Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility

Andrew J. Patton
Department of Economics
and
Oxford-Man Institute of Quantitative Finance
University of Oxford
andrew.patton@economics.ox.ac.uk

Kevin Sheppard*
Department of Economics
and
Oxford-Man Institute of Quantitative Finance
University of Oxford
kevin.sheppard@economics.ox.ac.uk

March 2009

Abstract

This paper examines the role that negative returns and their associated volatility play in determining future volatility. Measures of quadratic variation are decomposed into signed components which are further decomposed into signed-jump and continuous components using a simple transformation. Using data for both the S&P 500 SPDR and the individual components of the S&P 100, we find that jumps play an important role in future volatility. Moveover, we document that the effect of jumps is highly asymmetric where negative jumps lead to long lasting – almost permanent – increases in volatility while positive jumps lead to long term lower volatility.

*Contact author. Code used in this paper for computing realized quantities is available at www.kevinsheppard.com.

1 Introduction

The introduction of Realized Variance has revitalized interest in measuring and modeling conditional variance (Andersen & Bollerslev 1998, Andersen, Bollerslev, Diebold & Labys 2001, Barndorff-Nielsen & Shephard 2002*a*, Barndorff-Nielsen & Shephard 2002*b*). These papers all highlight how frequently sampled asset prices can be used to estimate important quantities of interest – namely the integrated variance and quadratic variation – for measuring the variability to asset prices. Substantial progress has been made developing estimators that are robust to market microstructure noise[1], capable of estimating the variation due to the continuous part of asset prices [2], and testing for jumps[3].

Realized Variance, and related estimators, are also particularly compelling for modeling the time-series dynamics of return variation. Beginning with Andersen, Bollerslev, Diebold & Labys (1999) and Andersen, Bollerslev, Diebold & Labys (2003), the forecasting potential of high-frequency based measures has been apparent. An incomplete list of forecasting methods and applications includes Fleming, Kirby & Ostdiek (2003), Corsi (2004), Liu & Maheu (2005), Lanne (2006*b*), Lanne (2006*a*), Chiriac & Voev (2007), Andersen et al. (2007), Chen & Ghysels (2008) and Visser (2008). Among these Andersen et al. (2007) stands out as a paper that uses the insights only available though the development of non-parametric estimators of quadratic variation and integrated variance to provide a refined view into the dynamics of total variation by decomposing quadratic variation into its continuous sample path and jump components, and exploring persistence of each and the interaction between the two components. High frequency measures have also been used for evaluating variance forecasts as in Andersen & Bollerslev (1998), Andersen, Bollerslev & Meddahi (2005) and Patton (2008).

The forecasting literature utilizing Realized Variance and related estimators has diverged from the previous two decades of literature which were dominated by ARCH-family models. Beginning with Engle's (1982) ARCH model and Bollerslev's (1986) generalization, the ARCH world has developed a virtual alphabet soup of related acronyms. Among the many specification developed for modeling the conditional variance using daily or lower frequency return series, one feature stands out as broadly empirically supported: the leverage effect – where the volatility subsequent to a negative return is higher than that subsequent to a positive return of the same size. Well know ARCH-family models capable of capturing the leverage effect include GJR-GARCH (Glosten, Jagannathan & Runkle 1993), TARCH (Zakoian 1994), and EGARCH (Nelson 1991).

The leverage effect has been mostly ignored in the Realized Variance-based forecasting liter-

[1]See Zhang, Mykland & Aït-Sahalia (2004), Barndorff-Nielsen, Hansen, Lunde & Shephard (2008*a*), Bandi & Russell (2008), *inter alia*.

[2]See Barndorff-Nielsen & Shephard (2004), Christensen, Oomen & Podolskij (2008), Andersen, Dobrev & Schaumburg (2008), *inter alia*.

