distributional implications. *Journal of Econometrics* **138**: 125–180.

Andersen T, Bollerslev T, Frederiksen P, Ørregaard Nielsen M. 2010. Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. *Journal of Applied Econometrics* **25**: 233–261.

Bae J, Kim C, Nelson C. 2007. Why are stock returns and volatility negatively correlated? *Journal of Empirical Finance* **14**: 41–58.

Baillie R, Bollerslev T, Mikkelsen H. 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **74**: 3–30.

Bandi F, Reno R. 2010. Time-varying leverage effects. *Journal of Econometrics* **Forthcoming**.

Barndorff-Nielsen O, Kinnebrock S, Shephard N. 2010. Measuring downside risk - realized semivariance. In Bollerslev T, Russel J, Watson M (eds.) *Volatility and Time Series Econometrics*, chapter 7. Oxford University Press, 117–137.

Barndorff-Nielsen O, Shephard N. 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* **64**: 253–280.

Barndorff-Nielsen O, Shephard N. 2004. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* **2**: 1–37.

Barndorff-Nielsen O, Shephard N. 2006. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* **4**: 1–30.

Bekaert G, Wu G. 2000. Asymmetric volatility and risk in equity markets. *Review of Financial Studies* **13**: 1–42.

Bollerslev T, Kretschmer U, Pigorsch C, Tauchen G. 2009. A discrete-time model for daily s&p500 returns and realized variations: Jumps and leverage effects. *Journal of Econometrics* **150**: 151–166.

Bollerslev T, Litvinova J, Tauchen G. 2006. Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics* **4**: 353.

Broadie M, Chernov M, Johannes M. 2007. Model specification and risk premia: Evidence from futures options. *Journal of Finance* **62**: 1453–1490.

Campbell J, Hentschel L. 1992. No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics* **31**: 281–318.

Chen X, Ghysels E. 2010. News - good or bad - and its impact on volatility predictions over multiple horizons. *Review of Financial Studies* **24**: 46–81.

Chernov M, Gallant A, Ghysels E, Tauchen G. 2003. Alternative models for stock price dynamics. *Journal of Econometrics* **116**: 225–257.

Christie A. 1982. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics* **10**: 407–432.

Clark T, McCracken M. 2001. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics* **105**: 85–110.

Clark T, West K. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* **138**: 291–311.

Corsi F. 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* **7**: 174–196.

Corsi F, Pirino D, Renò R. 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics* **159**: 276–288.

Corsi F, Renò R. 2010. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Working Paper* .

Diebold F, Mariano R. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* **13**: 253–263.

Ding Z, Granger C, Engle R. 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1**: 83–106.

Duffie D, Pan J, Singleton K. 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* **68**: 1343–1376.

Engle R, Bollerslev T. 1986. Modelling the persistence of conditional variances. *Econometric Reviews* **5**: 1–50.

Eraker B. 2004. Do stock prices and volatility jump? reconciling evidence from spot and option prices. *Journal of Finance* **59**: 1367–1404.

Eraker B, Johannes M, Polson N. 2003. The impact of jumps in volatility and returns. *Journal of Finance* **58**: 1269–1300.

Forsberg L, Ghysels E. 2007. Why do absolute returns predict volatility so well? *Journal of Financial Econometrics* **5**: 31–67. ISSN 1479-8409.

French K, Schwert G, Stambaugh R. 1987. Expected stock returns and volatility. *Journal of Financial Economics* **19**: 3–29.

Giacomini R, White H. 2006. Tests of conditional predictive ability. *Econometrica* **74**: 1545–1578.

Hansen P, Lunde A. 2006a. Consistent ranking of volatility models. *Journal of Econometrics* **131**: 97–121.

Hansen P, Lunde A. 2006b. Realized variance and market microstructure noise. *Journal of Business and Economic Statistics* **24**: 127–161.

Hansen P, Lunde A, Nason J. 2003. Choosing the best volatility models: The model confidence set approach. *Oxford Bulletin of Economics and Statistics* **65**: 839–861.

Hansen P, Lunde A, Nason J. 2011. The model confidence set. *Econometrica* **79**: 453–497.

Lee S, Mykland P. 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial studies* **21**: 2535–2563.

Mykland P, Zhang L. 2009. Inference for continuous semimartingales observed at high frequency. *Econometrica* **77**: 1403–1445.

Newey W, West K. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**: 703–708.

Patton A. 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* **160**: 246–256.

