Bulletin of Economic Research 68:2, 2016, 0307-3378

ALLOWING FOR JUMP MEASUREMENTS IN VOLATILITY: A HIGH-FREQUENCY FINANCIAL DATA ANALYSIS OF INDIVIDUAL STOCKS

Vassilios G. Papavassiliou
University College Dublin, School of Business, Blackrock, Co Dublin, Ireland
Rimini Centre for Economic Analysis (RCEA), Via Patara, Rimini, Italy

ABSTRACT

Following recent advances in the non-parametric realized volatility approach, we separately measure the discontinuous jump part of the quadratic variation process for individual stocks and incorporate it into heterogeneous autoregressive volatility models. We analyse the distributional properties of the jump measures vis-à-vis the corresponding realized volatility ones, and compare them to those of aggregate US market index series. We also demonstrate important gains in the forecasting accuracy of high-frequency volatility models.

Keywords: HAR-RV model, high-frequency data, realized volatility, volatility jumps

JEL classification numbers: C1, C14, C53, C58, G1

I. INTRODUCTION

The modelling of financial markets based on the assumption that the price series is a continuous-time diffusive process, is of central importance to finance. As Huang and Tauchen (2005) assert, there remains the open issue of whether diffusive models are consistent with the extreme movements seen in financial price series, and whether jump diffusions provide a more appropriate empirical model for financial volatility. Barndorff-Nielsen and Shephard (2003, 2004) develop a powerful toolkit for detecting the presence of jumps in higher frequency financial time series. Andersen *et al*. (2007) build on the theoretical results in Barndorff-Nielsen and Shephard involving the so-called bi-power variation measures constructed from the summation of appropriately scaled cross-products of adjacent high-frequency absolute returns. They develop a non-parametric procedure for separately measuring the continuous sample path variation and the discontinuous jump part of the quadratic variation process.

The majority of earlier findings were made exclusively for indices and there remains the open question whether these findings carry over to individual stocks. Rodrigues and Schlag (2009) provide ample evidence that the stylized facts found in studies on indices do not remain valid for individual stocks. The average size of jumps in the stock price is smaller in absolute terms

Correspondence: Vassilios G. Papavassiliou, Athens University of Economics and Business, Department of Management Science and Technology, 47A Evelpidon & 33 Lefkados, 104 34, Athens, Greece. Email: vpapavass@aueb.gr, Tel: +30 210 8203 670, Fax: +30 210 8828 078.

than for the index. Also, the frequency of jumps in prices is higher for individual stocks than for indices, in fact jumps tend to occur more than twice as often for the average stock. Also, there is a much less negative correlation between stock returns and volatility changes at the individual stock level than for the index.

Along the lines of empirical option pricing, Bollen and Whaley (2004) document that the implied volatility curves for stock market indices tend to be negatively sloped and much steeper than those for the component stocks. This finding emphasizes that results from the analyses of market indices cannot simply be extended to single stocks. Campbell *et al*. (2001) report that the variance at the individual firm-level has more than doubled between 1962 and 1997 whereas variances at the market level have remained fairly stable over that period.

To the best of our knowledge our paper represents the first study of volatility modelling with jumps for individual stocks listed on the Greek stock exchange. We verify that models estimated for indices are compatible to those estimated for individual stocks and that the appropriateness of jump diffusions is not restricted only to broad stock market indices. We also show that estimating price volatility using high-frequency data offers large gains in the forecasting accuracy of heterogeneous autoregressive models through the use of realized volatilities.

The rest of this paper is organized as follows. Section II describes the methodology and the data set. Section III presents the empirical results. Section IV concludes the paper.

II. METHODOLOGY AND THE DATA

Following Andersen *et al*. (2007) we estimate the contribution to the quadratic variation process due to jumps as follows. Assume that the logarithmic price p_t of a financial asset follows the continuous-time semimartingale jump diffusion process:

$$dp\,(t) = \mu\,(t)\,dt + \sigma\,(t)\,dW\,(t) + \kappa\,(t)\,dq\,(t)\,, \qquad 0 \leq t \leq T. \tag{1}$$

where $\mu\,(t)$ is the continuous mean process, the stochastic volatility process $\sigma\,(t)$ is positive and caglad, $W\,(t)$ denotes a standard Brownian Motion, $dq\,(t)$ is a counting process with $dq\,(t) = 1$ corresponding to a jump at time t and $dq\,(t) = 0$ otherwise with jump intensity $\lambda\,(t)$, and $\kappa\,(t)$ refers to the size of the corresponding jumps. The quadratic variation for the cumulative return process, $r\,(t) \equiv p\,(t) - p\,(0)$, is given by:

$$[r, r]_t = \int_0^t \sigma^2\,(s)\,ds + \sum_{0<s\leq t} \kappa^2\,(s) \tag{2}$$

