

# HAR volatility modelling with heterogeneous leverage and jumps\*

Fulvio Corsi<sup>†</sup>      Roberto Reno<sup>‡</sup>

July 8, 2009

## Abstract

We propose a dynamic model for financial market volatility with an heterogeneous structure for three components: continuous volatility, leverage and jumps. We find that each of the three components plays a significant role in volatility forecasting and neglecting one of them is detrimental to the forecasting performance. Importantly, we find remarkable forecasting power for the negative past returns at all the considered frequencies, which unveils a novel heterogeneous structure in the leverage effect. We also show, using simulation studies, that the presence of jumps is important for two distinct reasons: Firstly, explicitly modeling jumps has trimming effect on the dynamics of the persistent volatility component; secondly, they have a positive and significant impact on future volatility.

**JEL classification:** C13; C22; C51; C53

**Keywords:** Volatility Forecasting; High Frequency Data; Two-Scale Realized Volatility; HAR; Leverage Effect; Jumps.

---

\*This paper supersedes the previously circulating version "Volatility determinants: heterogeneity, leverage and jumps". We would like to acknowledge Davide Pirino for research assistance, Tim Bollerslev, Giampiero Gallo and Roel Oomen for useful suggestions, and the participants at the SITE Summer Segment in Stanford, June 2009. All errors and omissions are our own.

<sup>†</sup>Università di Siena, University of Lugano, and Swiss Finance Institute, E-mail: fulvio.corsi@lu.unisi.ch

<sup>‡</sup>Università di Siena, Dipartimento di Economia Politica, E-mail: reno@unisi.it

# 1 Introduction

The importance of financial market volatility led to a very large literature in which volatility dynamics has been modelled to take into account its most salient features (clustering, slowly decaying auto-correlation, asymmetric responses). This paper contributes to this literature by proposing a reduced form discrete-time model in which, after separating quadratic variation in the part due to the continuous variation and the part due to jumps,<sup>1</sup> volatility is assumed to depend on three components: continuous volatility, negative returns and jumps. In order to reproduce the persistent impact of each component observed in the data, we impose a common heterogeneous structure. In particular, while it is well known that volatility tends to increase more after a negative shock than after a positive shock of the same magnitude, see Christie (1982); Campbell and Hentschel (1992); Glosten et al. (1989) and more recently Bollerslev et al. (2006), we provide novel evidence on the fact that the impact of negative returns on future volatility (the so-called *leverage effect*) is also highly persistent and extends for a period of at least one month.

In Corsi (2009) a simple Heterogeneous Auto-Regressive (HAR) model has been proposed for realized volatility to capture the empirical memory persistence of volatility in a simple and parsimonious way. In this paper, we propose an extended version of the HAR model which considers asymmetric responses of the realized volatility not only to previous daily *negative* returns, but also to their weekly and monthly aggregation. Our main contribution is then to show that the heterogeneous structure applies to the leverage effect as well, thus reinforcing the Heterogeneous Market Hypothesis of Muller et al. (1997).

In addition we study the impact on future volatility of jumps (discontinuous variations of the asset price) measured over the same three different horizons. Given the inadequacy of bipower variation in measuring volatility in presence of jumps, we use the tests and measures introduced by Corsi et al. (2008) which provide a better identification and more precise measurement of jumps, and uncover the significant impact of jumps on future volatility. We confirm and reinforce this finding in a larger sample, at a larger aggregation frequency and out-of-sample. We also suggest, by means of a simulation study, that the presence of jumps is important for two distinct reasons. First, it has a direct impact on volatility dynamics which may be explained by the presence of contemporaneous jumps in price and volatility. Second, it has a trimming effect on the volatility series, which allows a better fit of the realized volatility, as suggested by Andersen et al. (2007).

Finally, we conduct robustness tests to check whether other volatility measures proposed in the literature (such as absolute variation, range, or semivariance) contain additional useful information which are not captured in our model specification. The results of these comparisons show that the other volatility measures contribute only marginally to the performance of the model.

The paper is organized as follows. Section 2 reviews the HAR model and presents its possible extensions

---

<sup>1</sup>Several techniques for separating the continuous variation from the discontinuous variation have been explored in Barndorff-Nielsen and Shephard (2004); Mancini (2009); Corsi et al. (2008); ? among others.

# Realized volatility memory

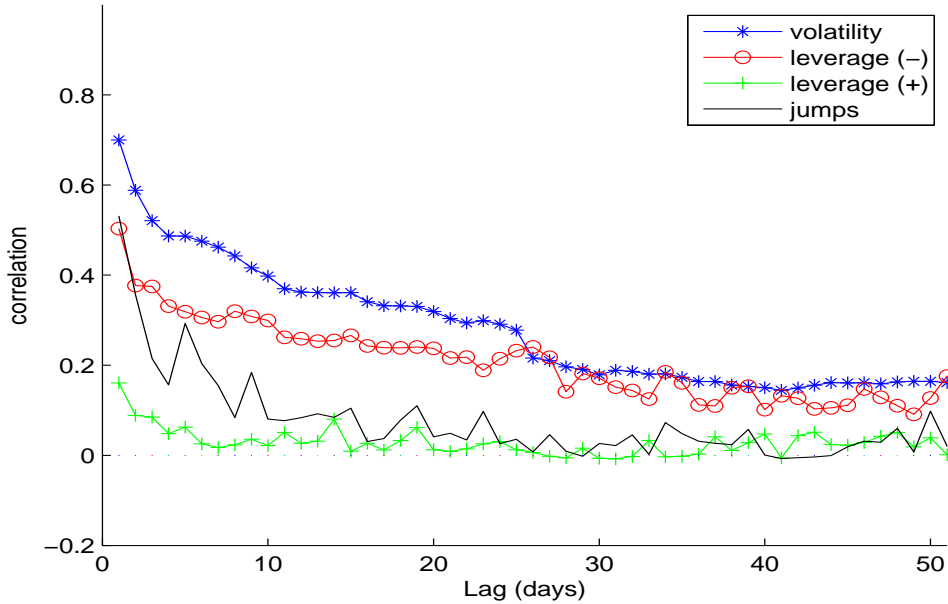


Figure 1:

with heterogeneous leverage and jumps. Section 3 describes the empirical in-sample and out-of-sample analysis on a long series of high frequency S&P500 futures data. Section 4 discuss the results by the light of two Monte Carlo simulations (with independent or contemporaneous jumps in price and volatility) and Section 5 contains concluding remarks.

## 2 Motivation and model specification

### 2.1 Motivation

Figure 1 shows, for the S&P 500 series studied in this paper, the lagged correlation function between the estimated daily integrated variance  $RV_{t+h}$  with  $X_t$  as a function of  $h$ , with  $X_t$  being  $RV_t$  itself, negative returns, positive returns and jumps (details on the data and the construction of all these quantities are provided in Section 3). When  $X_t = RV_t$  we are estimating the usual autocorrelation function of realized volatility, which displays the well known feature to be very slowly decaying and possibly long memory. When  $X_t$  are the negative returns, we observe that they have a large impact on future volatility, and this is the well known leverage effect. However, Figure 1 also shows that the impact of negative returns on future volatility is slowly decaying as well. Also jumps have a positive and large impact (as large as the negative returns when  $h = 1$ ), but with a faster decaying pace. Finally, positive returns have a very

small and almost insignificant impact on future volatility.

The slowly decaying impact of negative returns might well be a by-product of the slowly decaying auto-correlation function of volatility. However, since the same phenomenon is not observed with the jump component, it can be also suggestive of the fact that leverage effect might be very persistent (potentially long memory), a possibility which has been neglected so far.

This paper explores exactly this possibility. We follow Corsi (2009) in modelling the slowly decaying auto-correlation function by means of an heterogeneous structure induced by a volatility cascade. Since this has proven to be quite successful in modelling the volatility dynamics, we extend the heterogeneous structure to the negative returns and jumps, thus enforcing the views of Muller et al. (1997).