Patton A, Sheppard K. 2011. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Working Paper* .

Schwert G. 1989. Why does stock market volatility change over time? *Journal of Finance* **44**: 1115.

Taylor J. 2004. Volatility forecasting with smooth transition exponential smoothing. *International Journal of Forecasting* **20**: 273–286.

Todorov V, Tauchen G. 2010. Volatility jumps. *Forthcoming in Journal of Business and Economic Statistics* .

Veronesi P. 1999. Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies* **12**: 975–1007.

Wu G. 2001. The determinants of asymmetric volatility. *Review of Financial Studies* **14**: 837–859.

Year	n. Days with Return Jump	% Days with Return Jump	Mean Jump Intensity (One Day)	Max Jump Intensity (One Day)	% Days with Volatility Jump
1982*	31	17.92	1.10	3	6.94
1983	29	11.51	1.38	6	1.98
1984	48	19.05	1.13	2	3.57
1985	50	19.76	1.22	4	0.40
1986	48	18.97	1.60	7	5.14
1987	36	14.63	1.89	10	5.28
1988	45	17.86	1.47	6	3.97
1989	42	16.67	1.48	8	2.38
1990	47	18.65	1.51	5	3.17
1991	48	19.05	1.46	8	2.78
1992	60	23.9	1.47	8	1.20
1993	67	26.59	1.51	6	0.40
1994	57	22.71	1.47	5	1.59
1995	48	19.2	1.38	4	0.40
1996	49	19.6	1.49	5	4.40
1997	48	19.35	1.77	6	7.26
1998	48	19.2	1.33	4	5.20
1999	49	19.6	1.39	4	5.60
2000	39	15.6	1.59	7	6.00
2001	39	15.92	1.41	3	5.71
2002	33	13.31	1.12	3	7.66
2003	50	20.16	1.12	3	4.03
2004	47	19.34	1.43	4	1.23
2005	67	27.13	1.33	4	0.00
2006	50	20.24	1.32	4	0.81
2007	60	24.19	1.65	7	3.63
2008	48	19.43	1.38	4	7.29
2009	44	17.6	1.41	5	4.80
2010**	35	23.33	1.49	4	4.67
Full Sample		19.29	1.43		3.65

Table 1: Jump Descriptive Statistics

NOTE: The table reports the number or days with return jumps, their percentage, the average daily jump intensity conditional on the presence of at least one jump, the maximum daily jump intensity, and the percentage of days with jumps in continuous variation. All the statistics are sorted by year. The tests for jumps in return and in continuous variation are conducted with significance level $\alpha = 0.01$. * The sample period for 1982 starts on April, 28. ** The sample period for 2010 stops on August, 6.

| Horizon | Model | Associated Variables | | | | | | | | | | | |
		Cons	CSV+	CSV-	JSV+	JSV-	VolJ	cRet-	cRet-(t-5,t)	jRet	RV (t-5,t)	RV (t-22,t)	adj-R2
1 day	1	0.030	0.1365	0.1458	-0.074	0.1946	-0.001	-0.194	-0.173		0.330	0.263	0.727
		[1.16]	[6.29]	[7.41]	[-0.91]	[2.17]	[-0.04]	[-9.6]	[-4.44]		[10.64]	[11.37]	
	2	0.098	0.1581	0.136				-0.237		-0.099	0.375	0.235	0.726
		[4.89]	[7.78]	[7.1]				[-14.66]		[-2.64]	[12.92]	[10.59]	
5 days	1	-0.004	0.113	0.0875	-0.089	0.092	0.0348	-0.125	-0.128		0.321	0.349	0.798
		[-0.14]	[7.01]	[5.44]	[-1.49]	[1.69]	[1.26]	[-7.48]	[-3.29]		[8.37]	[10.61]	
	2	0.047	0.1295	0.082				-0.157		-0.053	0.352	0.328	0.797
		[2.27]	[8.39]	[5.28]				[-10.39]		[-2.21]	[9.73]	[10.29]	
15 days	1	-0.064	0.096	0.0497	-0.061	0.061	0.0433	-0.089	-0.130		0.271	0.413	0.777
		[-1.92]	[6.09]	[3.14]	[-1.19]	[1.39]	[1.74]	[-6.92]	[-2.75]		[6.23]	[9.58]	
	2	-0.010	0.1134	0.046				-0.123		-0.034	0.301	0.392	0.776
		[-0.35]	[6.92]	[2.83]				[-7.1]		[-1.84]	[7.15]	[9.11]	
22 days	1	-0.078	0.084	0.0437	-0.054	0.051	0.0456	-0.080	-0.099		0.252	0.433	0.754
		[-2.13]	[5.42]	[2.62]	[-0.97]	[1.16]	[1.86]	[-6.3]	[-2.01]		[5.27]	[9.63]	
	2	-0.034	0.0985	0.0416				-0.106		-0.026	0.2732	0.4164	0.754
		[-1.05]	[5.97]	[2.53]				[-5.97]		[-1.41]	[5.93]	[9.14]	