In the absence of jumps the quadratic variation equals the integrated volatility. In order to separately identify the two components in equation (2), Andersen *et al*. (2007) utilize a new non-parametric high-frequency approach. Let the daily returns denoted by:

$$r_t = p_t - p_{t-1}, \quad t = 1, \ldots. \tag{3}$$

The daily realized variance is defined as the summation of intradaily squared returns:

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2 \tag{4}$$

By the theory of quadratic variation the realized variance converges uniformly in probability to the quadratic variation as the sampling frequency of the underlying returns approaches infinity:

$$RV_t \rightarrow \int_{t-1}^{t} \sigma^2(s)\,ds + \sum_{j=N(t-1)+1}^{N(t)} \kappa^2(s_j) \tag{5}$$

In other words, the realized variance affords an ex-post measure of the true price variation, including the discontinuous jump part (Bollerslev *et al.*, 2009). For the separate identification of the continuous variation from the jump component, Barndorff-Nielsen and Shephard (2004) propose the so-called Bipower variation measure, defined by:

$$BV_t = \frac{\pi}{2}\sum_{j=2}^{M} |r_{t,j}||r_{t,j-1}| \tag{6}$$

For increasingly finely sampled returns the Bipower variation measure becomes immune to jumps and, for increasing values of M, estimates the integrated variance:

$$BV_t \rightarrow \int_{t-1}^{t} \sigma^2(s)ds \tag{7}$$

Combining equations (5) and (7), the contribution to the quadratic variation process due to jumps in the underlying process, may be estimated by:

$$RV_t - BV_t \rightarrow \sum_{j=N(t-1)+1}^{N(t)} \kappa^2(s_j) \tag{8}$$

The jump (J_t) measure is in theory restricted to be non-negative, however, in practice BV_t may exceed RV_t so that J_t becomes negative. Andersen *et al.* (2007) impose a non-negativity truncation on the actual empirical jump measurements, $J_t \equiv \max[RV_t - BV_t, 0]$, which we also follow in this study.

The results of Barndorff-Nielsen and Shephard are asymptotics. Realized volatility has a considerable statistical error which can be reduced by estimating returns over short time intervals (Dacorogna *et al.*, 2001). However, the choice of small return intervals leads to a bias caused by microstructure effects. Indeed, in actual empirical applications market microstructure effects such as bid-ask bounce and infrequent trading put a limit on the number of return observations that may be used productively in the computation of the various realized volatility specifications. Thus, realized volatility is subject to a finite-sample measurement error vis-à-vis the true integrated volatility. This observation motivated Barndorff-Nielsen and Shephard (2002) to develop their asymptotic theory.

Briefly, assuming that the mean and volatility processes are jointly independent, it follows from Barndorff-Nielsen and Shephard (2002) that,

$$z_t \equiv \sqrt{h^{-1}}\frac{U_t(h)}{\sqrt{2IQ_t}} N(0,1) \tag{9}$$

where $U_t(h) \equiv RV_t(h) - IV_t$ and the integrated quarticity, IQ_t, is defined by $IQ_t \equiv \int_{t-1}^{t} \sigma_u^4 du$. It follows under the same assumptions that, the integrated quarticity may be estimated by the standardized realized quarticity

$$RQ_t(h) \equiv \frac{1}{h}\frac{1}{3}\sum_{i=1}^{1/h} r_{t-1+ih}^{(h)4} \tag{10}$$

The aforementioned asymptotics allow for model-free approximations to the distribution of the realized volatility error and justify the use of high-frequency realized volatility as an unbiased and convenient measure.