## 2.2 Extending the HAR model

We assume that the state variable  $X$ , which may be thought as an economic variable (an interest rate or a stock log-price), is driven by the stochastic process:

$$dX_t = \mu_t dt + \sigma_t dW_t + c_t dN_t \quad (2.1)$$

where  $\mu_t$  is predictable,  $\sigma_t$  is càdlàg and  $N_t$  is a doubly stochastic Poisson process<sup>2</sup> whose intensity is an adapted stochastic process  $\lambda_t$ , the random times of the corresponding jumps are  $(\tau_j)_{j=1, \dots, N_T}$  and  $c_j$  are i.i.d. adapted random variables measuring the size of the jump at time  $\tau_j$ . In practice, e.g. for risk management purposes, we are interested in forecasting the quadratic variation defined as:

$$\tilde{\sigma}_t = \int_t^{t+1} \sigma_s^2 ds + \sum_{t \leq \tau_j \leq t+1} c_{\tau_j}^2,$$

where the time unit is one day. We estimate quadratic variation using  $n$  observations of the state variable in the interval  $[0, T]$ . The most popular estimator is realized volatility, which, after defining  $\Delta_{t,j}X = X_{t+j/n} - X_{t+(j+1)/n}$ , is given by:

$$\text{RV}_t = \sum_{j=0}^{n-1} (\Delta_{t,j}X)^2 \quad (2.2)$$

which is a consistent estimator, as  $n \rightarrow \infty$ , of  $\tilde{\sigma}_t$ , see Andersen et al. (2003) for a review. Other estimators have been devised, such as the range (see e.g. Alizadeh et al. 2002 and Brandt and Jones 2006) or refinement of realized volatility to account for the presence of microstructure noise, as those in Zhang et al. (2005), Barndorff-Nielsen et al. (2008), or Jacod et al. (2007). We indicate by  $\widehat{V}_t$  a generic unbiased estimator of quadratic variation such that:

$$\tilde{\sigma}_t = \widehat{V}_t + \omega_t \quad (2.3)$$

---

<sup>2</sup>We could also consider a wider class of jumps, such as Lévy, in the case in which they have a finite quadratic variation process.

where  $\omega_t$  is a zero mean and finite variance measurement error.

The need for heterogeneity of volatility components, advocated by Muller et al. (1997), has been reconsidered in the work of Corsi (2009) by making use of the concept of volatility cascades. In what follows, we briefly review this latter approach. Consider the aggregated values of  $\widehat{V}_t$ , defined as:

$$\widehat{V}_t^{(n)} = \frac{1}{n} \left( \widehat{V}_t + \dots + \widehat{V}_{t-n+1} \right) \quad (2.4)$$

and assume having two different time scales, of length  $n_1$  and  $n_2$ , with  $n_1 > n_2$ . For the largest time scale, assume that  $\tilde{\sigma}_t$ , once aggregated as in (2.4) is determined by:

$$\tilde{\sigma}_{t+n_1}^{(n_1)} = c^{(n_1)} + \beta^{(n_1)} \widehat{V}_t^{(n_1)} + \varepsilon_{t+n_1}^{(n_1)} \quad (2.5)$$

where  $\varepsilon_t^{(n_1)}$  is IID zero mean and finite variance noise independent on  $\omega_t$ , and  $c^{(n_1)}$  and  $\beta^{(n_1)}$  are constants.

It has been recently suggested that volatility over longer time intervals has stronger influence on those over shorter time intervals than conversely, suggesting a volatility cascade from low to high frequencies.<sup>3</sup> This can be explained by assuming that for short-term traders the level of long term volatility matters because it determines the expected future size of trends and risk. On the other hand, the level of short-term volatility does not affect the trading strategies of long-term traders. The shorter time scale ( $n_2$ ) is assumed to be influenced by the expected future value of the largest time scale ( $n_1$ ), so that:

$$\tilde{\sigma}_{t+n_2}^{(n_2)} = c^{(n_2)} + \beta^{(n_2)} \widehat{V}_{t+n_2}^{(n_2)} + \delta^{(n_2)} \mathbf{E}_t \left[ \tilde{\sigma}_{t+n_1}^{(n_1)} \right] + \varepsilon_{t+n_2}^{(n_2)}; \quad (2.6)$$

with  $\varepsilon_t^{(n_2)}$  IID zero mean and finite variance noise independent on  $\varepsilon_t^{(n_1)}$  and  $\omega_t$  and  $\delta^{(n_2)}$  is a constant.

By substitution, and using equation (2.3), this gives:

$$\widehat{V}_{t+n_2}^{(n_2)} = c + \beta^{(n_2)} \widehat{V}_t^{(n_2)} + \beta^{(n_1)} \widehat{V}_t^{(n_1)} + \varepsilon_t; \quad (2.7)$$

where  $\varepsilon_t$  is IID noise depending on  $\varepsilon_t^{(n_1)}$ ,  $\varepsilon_t^{(n_2)}$ ,  $\omega_t$ . The model (2.7) can be easily extended to  $d$  horizons of length  $n_1 > n_2 > \dots > n_d$ . Typically, three components are used with length  $n_1 = 22$  (monthly),  $n_2 = 5$  (weekly),  $n_3 = 1$  (daily). Since volatility at shorter time horizons is influenced by volatility at longer horizons, the auto-correlation function of the model and hence its memory persistence increases. Thus, even if the HAR model does not formally belong to the class of long memory processes, it fits the persistence properties of financial data as well as true long memory models, such as the fractionally integrated one, which, however, are much more complicated to estimate and to deal with (see the review of Banerjee and Urga 2005). For these reasons, the HAR model has been employed in several applications in the literature. Corsi et al. (2008) use it to study the volatility of realized volatility; Ghysels et al. (2006) and Forsberg and Ghysels (2007) compare this model with the MIDAS model; Andersen et al. (2007) use an extension of this model to forecast the volatility of stock prices, foreign exchange rates and

---

<sup>3</sup>See Müller et al. (1997), Arneodo et al. (1998), Lynch and Zumbach (2003). However, the HAR model would hold even if we allow the short-term volatility to affect the long-term volatility, although this would be at odds with the empirical findings (see Section 3.1).

bond prices; Clements et al. (2008) implement it for risk management with VaR measures; Bollerslev et al. (2008) use it to analyze the risk-return tradeoff; Bianco et al. (2009) use it to study the relation between intraday serial correlation and volatility.

It is then natural to extend the Heterogeneous Market Hypothesis approach to leverage effect. We assume that asymmetric responses of realized volatility not only to previous daily returns but also to past weekly and monthly returns. We model such heterogeneous leverage effects by introducing asymmetric return-volatility dependence at each level of the cascade considered in the above section. Define daily returns  $r_t = X_t - X_{t-1}$  and past aggregated negative returns as:

$$r_t^{(n)-} = \frac{1}{n} (r_t + \dots + r_{t-n+1}) I_{\{(r_t + \dots + r_{t-n+1}) < 0\}} \quad (2.8)$$

where  $I_{\{\cdot\}}$  denotes the indicator function. We assume that integrated volatility is determined by the cascade:

$$\begin{aligned} \tilde{\sigma}_{t+n_1}^{(n_1)} &= c^{(n_1)} + \beta^{(n_1)} \widehat{V}_t^{(n_1)} + \gamma^{(n_1)} r_t^{(n_1)-} + \varepsilon_{t+n_1}^{(n_1)} \\ \tilde{\sigma}_{t+n_2}^{(n_2)} &= c^{(n_2)} + \beta^{(n_2)} \widehat{V}_t^{(n_1)} + \gamma^{(n_2)} r_t^{(n_2)-} + \delta^{(n_2)} \mathbf{E}_t \left[ \tilde{\sigma}_{t+n_1}^{(n_1)} \right] + \varepsilon_{t+n_2}^{(n_2)} \end{aligned}$$

where  $\gamma^{(n_1,2)}$  are constants. This now gives:

$$\widehat{V}_{t+n_2}^{(n_2)} = c + \beta^{(n_2)} \widehat{V}_t^{(n_2)} + \beta^{(n_1)} \widehat{V}_t^{(n_1)} + \gamma^{(n_2)} r_t^{(n_2)-} + \gamma^{(n_1)} r_t^{(n_1)-} + \tilde{\varepsilon}_t \quad (2.9)$$