Table 2: In Sample Regressions

NOTE: The table contains estimates of the coefficients and t-statistics, in square brackets, based on Newey-West HAC consistent standard errors, for the two model specifications. Model 1 and model 2 correspond to equations (28) and (29), respectively. The models are estimated for forecasting horizons of 1, 5, 15 and 22 days.

	Model	Horizon			
		1 day	5 days	15 days	22 days
DM Test Statistics	1	4.87	4.34	2.38	1.80
with QLIKE Losses	2	4.44	4.65	3.16	3.13
CW Test Statistics	1	11.34	10.83	7.72	6.67
with MSE Losses	2	11.70	10.84	7.73	6.79

Table 3: Out-of-Sample Test

NOTE: The table reports results from the out-of-sample pairwise forecast performance comparison. The performance is based on the QLIKE and MSE loss functions for models 1 and 2 corresponding to equations (28) and (29), respectively. Each model is evaluated against the benchmark HAR-RV (eq. 26). Forecast horizons considered are 1 day, 5 days, 15 days, and 22 days. The upper panel reports the Diebold-Mariano t-statistics on QLIKE losses and the bottom panel reports the Clark-West test statistics on MSE losses. The out-of-sample performance period ranges from May 1990 to August 2010.

| Model | Statistics | QLIKE Loss Function | | | | MSE Loss Function (on log RV) | | | |
| | | Forecasting Horizon | | | | Forecasting Horizon | | | |
		1 Day	5 Days	15 Days	22 Days	1 Day	5 Days	15 Days	22 Days
HAR (log RV)	Mean loss (Full Sample)	0.524	0.453	0.455	0.465	0.260	0.158	0.163	0.176
	MCS (R-Stat)								
	MCS (SQ-Stat)			+	++			+	++
HAR (RV)	Mean loss (Full Sample)	0.614	0.682	0.743	0.773	0.650	0.629	0.661	0.682
	MCS (R-Stat)								
	MCS (SQ-Stat)								
Mod. 1	Mean loss (Full Sample)	0.514	0.449	0.453	0.463	0.245	0.152	0.159	0.174
	MCS (R-Stat)	++	++	++	++		++	++	++
	MCS (SQ-Stat)	++	++	++	++		++	++	++
Mod. 2	Mean loss (Full Sample)	0.515	0.449	0.453	0.463	0.244	0.152	0.159	0.173
	MCS (R-Stat)	++	++	++	++	+	++	++	++
	MCS (SQ-Stat)	++	++	++	++	+	++	++	++
Semivariances - HAR 1 (PS, 2011)	Mean loss (Full Sample)	0.521	0.452	0.456	0.466	0.255	0.157	0.163	0.178
	MCS (R-Stat)								
	MCS (SQ-Stat)			++	+			++	++
Semivariances - HAR 2 (PS, 2011)	Mean loss (Full Sample)	0.521	0.452	0.455	0.464	0.255	0.156	0.162	0.176
	MCS (R-Stat)								
	MCS (SQ-Stat)			++	++			++	++
HAR-RV-J (ABD, 2007)	Mean loss (Full Sample)	0.524	0.453	0.455	0.465	0.259	0.158	0.162	0.176
	MCS (R-Stat)								
	MCS (SQ-Stat)			+	++			++	++
HAR-RV-CJ (ABD, 2007)	Mean loss (Full Sample)	0.524	0.452	0.455	0.465	0.258	0.157	0.161	0.176
	MCS (R-Stat)			++				++	
	MCS (SQ-Stat)			++	++			++	++
LHAR-CJ (CR, 2010)	Mean loss (Full Sample)	0.513	0.449	0.454	0.464	0.242	0.151	0.160	0.175
	MCS (R-Stat)	++	++	++	++	++	++	++	++
	MCS (SQ-Stat)	++	++	++	++	++	++	++	++
LHAR-CJ+ (CR, 2010)	Mean loss (Full Sample)	0.513	0.449	0.454	0.464	0.242	0.151	0.160	0.175
	MCS (R-Stat)	++	++	++	++	++	++	++	+
	MCS (SQ-Stat)	++	++	++	++	++	++	++	++
Autoregressive 1 day	Mean loss (Full Sample)	0.562	0.502	0.513	0.524	0.325	0.247	0.265	0.282
	MCS (R-Stat)								
	MCS (SQ-Stat)								
Exponential Smoothing	Mean loss (Full Sample)	0.527	0.460	0.474	0.490	0.265	0.170	0.192	0.210
	MCS (R-Stat)								
	MCS (SQ-Stat)								