In a second step, we assess the added value of separately measuring the jump component in forecasting volatilities. We employ the simple Heterogeneous Autoregressive model of the Realized Volatility (HAR-RV) proposed by Corsi (2009). The daily HAR-RV model is expressed as:

$$RV_{t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \varepsilon_{t+1} \tag{11}$$

where RV_t, $RV_{t-5,t}$, $RV_{t-22,t}$ denote the daily, weekly and monthly realized volatilities, respectively.[1] Andersen *et al.* (2007) propose the new HAR-RV-J model, in which various specifications of the jump measurements J_t, $J_t^{1/2}$, $\log(J_t + 1)$ are included as additional explanatory variables:

$$RV_{t,t+h} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \beta_J J_t + \varepsilon_{t,t+h} \tag{12}$$

$$(RV_{t,t+h})^{1/2} = \beta_0 + \beta_D RV_t^{1/2} + \beta_W (RV_{t-5,t})^{1/2} + \beta_M (RV_{t-22,t})^{1/2} + \beta_J J_t^{1/2} + \varepsilon_{t,t+h} \tag{13}$$

$$\begin{aligned}\log(RV_{t,t+h}) = \beta_0 &+ \beta_D \log(RV_t) + \beta_W \log(RV_{t-5,t}) + \beta_M \log(RV_{t-22,t}) \\ &+ \beta_J \log(J_t + 1) + \varepsilon_{t,t+h}\end{aligned} \tag{14}$$

Similar to Andersen *et al.* (2007), we construct 'standard' HAR models in which the jump component is absent and the realized volatilities on the right-hand side of (11) are replaced by the corresponding lagged squared daily, weekly, and monthly returns.[2] Then, the resulting R^2's from the 'standard' HAR and HAR-RV-J models are compared, in an effort to highlight the gains afforded by the use of high-frequency data.

We utilize a sub-sample of the data set used by Papavassiliou (2012). It is based on quotation data, time-stamped to the nearest second, of the two most heavily traded and continuously listed blue-chip stocks of the FTSE/ATHEX20 index. The sample period extends from 23 September 2003 through to 31 March 2006, totalling 635 trading days. The stocks are: Alpha Bank (ticker symbol: ALPHA) and National Bank of Greece (ticker symbol: ETE). All pre-sessional quotations have been excluded, as well as zero and negative spreads. We employ midpoints of bid-ask quotes as price measures which are generally less noisy measures of the efficient price than are transaction prices (Bandi and Russell, 2006). Daily realized variance is computed using summed 5-minute squared returns. The 5-minute intraday returns are constructed from the linearly interpolated logarithmic midpoint of the continuously recorded bid and ask quotes.

III. EMPIRICAL RESULTS

Table 1 summarizes the distributional properties of the jump measures and those of the realized daily volatilities. It is evident that the skewness and kurtosis values for all jump measures are larger than the corresponding statistics of the realized volatility measures. The realized logarithmic standard deviations are closer to being normally distributed than the realized variance

[1]Multi-period realized volatilities are defined as the normalized sum of the one-period volatilities, $RV_{t,t+h} = h^{-1}[RV_{t+1} + RV_{t+2} + \cdots + RV_{t+h}]$.

[2]Weekly and monthly squared returns are calculated by summing up the one-period daily squared returns

TABLE 1
Distributional properties of daily jump measures and volatilities

	J_t	$J_t^{1/2}$	$\log(J_t+1)$	$\hat{\sigma}_t^2$	$\hat{\sigma}_t$	$\log(\hat{\sigma}_t)$
			ALPHA			
Mean	5.33E-05	0.005	0.006	0.0002	0.013	−4.426
S.D.	0.0001	0.005	0.004	0.0002	0.005	0.370
Skewness	6.748	2.813	2.773	3.997	1.603	0.174
Kurtosis	63.408	14.882	14.543	28.670	7.842	3.158
Q_{10}	452.11	540.58	539.26	496.07	570.96	500.51
			ETE			
Mean	5.30E-05	0.006	0.006	0.0002	0.013	−4.417
S.D.	0.0001	0.004	0.004	0.0002	0.005	0.363
Skewness	6.665	2.544	2.509	3.154	1.374	0.125
Kurtosis	69.221	12.740	12.439	19.098	6.168	3.107
Q_{10}	374.14	554.43	554.83	330.79	482.18	547.43

Note: The table summarizes the distributional properties of the daily jump measures and the daily realized volatilities. J_t, $J_t^{1/2}$, and $\log(J_t+1)$ denote the daily jump measures in variance, standard deviation, and logarithmic forms, respectively. $\hat{\sigma}_t^2$, $\hat{\sigma}_t$, and $\log(\hat{\sigma}_t)$ are the daily realized variance, realized standard deviation, and realized logarithmic standard deviation, respectively. The mean, standard deviation, skewness, kurtosis and the Ljung-Box test statistic for up to tenth order serial correlation are reported. The daily jumps and the realized volatilities are constructed from 5-minute linearly interpolated returns, spanning the period from 29 September 2003 through to 31 March 2006. The stocks analyzed are Alpha Bank (ALPHA) and National Bank of Greece (ETE).