[3]See Huang & Tauchen (2005) Andersen, Bollerslev & Diebold (2007), Bollerslev, Law & Tauchen (2008), Lee & Mykland (2008) *inter alia*.

ature, with the recent exceptions of Chen & Ghysels (2008) and Visser (2008). Chen & Ghysels (2008) construct an semi-parametric ARCH-like news impact curve to assess the relative contribution of returns of different signs and magnitudes. This is an ambitious undertaking but faced the difficulty of needing an estimate of the spot volatility, or a similar quantity, in-order construct devolatized residuals. Visser (2008) uses signed power variation to explore the its impact in forecasting S&P 500 return variation.

This paper uses a recently introduced estimator of Barndorff-Nielsen, Kinnebrock & Shephard (2008), Realized Semivariance, to provide insight into the components of conditional variance with a focus on the role of signed jumps. Like Realized Variance before it, Realized Semivariance estimates an integrated function of the variance process. If prices can be sampled frequently, Realized Semivariance converges to half the integrated variance plus the *signed* jump variation (can be either positive or negative, depending on the configuration of the estimators). We use a simple transformation to construct a new measure which we call *signed jump variation* that captures the variation due to jumps without requiring the estimation of the variation of the continuous component to asset prices.

Semivariance, and the broader class of downside risk measures, are not new concepts. Application of semivariance in finance include Hogan & Warren (1974) who study semivariance in a general equilibrium framework, Lewis (1990) who examined its role in option performance, and Ang, Chen & Xing (2006) who examined the role of semivariance and covariance in asset pricing. In many regards semivariance – or the variance of downside returns – is a more important measure of the relevant variability of asset prices than total variance. For more on semivariance and related measures, see Sortino & Satchell (2001).

We find that the information content of negative Realized Semivariance – that is the high frequency variation that corresponds to negative returns – is extremely informative for future variance using data on S&P 500 SPDRs and 105 firms that were in the S&P 100 between 1997 and 2008. Using models which extend up to 66 days in the future, we find that decomposing the usual Realized Variance into its signed components allows for more explanatory power at longer horizons than a measure which does not distinguish based on sign. We introduce a new measure which captures the variation due to either positive or negative signed jumps and find that jumps play a crucial role in future volatility, even at long horizons, and that the role of jumps is asymmetric – positive jump or "good volatility" lower long term variance while negative jumps or "bad volatility" raise long term variance.

The remainder of the paper is organized as follows. Section two described the stochastic environment and introduces the quantities that will be studied. Section three describes the data and section four explores the gains to decomposing Realized Variance into its two Realized Semivariances. Section five focuses on the role jumps play in future variance. Section six described the time-series properties of the new measure of variability and section seven concludes.

2 Stochastic Environment

Consider a continuous-time stochastic process for log-prices, p_t, which consists of a continuous component and a pure jump component

$$p_t = \int_0^1 \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + J_t. \tag{1}$$

where μ is a locally bounded predictable drift process, σ is a strictly positive cádlág process, J is a pure jump process and where the time interval has been normalized to 1. The quadratic variation of this process is defined

$$[p,p] = \int_0^1 \sigma_s^2 \mathrm{d}s + \sum_{0<s\leq 1} (\Delta p_s)^2 \tag{2}$$

where $\Delta p_s = p_s - p_{s-}$ captures a jump if present.

The natural estimator for the quadratic variation of a process is the sum of frequently sampled squared returns which is commonly known as realized variance (Andersen et al. 2001). For simplicity suppose that prices are observed $n+1$ times uniformly in time and denote these prices as $p_0, \ldots, p_n$.[4] Define the i^{th} return on day t as $r_{i,t} = p_{i,t} - p_{i-1,t}$, $i = 1, 2, \ldots, n$. Using these returns, the n-sample realized variance is the defined as

$$RV = \sum_{i=1}^{n} r_i^2. \tag{3}$$

As the time interval between observations becomes small, this estimator converges to the quadratic variation (Andersen et al. 2003),

$$\lim_{n\to\infty} RV \xrightarrow{p} [p,p]\,. \tag{4}$$

Barndorff-Nielsen & Shephard (2006) extended the study of estimating volatility functionals from simple estimators of the quadratic variation to a broader class which includes Bipower Variation,