Table 4: Model Confidence Set

NOTE: The table reports forecasting performance of models in the confidence set. A block bootstrap procedure with 5000 replications is used to establish equal predictive ability of the surviving models according to both "range" and "semi-quadratic" test statistics. Moreover, both QLIKE and MSE (on log realized variation) losses are used. Forecast horizons considered are 1 day, 5 days, 15 days, and 22 days. "+" denotes included models in the final confidence set with significance level 5%, and "++" denotes included models in the confidence set with significance level 15%, with $MCS_{15\%} \subset MCS_{5\%}$. Mod. 1 and Mod. 2 correspond to equations (28) and (29), respectively.
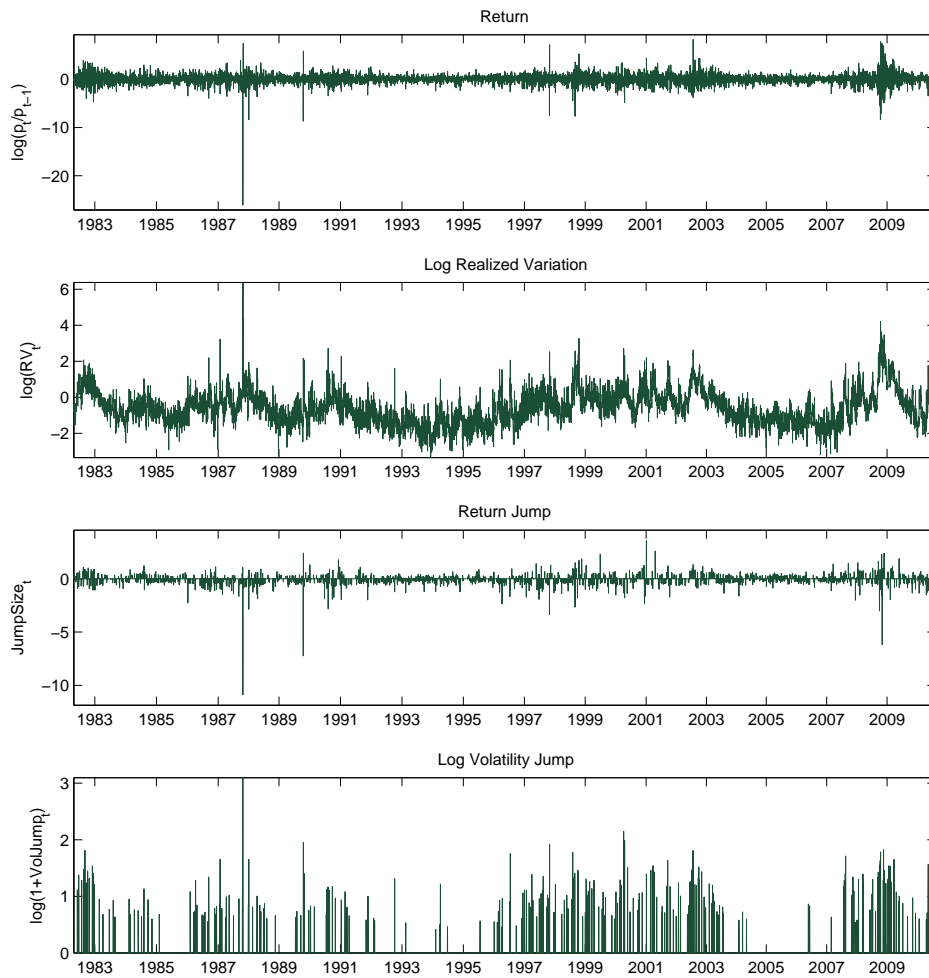
Figure 1: Time Series of Return, Volatility and Jumps