## 2.3 Jumps

The importance of jumps in financial econometrics is rapidly growing. Recent research focusing on jumps detection and volatility measuring in presence of jumps includes Barndorff-Nielsen and Shephard (2004); Mancini (2009), Lee and Mykland (2008), Jiang and Oomen (2008), Aït-Sahalia and Mancini (2008), ?, Christensen et al. (2008), and Boudt et al. (2008). Andersen et al. (2007) suggested that the continuous volatility and jump component have different dynamics and should thus be modelled separately. In this section, we follow closely Corsi et al. (2008) using the C-Tz statistics to detect the occurrence of the jump in a single day, and threshold bipower variation to measure the continuous part of integrated volatility, defined as:

$$\text{TBPV}_t = \frac{2}{\pi} \sum_{j=0}^{n-2} |\Delta_{t,j} X| \cdot |\Delta_{t,j+1} X| I_{\{|\Delta_{t,j} X|^2 \leq \vartheta_{j-1}\}} I_{\{|\Delta_{t,j+1} X|^2 \leq \vartheta_j\}} \quad (2.10)$$

where  $\vartheta_t$  is a threshold function which we estimate as in Corsi et al. (2008). This continuous volatility estimator has better finite sample properties than standard bipower variation and provides more accurate jump tests, which allows for a corrected separation of continuous and jump components. To this purpose, we fix a confidence level  $\alpha$  and estimate the jump component as:

$$J_t = I_{\{C\text{-Tz} > \Phi_\alpha\}} \cdot \left( \widehat{V}_t - \text{TBPV}_t \right)^+ \quad (2.11)$$

## Jump Contribution to Total Variation

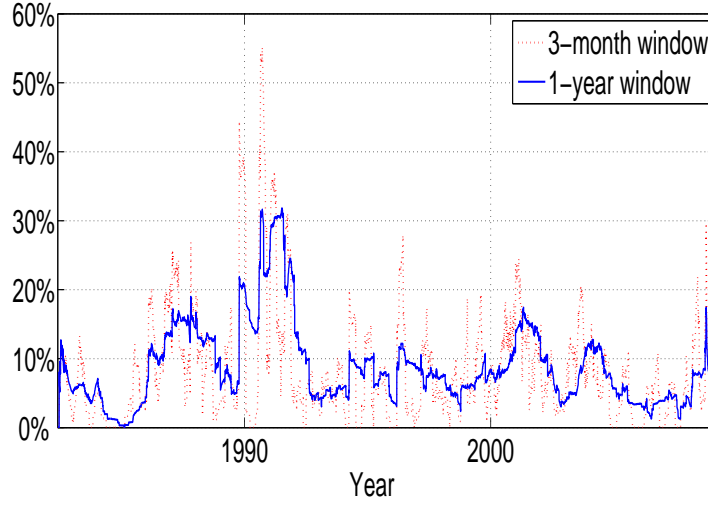


Figure 2: Percentage contribution of daily jump estimated by (2.11) to total quadratic variation measured over a moving window of 3-month (dotted line) and 1-year (solid line) for the S&P500 futures from 28 April 1982 to 5 February 2009 (6,669 observations) excluding the October 1987 crash. The C-Tz statistics in (2.11) is computed with a confidence interval  $\alpha = 99.9\%$ .

where  $\Phi_\alpha$  is the value of the standard Normal distribution corresponding to the confidence level  $\alpha$ , and  $x^+ = \max(x, 0)$ . The corresponding continuous component is defined as:

$$C_t = \widehat{V}_t - J_t, \quad (2.12)$$

which is equal to  $\widehat{V}_t$  if there are no jumps in the trajectory and to  $TBPV_t$  if a jump is detected with the C-Tz statistics.

Figure 2 reports the percentage contribution of jumps estimated by (2.11) to total quadratic variation computed on a 3-month and 1-year moving window for the full S&P 500 futures sample. In line with the results in Andersen et al. (2007) and Huang and Tauchen (2005) we find a jumps contribution varying between 2% and 20% of total variation (with an overall sample mean of about 6%), showing a higher level at the beginning of the sample and an increasing trend toward the end of the sample period.

Coherently with the above section, in the volatility cascade we assume that  $C_t$  and  $J_t$  enter separately at each level of the cascade, that is:

$$\begin{aligned} \widetilde{\sigma}_{t+n_1}^{(n_1)} &= c^{(n_1)} + \alpha^{(n_1)} J_t^{(n_1)} + \beta^{(n_1)} C_t^{(n_1)} + \gamma^{(n_1)} r_t^{(n_1)-} + \varepsilon_{t+n_1}^{(n_1)} \\ \widetilde{\sigma}_{t+n_2}^{(n_2)} &= c^{(n_2)} + \alpha^{(n_2)} J_t^{(n_2)} + \beta^{(n_2)} C_t^{(n_2)} + \gamma^{(n_2)} r_t^{(n_2)-} + \delta^{(n_2)} \mathbf{E}_t \left[ \widetilde{\sigma}_{t+1}^{(n_1)} \right] + \varepsilon_{t+n_2}^{(n_2)} \end{aligned}$$

originating the model:

$$\widehat{V}_{t+n_2}^{(n_2)} = c + \alpha^{(n_1)} J_t^{(n_1)} + \alpha^{(n_2)} J_t^{(n_2)} + \beta^{(n_2)} C_t^{(n_2)} + \beta^{(n_1)} C_t^{(n_1)} + \gamma^{(n_2)} r_t^{(n_2)-} + \gamma^{(n_1)} r_t^{(n_1)-} + \varepsilon_t$$

## 2.4 The LHAR-CJ model

Combining heterogeneity in realized volatility, leverage, and jumps we construct the *Leverage Heterogeneous Auto-Regressive with Continuous volatility and Jumps* (LHAR-CJ) model. As it is common in practice, we use three components for the volatility cascade: daily, weekly and monthly. Moreover, we use the logarithms of volatility measures. Hence, the proposed model reads:

$$\begin{aligned} \log \widehat{V}_{t+h}^{(h)} = c &+ \alpha^{(d)} \log(1 + J_t) + \alpha^{(w)} \log(1 + J_t^{(5)}) + \alpha^{(m)} \log(1 + J_t^{(22)}) \\ &+ \beta^{(d)} \log C_t + \beta^{(w)} \log C_t^{(5)} + \beta^{(m)} \log C_t^{(22)} \\ &+ \gamma^{(d)} r_t^- + \gamma^{(w)} r_t^{(5)-} + \gamma^{(m)} r_t^{(22)-} + \varepsilon_t^{(h)}, \end{aligned} \quad (2.13)$$

We estimate model (2.13) by OLS with Newey-West covariance correction for serial correlation. In order to make multiperiod predictions we will estimate the model considering the aggregated dependent variable  $\log \widehat{V}_{t+h}^{(h)}$  with  $h$  ranging from 1 to 22 i.e. from one day to one month. While, strictly speaking, models with  $h > 1$  would require a cascade specification with longer frequencies multiple of  $h$ , for simplicity and comparison purposes, we will always retain the standard cascade specification with the three natural frequencies of one day, one week and one month. This can be viewed as a simplifying approximation justified by its empirically good performance.