$$BV = \mu_1^{-2} \sum_{i=2}^{n} |r_i||r_{i-1}, \tag{5}$$

$\mu_1 = \mathrm{E}\left[|z|\right] = \sqrt{2/\pi}$ and where z is a standard normal. Unlike realized variance, Bipower Variation's limit only includes the component of the quadratic variation due to the continuous part of the price process, the integrated variance,

$$\lim_{n\to\infty} BV \xrightarrow{p} \int_0^1 \sigma_s^2 \mathrm{d}s \tag{6}$$

[4] Uniform sampling in time is not needed for any results used in this paper.

and so the difference between Realized Variance and Bipower Variation can be used to consistently estimate the variation due to jumps (if any) of quadratic variation,

$$\lim_{n\to\infty} RV - BV \xrightarrow{p} \sum_{0\leq s\leq t} \Delta p_s. \tag{7}$$

Barndorff-Nielsen, Kinnebrock & Shephard (2008) have recently introduced a new estimator which can capture the jump variation only due to negative or positive jumps using an estimator named Realized Semivariance. Negative Realized Semivariance, defined as,

$$RS^- = \sum_{i=1}^{n} r_i^2 I_{[r_i<0]} \tag{8}$$

and positive Realized Semivariance,

$$RS^+ = \sum_{i=1}^{n} r_i^2 I_{[r_i>0]} \tag{9}$$

provide a complete decompositions of the non-signed RV so that $RV = RS^+ + RS^-$. Like Realized Variance, their limits includes variation due to both the continuous part of the price process a well as the jump component (if any), but the imposition of the indicator function allows the sign of the jump to be extracted, and so

$$\lim_{n\to\infty} RS^+ \xrightarrow{p} \frac{1}{2}\int_0^1 \sigma_s^2 \mathrm{d}s + \sum_{0\leq s\leq 1} \Delta p_s I_{[\Delta p_s>0]} \tag{10}$$

$$\lim_{n\to\infty} RS^- \xrightarrow{p} \frac{1}{2}\int_0^1 \sigma_s^2 \mathrm{d}s + \sum_{0\leq s\leq 1} \Delta p_s I_{[\Delta p_s<0]} \tag{11}$$

by Corollary 1 of Barndorff-Nielsen, Kinnebrock & Shephard (2008). An interesting consequence of the limit of Realized Semivariances is that the variation due to the continuous component can be removed by simple polarization, and we define the *signed jump variation* as

$$\Delta J^2 \equiv RS^+ - RS^-, \tag{12}$$

and it follows directly from Corollary 1 of BNKS that

$$\lim_{n\to\infty} \Delta J^2 \xrightarrow{p} \sum_{0\leq s\leq 1} \Delta p_s I_{[\Delta p_s>0]} - \sum_{0\leq s\leq 1} \Delta p_s I_{[\Delta p_s<0]}. \tag{13}$$

The results of Proposition 2 of BNKS can be similarly extended to provide the asymptotic distribution of the signed jump variation estimator when there are no jumps, and so a new signed jump test could be constructed. The asymptotic theory of BNKS is, unfortunately, infeasible except if both the drift and the leverage between the price process and the volatility are 0. The first of these, that the drift is 0, is unlikely to cause significant problems. The second, however, is implausible for many equity series and so we will not further pursue a test based on ΔJ^2.

3 Data

The data used in this paper consist of high-frequency prices on the members of the S&P 100 plus the S&P 500 tracker ETF (SPDR) between June 23, 1997 and July 31, 2008. June 23, 1997 was selected as the start date since it corresponded to the first day that U.S. equities could trade with a spread less than $\frac{1}{8}$, and a commensurate increase trading volume and a large drop in the magnitude of microstructure noise that affects realized-type estimators is seen after this date (Aït-Sahalia & Yu 2009). The ending date was determined by data availability.