The figure plots time series of daily log return, log realized variation, jump in log returns and log jump in continuous variation. The underlying security is the S&P 500 index Future and the time period ranges from April 28, 1982 to August 6, 2010.
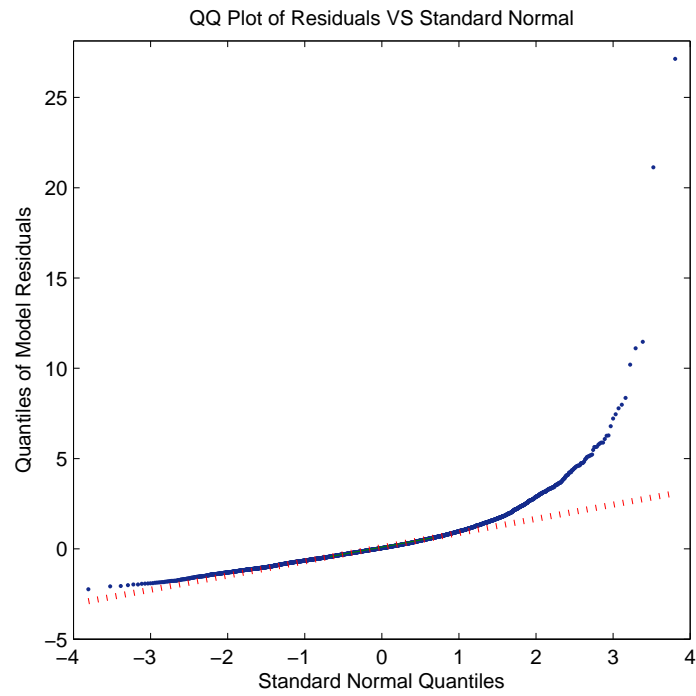
Figure 2: GARCH Residuals and Volatility Jumps

The figure represents the quantiles of the standardized residuals from the GARCH fitting against the quantiles of the standard normal distribution in blue dots. The red dots represent the quantiles of the standard normal distribution against itself.
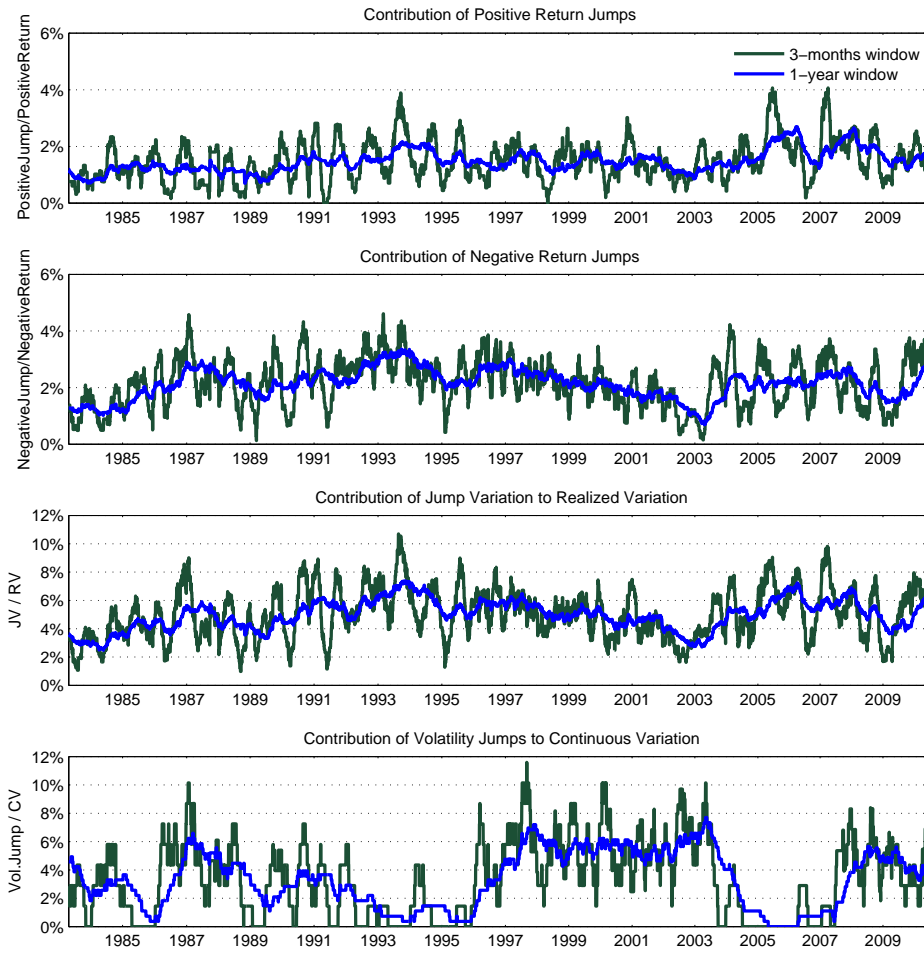
Figure 3: Jump Contributions