## 3 Empirical evidence

The purpose of this section is to empirically analyze the main determinants of future asset volatility. Our data set covers a long time span of almost 18 years of high frequency data for the S&P 500 futures from the January 1982 to February 2009 (6,669 days). In order to mitigate the impact of microstructure effects on our estimates, the daily volatilities  $\widehat{V}_t$  are computed with the two-scales estimator proposed by Zhang et al. (2005).<sup>4</sup> Aït-Sahalia and Mancini (2008) show that using the two-scales estimator instead of standard realized volatility measures yields significant gains in volatility forecasting. The TBPV measure (2.10) for jump detection and estimation of  $C_t$  and  $J_t$  is computed at the sampling frequency of 5 minutes (corresponding to 84 returns per day).

### 3.1 In-sample analysis

The LHAR-CJ model is estimated using, as a dependent variable, realized volatility aggregated at different horizons. The results of the estimation of the LHAR-CJ when forecasting the S&P500 realized volatility at 1 day, 1 week, 2 weeks and 1 month are reported in Table 1, together with their statistical significance

---

<sup>4</sup>The two-scales estimator combines two realized volatilities computed at two different frequencies, where the slower one is computed by subsampling and averaging while the faster one (being a proxy for the variance of microstructure noise) is used for bias correction. In our implementation of the two-scales estimator we use a slower frequency of ten ticks returns.

S&amp;P500 LHAR in-sample regression, period 1982–2009

Variable	One day	One week	Two weeks	One month
$c$	0.442* (10.699)	0.549* (9.258)	0.662* (8.525)	0.858* (7.756)
$C$	0.307* (16.983)	0.201* (14.158)	0.154* (12.984)	0.116* (10.590)
$C^{(5)}$	0.369* (13.908)	0.359* (11.251)	0.332* (9.166)	0.286* (6.784)
$C^{(22)}$	0.222* (10.958)	0.319* (10.913)	0.370* (10.198)	0.415* (9.344)
$J$	0.043* (7.057)	0.020* (4.453)	0.017* (4.485)	0.012* (3.804)
$J^{(5)}$	0.011* (3.373)	0.013* (3.112)	0.011* (2.256)	0.010 (1.913)
$J^{(22)}$	0.005* (2.199)	0.008* (2.106)	0.010* (2.205)	0.014* (2.336)
$r^-$	-0.007* (-9.669)	-0.005* (-10.435)	-0.004* (-8.298)	-0.003* (-5.518)
$r^{(5)-}$	-0.008* (-4.412)	-0.006* (-3.059)	-0.008* (-4.012)	-0.007* (-3.472)
$r^{(22)-}$	-0.009* (-2.845)	-0.012* (-2.314)	-0.009 (-1.481)	-0.004 (-0.467)
$R^2$	0.7664	0.8137	0.8030	0.7629
HRMSE	0.2168	0.1692	0.1699	0.1796

Table 1: OLS estimate for LHAR-CJ regressions, model (2.13), for S&P500 futures from 28 April 1982 to 5 February 2009 (6,669 observations) excluding the October 1987 crash. The LHAR-CJ model is estimated on 1-day, 1-week, 2-week and 1-month realized volatility. The significant jump are computed using a critical value of  $\alpha = 99.9\%$ . Reported in parenthesis are  $t$ -statistics based on Newey-West correction.

### t-statistics of daily jump coefficients

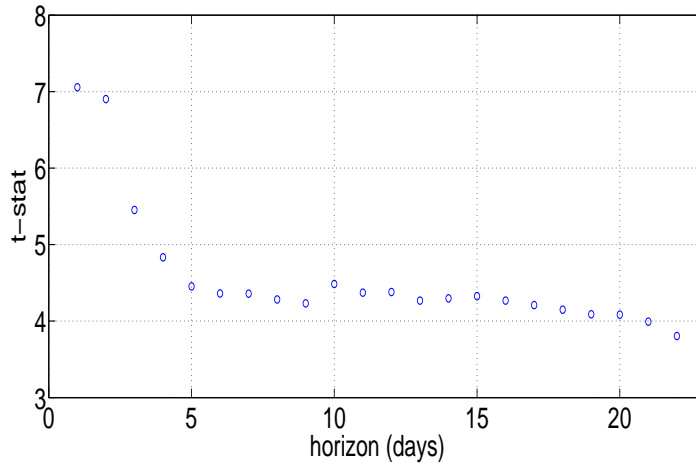


Figure 3: t-statistics of daily jump coefficients for LHAR-CJ model estimated on S&P500 from 28 April 1982 to 5 February 2009 (6,669 observations excluding the October 1987 crash) as a function of the forecasting horizon  $h$  .

evaluated with the Newey-West robust t-statistic. The forecasts of the different models are evaluated on the basis of the  $R^2$  of Mincer-Zarnowitz forecasting regressions, and the heteroskedasticity-adjusted root mean square error (HRMSE) proposed by Bollerslev and Ghysels (1996).

As usual, all the coefficients of the three continuous volatility components are positive and, in general, highly significant. It is however interesting to remark that the hierarchical asymmetric propagation of the volatility cascade presented in Section 2 is confirmed by the results. Indeed, the impact of daily and weekly volatility decreases with the forecasting horizon of future volatility, while the impact of monthly volatility increases. The coefficient which measures the impact of monthly volatility on future daily volatility is approximately double than that of daily volatility on future monthly volatility. This finding is consistent with Corsi (2009).

A similar hierarchical structure is present in the impact of jumps on future volatility. The daily and weekly jump components remain highly significant and positive especially for the shorter horizon realized volatility, and their impact declines with increasing horizon. The monthly jump component is also slightly significant at all horizons, but its impact increases with the horizon: the coefficient measuring the impact on monthly volatility being almost three times of that on daily volatility. Figure 3 plots the  $t$ -statistics of the impact of the daily jump on aggregated volatility at different time horizons, confirming, with its rapid decline, that daily jumps affects future volatilities much strongly over a short period of about one week, even if it remains highly significant at all the considered horizons.

The estimation of model (2.13) also reveals the strong significance (with an economically sound negative sign) of the negative returns at all the daily, weekly and monthly aggregation frequency, which unveils an

heterogeneous structure in the leverage effect as well. Not only daily negative returns affect the next day volatility (the well know leverage effect), but, in addition, also the negative returns of the past week and past month have an impact on forthcoming volatility. This finding suggests that the market aggregates daily, weekly and monthly memory, observing and reacting to price declines happened in the past week and month. To our knowledge, this is a novel empirical finding that further confirms the views of the Heterogeneous Market Hypothesis.

It is important to remark that the coefficients on the three components of continuous volatility and negative returns are roughly similar across the three frequencies, while the coefficients measuring the impact of jumps are decreasing with the frequency. This can be easily interpreted by looking again at Figure 1: the impact of jumps is significant till one month, but is decaying much faster than the impact of continuous volatility and negative returns. The fact that negative returns have a similar impact on future volatility across the three frequencies signals that the slowly decaying lagged correlation between negative returns and volatility is not only due to the high persistence of volatility, but to the high persistence of leverage effect as well.

In order to evaluate the relative contribution of the newly proposed model, we compare its in-sample prediction with those of the standard HAR model and the HAR-CJ model with jumps but with no leverage effects, as proposed by Andersen et al. (2007) but using threshold bipower variation as in Corsi et al. (2008). For each horizon, ranging from one day to one month, the forecasts are obtained by first estimating the parameters of the models on the full sample and then performing a series of static one-step-ahead forecasts.

The results, reported in Figures 4, which shows the  $R^2$  for different models at various horizons, and 5, which report Diebold-Mariano tests, show unambiguously that the inclusion of both the heterogeneous jumps and the heterogeneous leverage effects considerably improve the forecasting performance of the S&P 500 volatility at any forecasting horizon. In particular, the inclusion of heterogeneous leverage effect provides a relevant overall benefit in the in-sample performance. The fact that the HAR-CJ model does not outperform significantly the HAR model is not surprising since the difference between the two forecasts is mainly in the days which come after a jump, which are quite few.