Many of the constituents of the S&P 100 have changed over this period. Since the focus of this paper is on the time-series behavior of semi-variance and jumps, equities which were not continuously available for 4 years were excluded. This reduced the initial pool of 154 members to 105. The majority of the dropped assets were either dropped from the index early in the sample or added within 4 years of the sample ending date. Only a small number of firms were added to the index and spent less than 4 years in the index before being removed.

All prices used were transactions taken from NYSE TAQ. Trades were filtered to include only those occurring between 9:30:00 and 16:00:00 (inclusive) and were cleaned according to the rules detailed in Appendix A. Prices were not adjusted for splits or dividends since overnight returns are not used.

3.1 Estimator Implementation

All estimators were computed using returns sampled in business time so that prices were sampled approximately every 5-minutes. In the usual case where data is available from 9:30:00 to 16:00:00, this corresponds to sampling prices 79 times on a tick-time grid with the first sample is the first price and the final sample is the last price of the day. The remaining samples are constructed so that the number of time-stamps between each is the same. Business-time sampling is more natural than calendar time sampling and under certain conditions produces realized measures with superior statistical properties (Oomen 2005). To our knowledge this type of sampling, which we describe as *uniform sampling in business time*, has only been used in Barndorff-Nielsen, Hansen, Lunde & Shephard (2008*b*).

The choice to sample prices using an approximate 5-minute window was motivated by the desire to avoid bid-ask bounce type microstructure noise. Since the window is large relative to the tick frequency, sub-sampling can be used to improve the estimators and so all were computed by averaging the estimator over 5 uniformly spaced (in tick time) sub-samples. This procedure should produce a mild decrease in variance, and has been shown to be a reasonable choice for modeling the time-series properties of volatility in Andersen, Bollerslev & Meddahi (2006).

Denote the observed log-prices as $p_0, p_2, \ldots, p_m$ where $m+1$ is the number of unique time-stamps between 9:30:00 and 16:00:00 that have prices. Setting the number of price samples to 79, non-sub-sampled RV is computed uniformly in business time as

$$RV = \sum_{i=1}^{n} \left(p_{\lfloor ik \rfloor} - p_{\lfloor (i-1)k \rfloor}\right)^2 \tag{14}$$

where $k = m/79$ and $\lfloor \cdot \rfloor$ rounds down to the next integer. The sub-sampled version is computed by averaging over 5 uniformly spaced approximate 5-minute windows,

$$RV = \frac{1}{5} \sum_{j=0}^{4} \sum_{i=1}^{n} \left(p_{\lfloor ik + j\Delta \rfloor} - p_{\lfloor (i-1)k + j\Delta \rfloor}\right)^2 \tag{15}$$

where prices outside of the trading day are set to the close price.

Realized semi-variances, RS^+ and RS^+ were constructed in an analogous manner, only summing over the appropriately signed returns,

$$RS^- = \frac{1}{5} \sum_{j=0}^{4} \sum_{i=1}^{n} \left(p_{\lfloor ik + j\Delta \rfloor} - p_{\lfloor (i-1)k + j\Delta \rfloor}\right)^2 I_{[p_{\lfloor ik + j\Delta \rfloor} - p_{\lfloor (i-1)k + j\Delta \rfloor} < 0]} \tag{16}$$

In addition to sub-sampling, the estimator for bi-power variation was computed by averaging multiple "skip" versions. Skip versions of other estimators, namely tri-power quarticity and quad-power quarticity were found to possess superior statistical properties than returns computed using adjacent returns in Andersen et al. (2007). The standard bipower estimator, when computed using uniformly sampled prices in business time, is computed as

$$BV = \mu_1^{-2} \sum_{i=2}^{n} \left(p_{\lfloor ik \rfloor} - p_{\lfloor (i-1)k \rfloor}\right) \left(p_{\lfloor (i-1)k \rfloor} - p_{\lfloor (i-2)k \rfloor}\right). \tag{17}$$