The figure plots the average contributions of positive jumps to positive return, negative jumps to negative return, jump variation to realized variation and volatility jumps to continuous variation. The percentages are calculated by using rolling windows consisting of three months and one year.
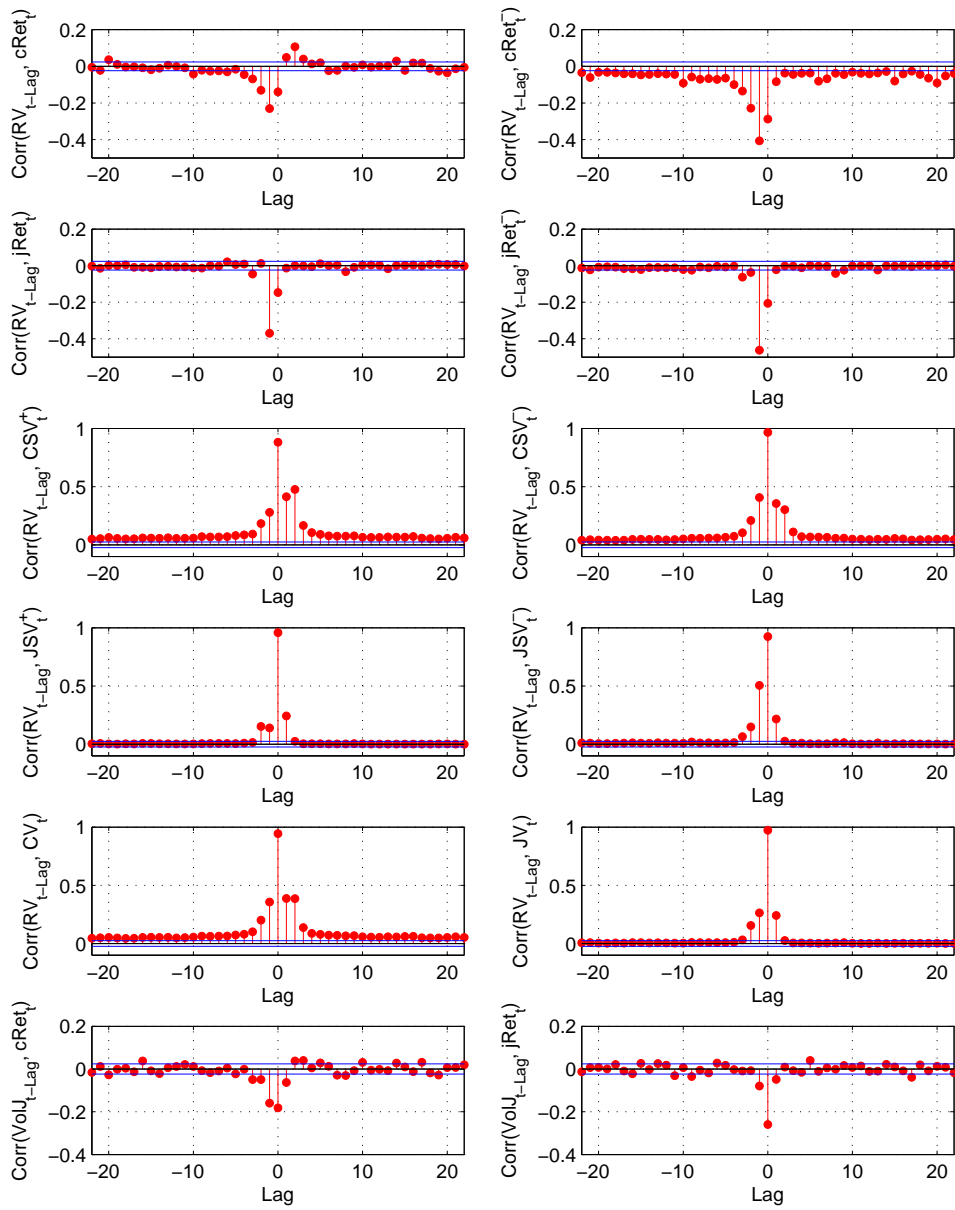
Figure 4: Sample Cross-correlation Between Realized Variation, Return and Jumps