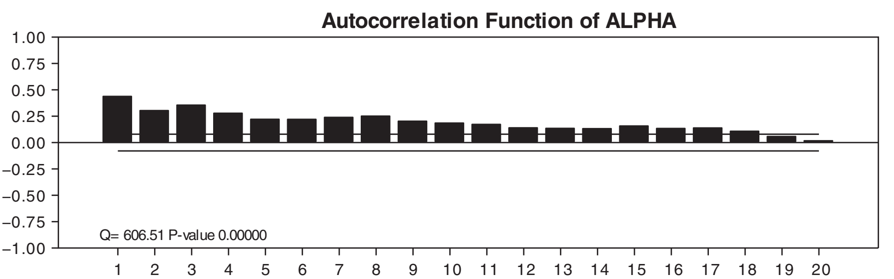

Fig. 1. Autocorrelations of the logarithm of realized volatility-ALPHA stock.

and realized standard deviation series. It is also evident that the Ljung-Box statistics for both the jumps and the realized volatility measures are highly significant up to tenth order serial correlation, especially for the series expressed in standard deviation and logarithmic form. Figure 1 shows the autocorrelations of $\log(\hat{\sigma}_t)$ for the ALPHA stock obtained at lags 1–20. They decay slowly, as previously emphasized by Andersen *et al.* (2001), and suggest we should consider a long memory model for volatility. The decay in the autocorrelations of realized volatility can be explained by a fractionally integrated process, supporting similar empirical evidence for ARCH and stochastic volatility models (see Dacorogna *et al.*, 2001). Figure 2 provides a visual illustration of daily realized volatilities and jumps of both assets under study.

A comparison of the results of Table 1 with those of Andersen *et al.*'s (2007) who deal with US S&P 500 market index data yields interesting observations. First, the average size of jumps for individual stocks is smaller than the index. Also, the jump series for individual stocks, regardless

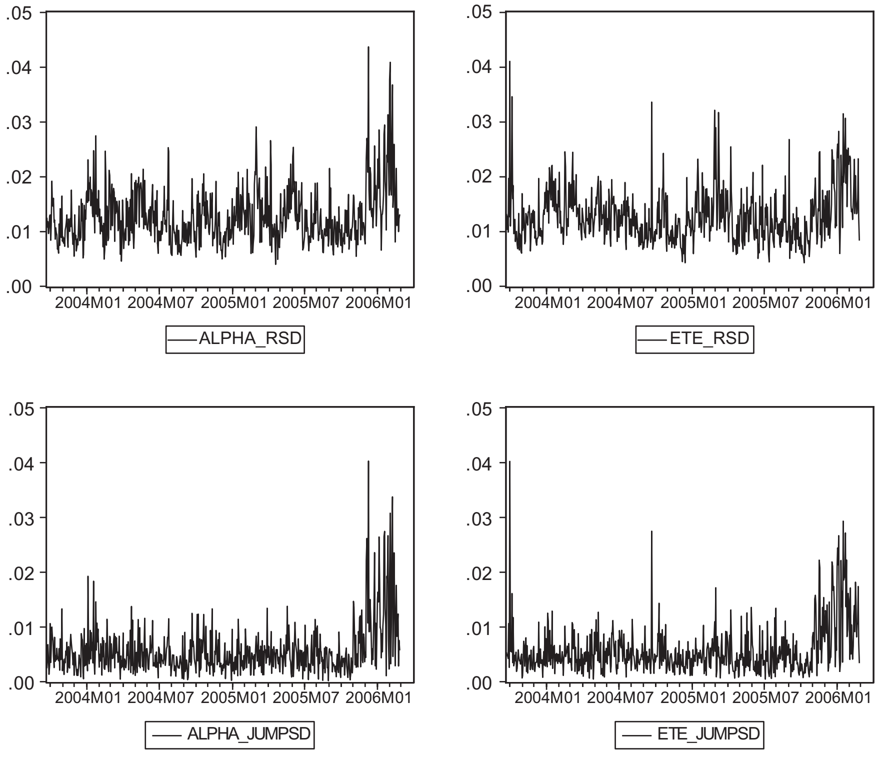

Fig. 2. Daily realized volatilities and jumps.
The upper panel visually illustrates the daily realized volatility in standard deviation form, of both stocks ALPHA and ETE. The lower panel graphs the jump measures of both stocks in standard deviation form. The data sample spans from 23 September 2003 through to 31 March 2006.

of their measurement in variance, standard deviation or logarithmic form, are less noisy than the corresponding index series, taking on smaller standard deviation values. The skewness and kurtosis values for individual stocks are smaller and closer to the normal distribution than the corresponding index values, in all specifications. Finally, the index series exhibit a higher degree of own serial correlation than the individual stocks, implying that aggregate market returns are more predictable than those of individual stocks.