### 3.2 Robustness to other volatility measures

In the literature many volatility measures have been proposed to better capture the dynamics of volatility. Forsberg and Ghysels (2007) proposed the use of realized absolute variation (RAV) which shows a more persistent dynamics than realized volatility being more robust to microstructure noise and jumps. The range has also been found to be significant by many authors, see e.g. Brandt and Jones (2006) and Engle and Gallo (2006). Motivated by the analysis of Bandi et al. (2008) who found liquidity to be a significant factor in asset pricing, we also compute the sum of squared tick-by-tick returns as a liquidity measure and employ it as a volatility factor. Recently, Barndorff-Nielsen et al. (2008) proposed the realized

### In-sample Mincer-Zarnowitz $R^2$

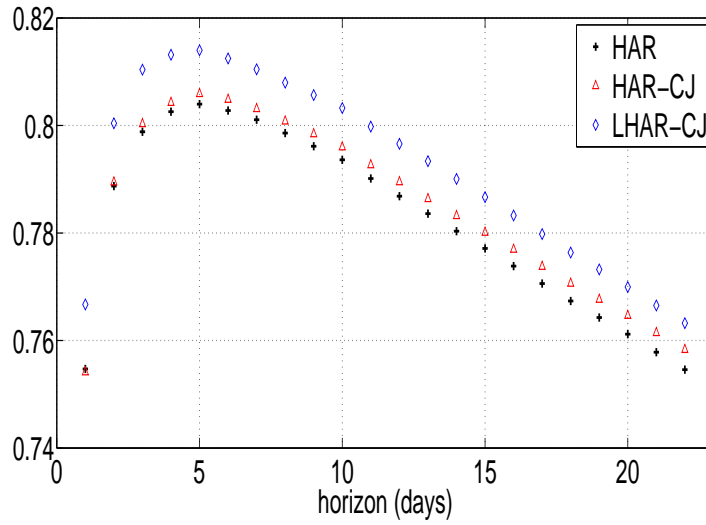


Figure 4:  $R^2$  of Mincer-Zarnowitz regressions for static in sample one-step ahead forecasts for horizons ranging from 1 day to 1 month of the S&P500 from April 1982 to February 2009 (6,669 observations). The forecasting models are the standard HAR with only heterogeneous volatility, the HAR-CJ with heterogeneous jumps and the LHAR-CJ model.

### In-sample Diebold-Mariano test

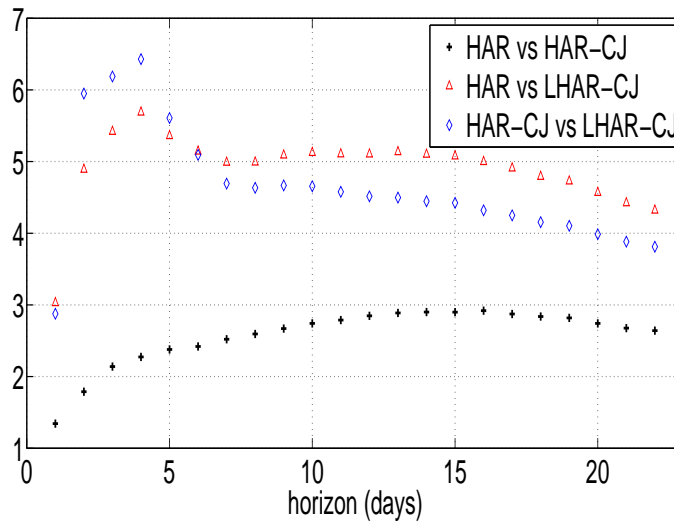


Figure 5: Diebold-Mariano test for the HRMSE in predicting the square root of  $\widehat{V}$  for horizons ranging from 1 day to 1 month of the S&P500 from April 1982 to February 2009 (6,669 observations). The forecasting models are the standard HAR with only heterogeneous volatility, the HAR-CJ with heterogeneous jumps and the LHAR-CJ model.

semivariance as the sum of square negative returns to capture the impact on volatility of downward price pressures. Visser (2008) combines RAV and semivariance by taking the sum of negative absolute squared returns.

In the spirit of Forsberg and Ghysels (2007), we compare the relative explanatory power of different volatility measures by estimating the following set of models (for space concerns we limit ourselves to the one day horizon) obtained by adding explanatory variables to model (2.13).

Some of the additional measures turn out to be fairly related to the daily continuous volatility and to the daily jump (such as range and semivariance). For those measures we also estimate models where the daily jump regressor is removed so that a direct performance comparison with the LHAR-CJ is possible. Estimation results are reported in Table 2.

The liquidity proxy (LQ) turn out to be not significant when included in the LHAR-CJ model. In line with previous literature, we find that the realized absolute variation (RAV) computed at 5-minute frequency and the range have a significant impact on future volatility. However, they seems to be mainly substitutes for continuous volatility and jumps, which is not totally surprising since they are estimators (though noisy) of total quadratic variation. Indeed, for instance, when the range replaces the jumps (LHAR-Range model, not reported), the coefficients of daily continuous volatility almost halves. The  $R^2$  of the two competing regressions (LHAR-Range and LHAR-CJ) is practically the same. When the range is inserted together with the jumps (LHAR-CJ-Range), both the coefficients of daily volatility and jumps decrease, although they remain highly significant. While, the significance of the heterogeneous leverage effect is untouched by the presence of the RAV and the range. The  $R^2$  of the encompassing regression increases marginally. We thus conclude that the RAV and the range, while partially proxying for both volatility and jump, are also able to capture (especially the range) some other effect which is not captured by the other variables in the model. We found similar results for the realized semivariance (semiRV) of Barndorff-Nielsen et al. (2008) and the downward absolute power variation of Visser (2008) (semiRAV). Realized semivariance and semi-power-variation are significant in explaining future volatility, and, again, they seems very correlated with both the daily two-scales estimator and the jumps (typically depleting the significance of the corresponding coefficients without totally removing it), while unrelated with the leverage. However, their contribution to the model performance is no significant (as measured by the Diebold-Mariano test). Moreover, when they are included in the all encompassing model they both remain insignificant.

Summarizing, the results of this section show that when the other volatility measures proposed in the literature are inserted in the baseline LHAR-CJ model they either not contribute significantly or only marginally contribute to the performance of the model. Moreover, they mainly act as substitutes of continuous volatility and jumps. Hence, the LHAR-CJ model seems to capture the main determinants of volatility dynamics.

### 3.3 Out-of-sample analysis

In this section, we evaluate the performance of the LHAR-CJ model on the basis of a genuine out-of-sample analysis. For the out-of-sample forecast of  $\widehat{V}_t$  on the  $[t, t+h]$  interval we keep the same forecasting horizons ranging from one day to one month and re-estimate the model at each day  $t$  on an increasing window of all the observation available up to time  $t-1$ . The out of sample forecasting performance for the square root of  $\widehat{V}$  in terms of Mincer-Zarnowitz  $R^2$  is reported in Figure 7. The superiority of the LHAR-CJ model at all horizons is confirmed, validating the importance of including both the heterogeneous leverage effects and jumps in the forecasting model.

Finally, it is important to note that the inclusion of the jump component helps also in forecasting longer horizon volatility. To clarify this issue, we perform a Monte Carlo simulation analysis described in the following section.