The figure plots pairwise sample cross-correlation of daily realized variations with daily returns and return jumps (first row), signed returns and jumps (second and third rows), continuous and jump variations (fourth row) and the cross-correlation between volatility jumps and continuous and jump returns (last row). Lags and leads of up to 22 days are considered.
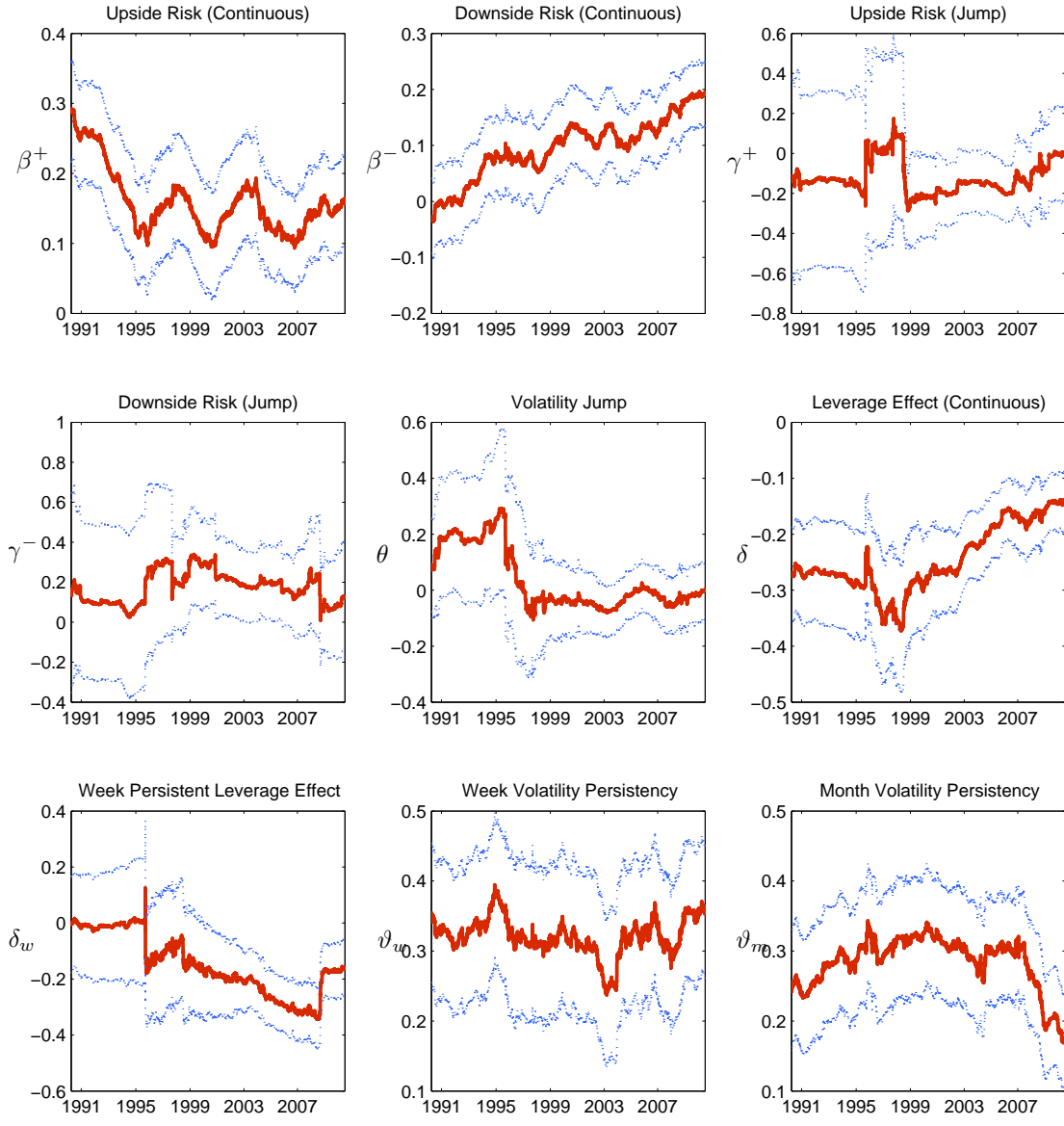
Figure 5: Rolling Window Coefficients for One Period Forecasts