Table 2 presents the daily, weekly, and monthly HAR-RV-J and 'standard' HAR model regressions. The models have been estimated using the Newey-West HAC consistent covariances. The estimates of the jump component β_J are all negative in the ALPHA stock models and statistically significant in the longer-run monthly regressions, whereas evidence is mixed in the second stock. The β_D and β_M coefficients are the largest in the daily models. Also, the coefficients in the logarithmic realized volatility models are larger in magnitude than the respective coefficients in the other two volatility model specifications. The explained variation measured by the R^2's, is lower in the case of the 'standard' HAR model where the jump component is absent.[3] Also, the gains afforded by the use of realized volatilities constructed from high-frequency data are

[3]As discussed in Andersen *et al*. (2005, 2007), it should be noted that although the relative magnitudes of the R^2's for a given volatility series are directly comparable across the two models, the measurement errors in the left-hand side realized volatilities constantly result in a systematic downward bias in the reported R^2's vis-à-vis the inherent predictability in the true latent quadratic variation process.

TABLE 2

HAR-RV-J and 'standard' HAR model regressions

ALPHA									
	$RV_{t,t+h}$			$(RV_{t,t+h})^{1/2}$			$\log(RV_{t,t+h})$		
h	1	5	22	1	5	22	1	5	22
β_0	0.00009	0.00008	0.00007	0.010	0.009	0.009	−4.630	−4.695	−4.615
	(0.00003)	(0.00004)	(0.00009)	(0.002)	(0.002)	(0.0004)	(0.111)	(0.167)	(0.038)
β_D	0.082	0.063	0.021	3.299	3.112	0.823	280.262	322.276	76.930
	(0.040)	(0.033)	(0.023)	(1.583)	(1.443)	(1.011)	(137.232)	(133.596)	(89.494)
β_W	−0.017	0.014	0.005	−0.738	0.580	−0.051	−65.305	53.688	−23.138
	(0.009)	(0.010)	(0.003)	(0.357)	(0.460)	(0.196)	(30.553)	(45.337)	(21.881)
β_M	0.008	0.002	0.007	0.354	0.068	0.304	29.987	4.621	28.427
	(0.004)	(0.004)	(0.001)	(0.178)	(0.185)	(0.051)	(15.176)	(16.919)	(4.498)
β_J	−0.197	−0.191	−0.448	−0.132	−0.122	−0.294	−13.162	−14.537	−34.468
	(0.153)	(0.291)	(0.054)	(0.120)	(0.174)	(0.045)	(11.180)	(17.009)	(5.196)
$R^2_{HAR-RV-J}$	0.333	0.116	0.195	0.328	0.122	0.188	0.300	0.133	0.220
R^2_{HAR}	0.325	0.103	0.105	0.317	0.109	0.097	0.287	0.113	0.094

(*Continued*)

TABLE 2
Continued

ETE									
	$RV_{t,t+h}$			$(RV_{t,t+h})^{1/2}$			$\log(RV_{t,t+h})$		
h	1	5	22	1	5	22	1	5	22
β_0	−0.00005	0.00007	0.0001	0.005	0.008	0.012	−4.981	−4.773	−4.427
	(0.00006)	(0.00002)	(0.00001)	(0.002)	(0.0009)	(0.0006)	(0.101)	(0.090)	(0.055)
β_D	0.234	0.019	0.001	6.499	0.512	0.006	424.194	48.281	11.583
	(0.055)	(0.020)	(0.005)	(1.277)	(0.837)	(0.224)	(74.285)	(78.597)	(18.609)
β_W	0.006	0.027	−0.018	0.608	0.951	−0.881	71.938	70.993	−84.370
	(0.019)	(0.019)	(0.005)	(0.586)	(0.734)	(0.247)	(49.112)	(59.138)	(23.957)
β_M	0.030	0.001	−0.00005	0.765	0.024	0.003	45.681	−1.260	−1.000
	(0.015)	(0.005)	(0.0009)	(0.373)	(0.152)	(0.042)	(23.405)	(10.884)	(3.641)
β_J	−0.041	0.090	0.089	−0.007	0.174	0.085	3.042	15.657	5.694
	(0.042)	(0.009)	(0.007)	(0.065)	(0.021)	(0.038)	(4.181)	(2.351)	(3.473)
$R^2_{HAR-RV-J}$	0.454	0.247	0.321	0.468	0.239	0.197	0.415	0.196	0.183
R^2_{HAR}	0.451	0.122	0.126	0.467	0.088	0.143	0.412	0.059	0.156