The "skip-q " bipower variation estimator is defined as

$$BV_q = \mu_1^{-2} \sum_{i=q+2}^{n} \left(p_{\lfloor ik \rfloor} - p_{\lfloor (i-1)k \rfloor}\right) \left(p_{\lfloor (i-1-q)k \rfloor} - p_{\lfloor (i-2-q)k \rfloor}\right). \tag{18}$$

and so $BV_0 \equiv BV$ is the usual estimator. We construct our estimator of bipower variation by averaging the skip-0 through skip-4 estimators, which represents a tradeoff between locality (skip-0) and robustness to both market microstructure noise and jumps that are not contained in a single sample (skip-4).[5] Using a skip estimator was advocated in Huang & Tauchen (2005)

[5] Events which are often identified in jumps in US equity data correspond to periods of rapid price movement

as an important correction to bipower which may be substantially biased in small samples, although to our knowledge the use of an average over multiple skip-q estimators is novel. The combination can be formally justified as an asymptotic minimum-variance combination of K skip Bipower estimators.

Proposition 3.1
Let the skip bipower variation estimator be defined as in eq. (18) and assume there are no jumps. Then the joint asymptotic distribution of K such estimators is

$$\begin{bmatrix} BV_0 \\ \vdots \\ BV_K \end{bmatrix} - \int_0^1 \sigma_s^2 ds \xrightarrow{d} MN\left(\mathbf{0}, 2.69\mathbf{R}\right)$$

where $\mathbf{R}$ *is an equicorrelation matrix,*

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \vdots & \vdots \\ \rho & \dots & \rho & 1 \end{bmatrix}$$

where $\rho = .878$.

The an average is the minimum variance combination follows from directly from that structure of the correlation matrix. It is also obvious that the maximal gain for a combination would be to reduce the variance by a factor of ρ, which is fairly modest. It should also be noted that this result is only valid for finite K so that asymptotically the time-lag between the return used to compute the longest skip estimator and the contemporaneous return is negligible. If K were allowed to grow at some rate it would be expected that an asymptotic bias term would arise since some of the returns were not local to the contemporaneous return.

Noting that BV may be biased in an important way in the presence of jumps, we also include a new jump robust measure of IV developed by Andersen et al. (2008) known as median RV or $MedRV$, which is defined as

$$MedRV = \frac{n}{n-2} \sum_{i=3}^{n} \frac{\pi}{6 - 4\sqrt{3} + \pi} \text{med}\left(r_{i-2}^2, r_{i-1}^2, r_i^2\right)$$

where the term $\pi/(6 - 4\sqrt{3} + \pi)$ is a normalization constant which is the expected value of the median square of three standard normal random variables. Because MedRV will exclude the largest of a series of 3 consecutive returns it may be less affected by jumps than BV.

although they are usually characterized by multiple trades during the movement due to continuity of price rules on market makers.

Levels

$$RV_{i,t+h} = \mu_i + \phi_1 RV_{i,t} + \phi_5 \overline{RV}_{5,i,t} + \phi_{22} \overline{RV}_{22,i,t} + \epsilon_{i,t}$$

	SPDR			Panel		
h	ϕ_1	ϕ_5	ϕ_{22}	ϕ_1	ϕ_5	ϕ_{22}
1	0.404	0.226	0.247	0.392	0.294	0.185
	(.000)	(.006)	(.000)	(.000)	(.002)	(.006)
5	0.062	0.370	0.323	0.273	0.203	0.308
	(.302)	(.003)	(.003)	(.000)	(.221)	(.007)
22	−0.008	0.297	0.229	0.123	0.120	0.400
	(.851)	(.067)	(.049)	(.000)	(.343)	(.000)
66	0.008	0.061	0.197	0.069	0.054	0.425
	(.629)	(.183)	(.025)	(.000)	(.298)	(.000)

Logs

$$\ln RV_{i,t+h} = \mu_i + \phi_1 \ln RV_{i,t} + \phi_5 \ln \overline{RV}_{5,i,t} + \phi_{22} \ln \overline{RV}_{22,i,t} + \epsilon_{i,t}$$