The figure plots rolling window coefficient estimates with the correspective 95% confidence interval, based on Newey-West HAC consistent standard errors. The coefficients are the ones associated with the model in eq. (28) for one period ahead forecasts. The window size consists of 2000 observations.
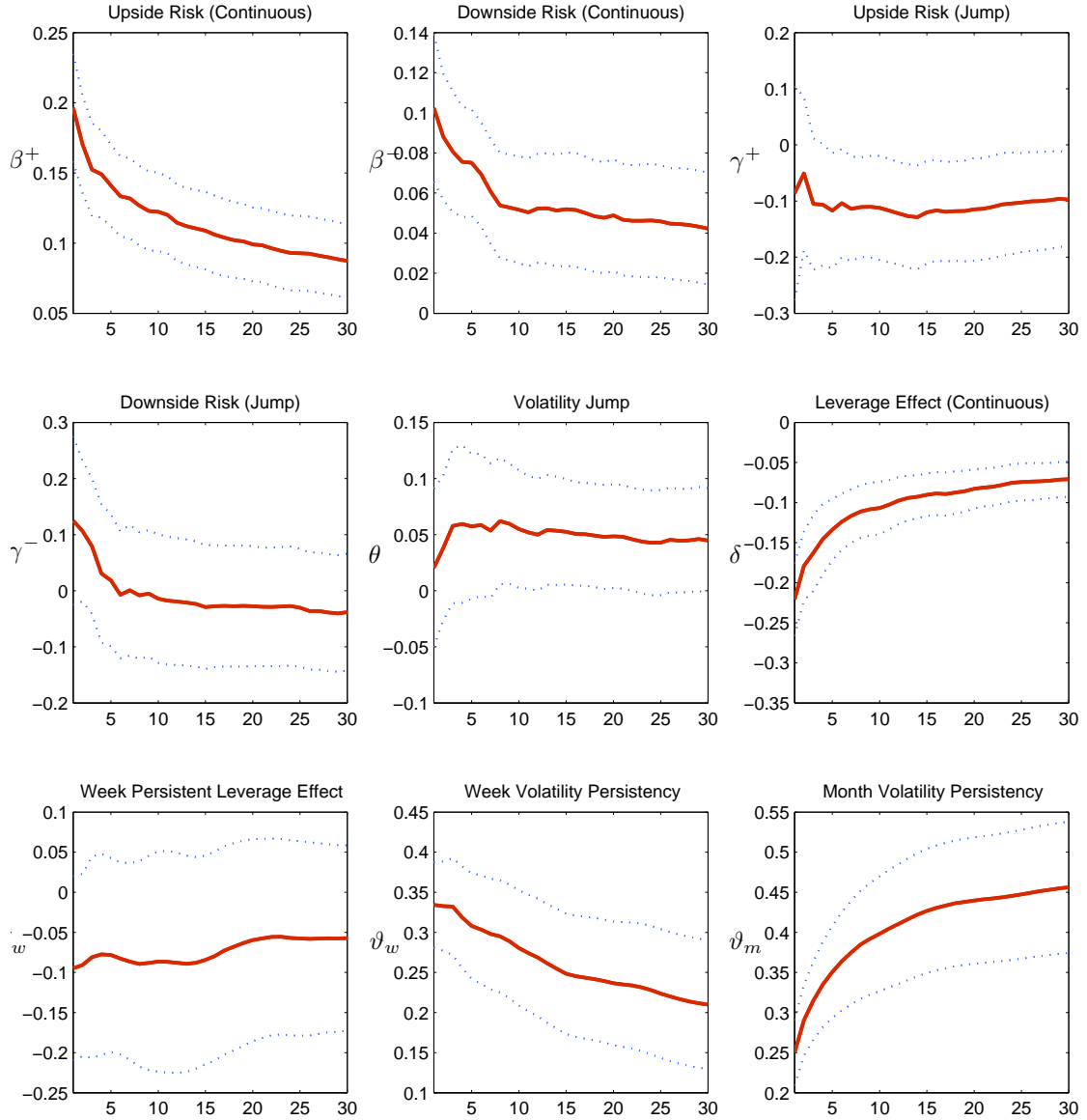
Figure 6: Estimated Coefficients to Different Forecasting Horizons

The figure plots coefficient estimates with the correspective 95% confidence interval, based on Newey-West HAC consistent standard errors, for the model in eq. (28). The forecasting horizons range from 1 to 30 days.