## 4 A simulation study

We evaluate our empirical results for jumps through the lens of a Monte Carlo simulations. We simulate the stock index price with the flexible specification of Eraker et al. (2003), that is:

$$\begin{pmatrix} dY_t \\ dV_t \end{pmatrix} = \begin{pmatrix} \mu \\ \kappa(\theta - V_{t-}) \end{pmatrix} dt + \sqrt{V_{t-}} \begin{pmatrix} 1 & 0 \\ \sigma_v \rho & \sigma_v \sqrt{1 - \rho^2} \end{pmatrix} dW_t + \begin{pmatrix} \xi^y dN_t^y \\ \xi^v dN_t^v \end{pmatrix} \quad (4.1)$$

where  $W_t$  is a bidimensional Brownian motion and  $dN^y$  and  $dN^v$  are Poisson processes with intensity  $\lambda^y$  and  $\lambda^v$  respectively;  $\xi^y$  is normally distributed, while  $\xi^v$  has an exponential law. As in Eraker et al. (2003), we consider two cases: the case in which  $dN^y$  is independent from  $dN^v$  (what they name the SVIJ model) and the case in which  $dN^y = dN^v$  (what they name the SVCJ model), and we hold their terminology. We use exactly the parameters estimated by Eraker et al. (2003) for the S&P500 time series.

Figure 8 and 9 report the results. Explicitly including the jump component has a direct benefit for both the SVIJ (independent jumps) and the SVCJ (contemporaneous jumps) specification. In the SVIJ case, jumps has no impact on future volatility, but there is still a benefit in removing the jump component. Indeed, in this model the persistence is conveyed only by the continuous volatility, while total quadratic variation (which is estimated by realized volatility) also contain the memoryless jumps. Thus, by separating the jumps from the persistent part in the explanatory variables, a better model specification is obtained. We conclude that, when the memory of volatility is mainly contained in the continuous part of quadratic variation, there is still a potential benefit in removing jumps even if they do not impact on future volatility. This benefit persist also for long horizon forecasts. Importantly, in this case, the jump component is found to be insignificant.

When a jump occurs in price, it also occurs in volatility and it is positive. Thus, when there is a jump in

### Out-of-sample Mincer-Zarnowitz $R^2$

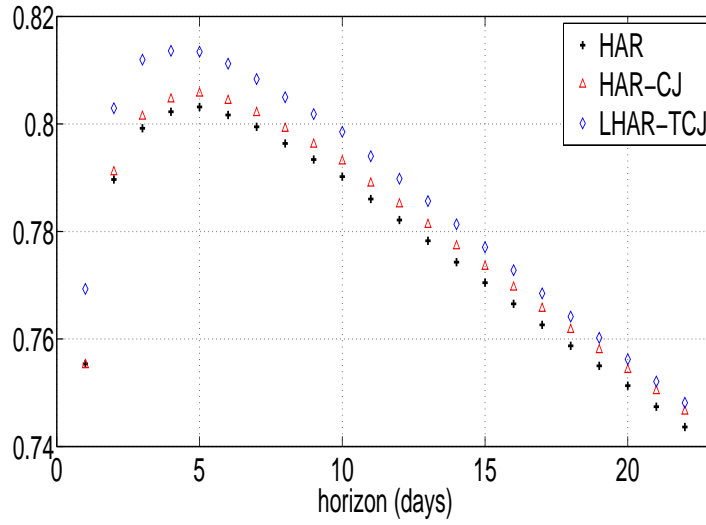


Figure 6:  $R^2$  of Mincer-Zarnowitz regressions for out-of-sample forecasts for horizons ranging from 1 day to 1 month of the S&P500 from January 1982 to February 2009 (6,669 observations, the first 2000 observation are used to initialize the models). The forecasting models are the standard HAR with only heterogeneous volatility, the HAR-CJ with heterogeneous jumps and the LHAR-CJ model.

### Out-of-sample Diebold-Mariano test

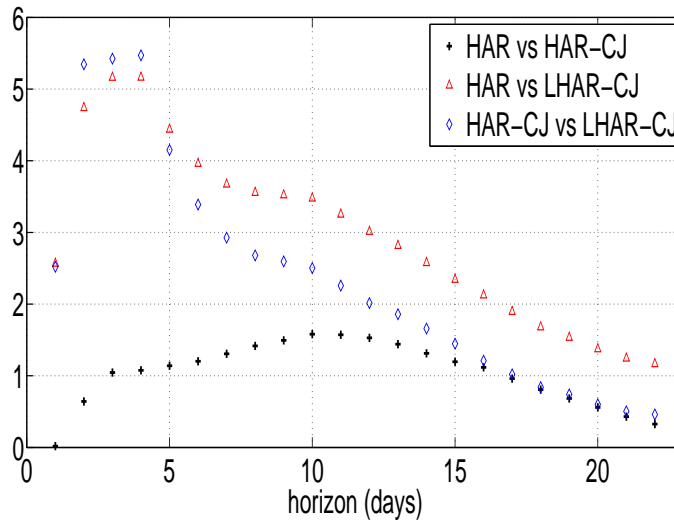


Figure 7: Diebold-Mariano test for the out-of-sample RMSE in predicting the square root of  $\widehat{V}$  for horizons ranging from 1 day to 1 month of the S&P500 from January 1982 to February 2009 (6,669 observations, the first 2000 observation are used to initialize the models). The forecasting models are the standard HAR with only heterogeneous volatility, the HAR-CJ with heterogeneous jumps and the LHAR-CJ model.

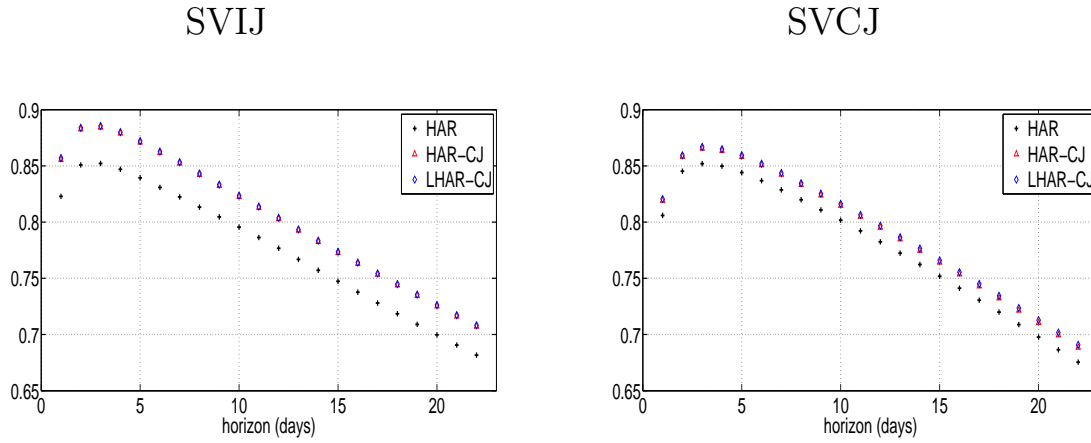


Figure 8:  $R^2$  of Mincer-Zarnowitz regressions for realized volatility forecast ranging from 1 day to 1 month of 4000 days simulated data from SVIJ (left) and SVCJ(right) model. The forecasting models are the standard HAR with only heterogeneous volatility, the HAR-CJ with heterogeneous jumps and the LHAR-CJ model.