Note: The table reports the OLS estimates for daily (h = 1), overlapping weekly (h = 5) and monthly (h = 22) HAR-RV-J volatility regressions. Newey-West HAC standard errors are reported in parentheses. $RV_{t,t+h}$, $(RV_{t,t+h})^{1/2}$, and $\log(RV_{t,t+h})$ denote the realized measures in variance, standard deviation, and logarithmic form, respectively. The realized volatilities and jumps are constructed from 5-minute linearly interpolated returns, covering the period from 29 September 2003 through to 31 March 2006. The stocks analyzed are Alpha Bank (ALPHA) and National Bank of Greece (ETE). The last two rows in each panel report the coefficients of multiple determination (R^2) for HAR-RV-J and 'standard' HAR volatility regressions.

extended to weekly and monthly horizons. These results are in line with the findings in Andersen *et al*. (2007) providing evidence that market-wide models can be used interchangeably with models at the firm-level.

IV. CONCLUSIONS

Following recent econometric advances, this note extends our understanding on financial volatility forecasting by employing jump measures on individual stocks. Our results generalize the research findings from earlier studies on the aggregate US stock market, confirming the significance of separately measuring the jump component in forecasting volatilities. It is also emphasized that the use of high-frequency data adds considerable value in model-free, nonparametric financial econometric procedures.

REFERENCES

Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2007). 'Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility', *Review of Economics and Statistics*, 89, pp. 701–20.

Andersen, T. G., Bollerslev, T. and Meddahi, N. (2005). 'Correcting the errors: volatility forecast evaluation using high-frequency data and realized volatilities', *Econometrica*, 73, pp. 279–96.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001). 'The distribution of realized stock return volatility', *Journal of Financial Economics*, 61, pp. 43–76.

Bandi, F. M. and Russell, J. R. (2006). 'Separating microstructure noise from volatility', *Journal of Financial Economics*, 79, pp. 655–92.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). 'Econometric analysis of realized volatility and its use in estimating stochastic volatility models', *Journal of the Royal Statistical Society, Ser. B*, 64, pp. 253–80.

Barndorff-Nielsen, O. E. and Shephard, N. (2003). 'Realized power variation and stochastic volatility', *Bernoulli*, 9, pp. 243–65.

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

Bollen, N. P. and Whaley, R. E. (2004). 'Does net buying pressure affect the shape of implied volatility functions?', *Journal of Finance*, 59, pp. 711–53.

Bollerslev, T., Kretschmer, U., Pigorsch, C. and Tauchen, G. (2009). 'A discrete-time model for daily S&P500 returns and realized variations: jumps and leverage effects', *Journal of Econometrics*, 150, pp. 151–66.

Campbell, J. Y., Lettau, M., Malkiel, B. G. and Xu, Y. (2001). 'Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk', *Journal of Finance*, 56, pp. 1–43.

Corsi, F. (2009). 'A simple approximate long-memory model of realized volatility', *Journal of Financial Econometrics*, 7, pp. 174–96.

Dacorogna, M. M., Gençay, R., Müller, U., Olsen, R. B. and Pictet, O. V. (2001). '*An introduction to high-frequency finance*', San Diego, CA: Academic Press.

Huang, X. and Tauchen, G. (2005). 'The relative contribution of jumps to total price variance', *Journal of Financial Econometrics*, 3, pp. 456–99.

Papavassiliou, V. G. (2012). 'The efficiency of the realized range measure of daily volatility: evidence from Greece', *Economic Notes*, 41, pp. 173–82.

Rodrigues, P. and Schlag, C. (2009). 'A jumping index of jumping stocks? an MCMC analysis of continuous-time models for individual stocks', Working Paper, Goethe University Frankfurt.