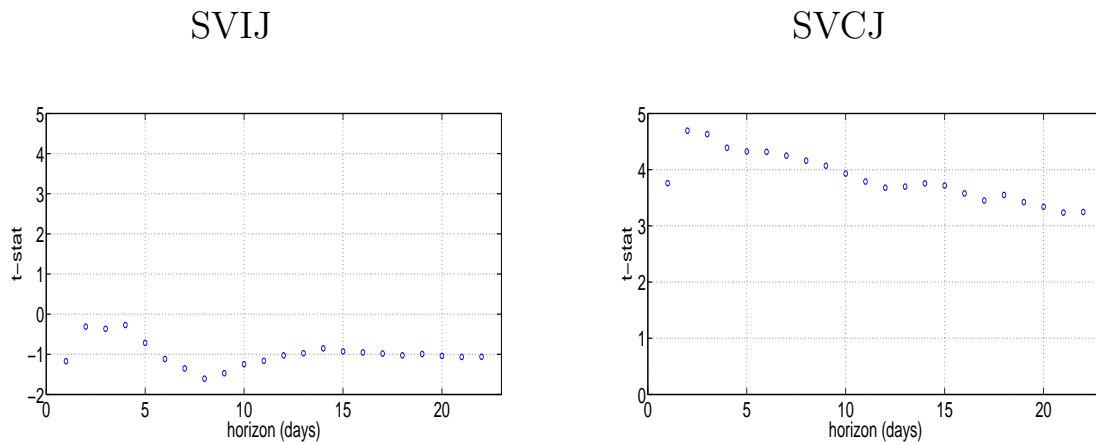


Figure 9: t-statistics of daily jump coefficients for LHAR-CJ model estimated on 4000 simulated daily data from SVIJ (left) and SVCJ (right) model as a function of the forecasting horizon  $h$ .

price, volatility becomes higher and it stays higher because of its memory persistence. That is why jumps are (contrary to the SVIJ case) found to be significant in explaining future volatility in SVCJ models. Hence, our simulation results show that a possible mechanism explaining the significant impact of jumps on future volatility is given by contemporaneous jumps in price and volatility. Moreover, it confirms that the significance of the jumps coefficients on our empirical analysis provides an indication of the presence of a genuine forecasting power of jumps on future volatility. Hence, the similarity between the figures reporting the Newey-West corrected t-statistics of the daily jump coefficient estimated on the simulated SVCJ model (Figure 9 right panel) and on the empirical S&P500 (Figure 3), confirms the presence of a genuine forecasting power of jumps on future S&P volatility.

On the other hand, the heterogeneous leverage effect found in real data cannot be completely explained by model 4.1. Indeed, the presence of a negative coefficient  $\rho \approx -0.5$  (estimated on S&P 500 data) is able to explain only short-period leverage effect, by propagating negative returns into contemporaneous, and by memory persistence, future volatility. While, in the real data, we provided evidence for strong heterogeneous leverage effect, being also the weekly and monthly negative components highly significant. The model specification 4.1 is then insufficient to explain our results which demand for a more complicated continuous process with a richer specification.

## 5 Conclusions

This paper presents a new model for volatility forecasting which extends the HAR model by isolating three main determinants of volatility dynamics, namely heterogeneous lagged continuous volatility, heterogeneous lagged negative returns and heterogeneous lagged jumps. We find that each component plays a different role at different forecasting horizons, but all the three are highly significant and neglecting each one of them is detrimental to the forecasting performance of the model. Moreover, when other volatility measures proposed in the literature are inserted in the LHAR-CJ model they either drop out or only marginally contribute to the performance of the model confirming the ability of the LHAR-CJ model to capture the main determinants of volatility dynamics.

Explicitly modelling the jump component is important for two distinct reasons. First, it has a trimming effect on the dynamics of the persistent component of volatility which allows a better prediction of future volatility, confirming Andersen et al. (2007). Secondly, as suggested in Corsi et al. (2008), they have a direct positive and significant impact on future volatility. Moreover, there are evidences that this direct impact of jumps is of a short lived nature. If, as it seems reasonable, volatility is a measure of the uncertainty of the market about its fundamental values, our findings can be interpreted as follows: after a jump (usually a market crash) the market takes a longer period to reassess its fundamental value by dissipating the uncertainty created by the jumps; during this period residual uncertainty generate higher volatility. Our simulated experiments indicated that this mechanism can be statistically reproduced by

a model having contemporaneous jumps in price and volatility.

On the other hand, while the mechanism of leverage effects on volatility dynamics is still not well understood, we find that not only daily but also weekly and monthly negative past returns are highly significant and have a significant forecasting power on future volatility, suggesting a slowly decaying impact of negative returns similar to that of continuous volatility. This novel effect seems to confirm the heterogeneous structure of the market and cannot be explained by continuous-time models, though flexible, as the ones specified so far in the literature.

We conclude by noting that our model is very simple to implement, as it does not require sophisticated computational technique. The estimation of the model parameters can be performed through a simple OLS regression, and the computation of the explanatory variables is trivial. We think that, for all the aforementioned reasons, this model we proposed may be effectively used for risk management.

S&P500 in-sample estimates, period 1982–2009

Variable	LHAR- CJ	LHAR- CJ-LQ	LHAR- CJ-RAV	LHAR- CJ- Range	LHAR- CJ- SemiRV	LHAR- CJ- SemiRAV	LHAR- CJ-All
const	0.442* (10.699)	0.432* (9.968)	0.593* (12.419)	0.444* (10.860)	0.470* (11.029)	0.554* (9.869)	0.616* (8.914)
C	0.307* (16.983)	0.301* (14.672)	0.116* (3.355)	0.207* (10.035)	0.249* (9.588)	0.252* (9.489)	0.085* (2.449)
C <sup>(5)</sup>	0.369* (13.908)	0.368* (13.786)	0.374* (14.125)	0.384* (14.568)	0.372* (14.074)	0.372* (14.032)	0.389* (14.693)
C <sup>(22)</sup>	0.222* (10.958)	0.221* (10.941)	0.221* (10.943)	0.223* (11.071)	0.223* (11.003)	0.222* (10.963)	0.223* (11.035)
J	0.043* (7.057)	0.042* (6.912)	0.018* (2.519)	0.024* (3.893)	0.033* (4.850)	0.037* (5.620)	0.010 (1.331)
J <sup>(5)</sup>	0.011* (3.373)	0.011* (3.351)	0.012* (3.717)	0.012* (3.789)	0.011* (3.409)	0.011* (3.413)	0.012* (3.959)
J <sup>(22)</sup>	0.005* (2.199)	0.005* (2.228)	0.006* (2.334)	0.005* (2.207)	0.006* (2.266)	0.006* (2.277)	0.006* (2.329)
r <sup>-</sup>	-0.007* (-9.669)	-0.007* (-9.705)	-0.007* (-10.156)	-0.006* (-8.193)	-0.006* (-8.197)	-0.006* (-7.792)	-0.005* (-5.630)
r <sup>(5)-</sup>	-0.008* (-4.412)	-0.008* (-4.382)	-0.007* (-4.147)	-0.008* (-4.847)	-0.008* (-4.368)	-0.008* (-4.278)	-0.008* (-4.582)
r <sup>(22)-</sup>	-0.009* (-2.845)	-0.009* (-2.845)	-0.009* (-2.687)	-0.009* (-2.868)	-0.010* (-2.980)	-0.010* (-2.951)	-0.010* (-2.853)
LQ		0.009 (0.555)					
RAV			0.185* (6.533)				0.077 (1.867)
Range				0.088* (9.364)			0.086* (7.911)
SemiRV					0.058* (3.305)		-0.011 (-0.330)
SemiRAV						0.054* (3.141)	0.056 (1.455)
R <sup>2</sup>	0.7664	0.7664	0.7681	0.7696	0.7668	0.7668	0.7704
HRMSE	0.2168	0.2168	0.2158*	0.2148*	0.2165	0.2165	0.2142*
DM		(-0.1961)	(2.7060)	(4.6323)	(1.5020)	(1.8539)	(4.6778)

Table 2: Estimated parameters, Mincer-Zarnowitz  $R^2$ , and Heteroskedasticity adjusted RMSE (HRMSE), of alternative specifications of the baseline LHAR-CJ model; t-statistic and Diebold-Mariano test for HRMSE are in parenthesis.

SVIJ regression					SVCJ regression				
	1 day	1 week	2 weeks	1 month		1 day	1 week	2 weeks	1 month
$c$	0.235*	0.441*	0.678*	1.220*	$c$	0.291*	0.521*	0.814*	1.476*
	(6.325)	(6.816)	(6.885)	(7.301)		(6.364)	(7.301)	(7.817)	(8.131)
$C$	0.638*	0.594*	0.549*	0.479*	$C$	0.440*	0.429*	0.398*	0.370*
	(16.288)	(15.437)	(14.403)	(14.852)		(12.103)	(11.229)	(10.495)	(10.797)
$C^{(5)}$	0.341*	0.340*	0.345*	0.339*	$C^{(5)}$	0.537*	0.540*	0.555*	0.516*
	(7.861)	(7.691)	(6.817)	(6.022)		(13.527)	(10.555)	(9.306)	(7.375)
$C^{(22)}$	-0.024	-0.017	-0.022	-0.047	$C^{(22)}$	-0.036*	-0.074*	-0.117*	-0.182*
	(-1.516)	(-0.746)	(-0.595)	(-0.836)		(-2.048)	(-2.453)	(-2.811)	(-2.880)
$J$	-0.012	-0.006	-0.011	-0.009	$J$	0.042*	0.044*	0.040*	0.038*
	(-1.174)	(-0.713)	(-1.249)	(-1.062)		(3.848)	(4.288)	(3.924)	(3.154)
$J^{(5)}$	-0.001	-0.001	0.014*	0.012	$J^{(5)}$	0.019*	0.019*	0.023*	0.022*
	(-0.296)	(-0.120)	(2.184)	(1.376)		(4.922)	(4.306)	(3.928)	(3.487)
$J^{(22)}$	-0.000	-0.001	-0.004	-0.009	$J^{(22)}$	0.006*	0.007	0.008	0.010
	(-0.206)	(-0.235)	(-0.694)	(-1.118)		(2.080)	(1.720)	(1.363)	(1.071)
$r^-$	-0.052*	-0.038*	-0.039*	-0.041*	$r^-$	-0.042*	-0.040*	-0.038*	-0.038*
	(-5.444)	(-5.417)	(-5.496)	(-5.224)		(-4.277)	(-5.308)	(-4.829)	(-4.347)
$r^{(5)-}$	-0.016	-0.031	0.001	0.018	$r^{(5)-}$	-0.029	-0.024	-0.038	-0.065
	(-0.994)	(-1.474)	(0.048)	(0.598)		(-1.388)	(-0.893)	(-1.390)	(-1.754)
$r^{(22)-}$	0.040	0.057	0.033	0.002	$r^{(22)-}$	0.079	0.053	0.025	-0.002
	(1.651)	(1.603)	(0.653)	(0.029)		(1.569)	(0.835)	(0.320)	(-0.015)
$R^2$	0.8567	0.8718	0.8232	0.7076	$R^2$	0.8197	0.8587	0.8152	0.6895
HRMSE	0.1325	0.1448	0.1661	0.2033	HRMSE	0.1254	0.1298	0.1460	0.1790

Table 3: OLS estimate for baseline LHAR-CJ model, for S&P500 futures from January 1990 to December 2007 (4344 observations). The LHAR-CJ model is estimated on 1-day, 1-week, 2-week and 1-month realized volatility. The significant jumps are computed using a critical value of  $\alpha = 99.9\%$ . Reported in parenthesis are  $t$ -statistics based on Newey-West correction.

## References

- Ait-Sahalia, Y. and L. Mancini (2008). Out of sample forecasts of quadratic variation. *Journal of Econometrics* 147(1), 17–33.
- Alizadeh, S., M. Brandt, and F. Diebold (2002). High and low frequency exchange rate volatility dynamics: Range-based estimation of stochastic volatility models. *Journal of Finance* 57, 1047–1092.
- Andersen, T., T. Bollerslev, and F. X. Diebold (2003). Parametric and nonparametric volatility measurement. In L. P. Hansen and Y. Ait-Sahalia (Eds.), *Handbook of Financial Econometrics*. Amsterdam: North-Holland.
- Andersen, T., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics* 89, 701–720.
- Arneodo, A., J. Muzy, and D. Sornette (1998). Casual cascade in stock market from the "infrared" to the "ultraviolet". *European Physical Journal B* 2, 277.
- Bandi, F., C. Moise, and J. Russell (2008). Market volatility, market frictions, and the cross-section of stock returns. Working paper.
- Banerjee, A. and G. Urga (2005). Modelling structural breaks, long memory and stock market volatility: an overview. *Journal of Econometrics* 129(1-2), 1–34.
- Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*. Forthcoming.
- Barndorff-Nielsen, O., S. Kinnebrock, and N. Shephard (2008). Measuring downside risk – realised semivariance. Working paper.
- Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* 2, 1–48.
- Bianco, S., F. Corsi, and R. Renò (2009). Intraday LeBaron Effects. *Proceedings of the National Academy of Science of the USA*. Forthcoming.
- Bollerslev, T. and E. Ghysels (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics* 14(2), 139–151.
- Bollerslev, T., J. Litvinova, and G. Tauchen (2006). Leverage and Volatility Feedback Effects in High-Frequency Data. *Journal of Financial Econometrics* 4(3), 353.
- Bollerslev, T., G. Tauchen, and H. Zhou (2008). Expected stock returns and variance risk premia. Working Paper.
- Boudt, K., C. Croux, and S. Laurent (2008). Outlyingness Weighted Quadratic Covariation. Working Paper.
- Brandt, M. and C. Jones (2006). Volatility Forecasting With Range-Based EGARCH Models. *Journal of Business and Economic Statistics* 24(4), 470.
- Campbell, J. Y. and L. Hentschel (1992). No news is good news: a asymmetric model of changing volatility in stock returns. *Journal of Financial Economics* 31, 281–318.
- Christensen, K., R. Oomen, and M. Podoslki (2008). Realised quantile-based estimation of the integrated variance. Working paper.
- Christie, A. (1982). The stochastic behavior of common stock variances: value, leverage and interest rate effects. *Journal of Financial Economics* 10, 407–432.
- Clements, M., A. Galvão, and J. Kim (2008). Quantile Forecasts of Daily Exchange Rate Returns from Forecasts of Realized Volatility. *Journal of Empirical Finance*. Forthcoming.
- Corsi, F. (2009). A simple approximate long-memory model of realized-volatility. *Journal of Financial Econo-*

- metrics* 7, 174–196.
- Corsi, F., S. Mittnik, C. Pigorsch, and U. Pigorsch (2008). The volatility of realized volatility. *Econometric Reviews* 27(1-3), 1–33.
- Corsi, F., D. Pirino, and R. Renò (2008). Volatility forecasting: the jumps do matter. Working paper.
- Engle, R. and G. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics* 131(1-2), 3–27.
- Eraker, B., M. Johannes, and N. Polson (2003). The impact of jumps in equity index volatility and returns. *Journal of Finance* 58, 1269–1300.
- Forsberg, L. and E. Ghysels (2007). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics* 5, 31–67.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131(1-2), 59–95.
- Glosten, L., R. Jagannathan, and D. Runkle (1989). On the relation between the expected value of the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–1801.
- Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics* 3(4), 456–499.
- Jacod, J., Y. Li, P. Mykland, M. Podolskij, and M. Vetter (2007). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and Their Applications*. Forthcoming.
- Jiang, G. and R. Oomen (2008). Testing for jumps when asset prices are observed with noise—a “swap variance” approach. *Journal of Econometrics* 144(2), 352–370.
- Lee, S. and P. Mykland (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial studies* 21(6), 2535.
- Lynch, P. and G. Zumbach (2003). Market heterogeneities and the causal structure of volatility. *Quantitative Finance* 3(4), 320–331.
- Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics* 36(2), 270–296.
- Muller, U., M. Dacorogna, R. Davé, R. Olsen, O. Pictet, and J. von Weizsacker (1997). Volatilities of different time resolutions - analyzing the dynamics of market components. *Journal of Empirical Finance* 4, 213–239.
- Visser, M. (2008). Forecasting S&P 500 Daily Volatility using a Proxy for Downward Price Pressure. Working Paper.
- Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100, 1394–1411.