	SPDR			Panel		
h	ϕ_1	ϕ_5	ϕ_{22}	ϕ_1	ϕ_5	ϕ_{22}
1	0.435	0.347	0.167	0.359	0.333	0.257
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
5	0.239	0.352	0.292	0.235	0.338	0.346
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
22	0.132	0.169	0.446	0.159	0.259	0.437
	(.000)	(.079)	(.000)	(.000)	(.000)	(.000)
66	0.067	0.086	0.477	0.112	0.161	0.512
	(.044)	(.261)	(.000)	(.000)	(.000)	(.000)

Table 1: Reference model parameter estimates and p-values (in parentheses). The top panel contains results form the model in levels and the bottom contains estimates from the log specification. In all cases the h-step ahead realized variance or log realized variance was regressed on time-t measurable variables. These results broadly agree with those of Andersen, Bollerslev and Diebold (2007), although ABD regressed the h-step cumulative realized variance on time-t regressors.

4 The Persistence of Volatility

Before moving into models which decompose volatility into signed components, it is instructive to establish a set of reference results. We fit a reference specification in both levels,

$$RV_{t+h} = \mu + \phi_1 RV_t + \phi_5 \overline{RV}_{5,t} + \phi_{22} \overline{RV}_{22,t} + \epsilon_t \tag{19}$$

where $\overline{RV}_{5,t}$ is the average of the past 5 RVs and $\overline{RV}_{22,t}$ is the average of the past 22 days of RV, and in logs,

$$\ln RV_{t+h} = \mu + \phi_1 \ln RV_t + \phi_5 \ln \overline{RV}_{5,t} + \phi_{22} \ln \overline{RV}_{22,t} + \epsilon_t. \tag{20}$$

which is similar to the specifications studied in Andersen et al. (2007), both of which are Heterogeneous Autoregressions (HAR) introduced to the realized volatility literature by Corsi (2004) as an adaptation of Müller, Dacorogna, Dav, Olsen, Pictet & von Weizsacker (1997). We differ from ABD in two ways: first we model the direct effect on the volatility h-days ahead, rather than the h-day cumulative realized variance, and second we include leads up to 66 days (1 quarter).

We also extend the standard HAR specification to an unbalanced, pooled panel HAR with fixed effects in both levels,

$$RV_{i,t+h} = \mu_i + \phi_1 RV_{i,t} + \phi_5 \overline{RV}_{5,i,t} + \phi_{22} \overline{RV}_{22,i,t} + \epsilon_t \tag{21}$$

and in logs,

$$\ln RV_{i,t+h} = \mu_i + \phi_1 \ln RV_{i,t} + \phi_5 \ln \overline{RV}_{5,i,t} + \phi_{22} \ln \overline{RV}_{22,i,t} + \epsilon_t \tag{22}$$

which facilitates performing inference on the average effect among the 105 S&P 100 constituents included in this study. Inference on these regressions is standard although in models where the dependant variable is more than 1-step ahead require an autocorrelation robust estimator. To implement the HAC estimator, we use a Newey & West (1987) estimator with $2(h-1)$ lags. This choice guarantees that at least half of the relevant covariance will be included at the most distant lag. Implementation details of the imbalanced panel regression and inference are presented in Appendix A.

Results of these reference models are presented in table 1. The left column correspond to the model fit on the S&P 500 SPDR and the right columns correspond to the model fit on the panel. Both sets of results show that there is substantial short term predictability of RV using recent RV, and that both the economic and statistical significance of recent RV on predicting medium to long-term RV diminishes substantially. For both models the log-specification is able to find stronger relationships and even very long horizon log RV has some predictability. The differences between these two are primarily driven by the strong heteroskedasticity present in the residuals of the levels model which is mostly eliminated by the log transformation. While we do not pursue the idea further, there may be large gains to estimating the levels specification using FGLS or as a multiplicative error model (Engle 2002).

These results are also closely in-line with the findings of ABD despite using a different dependant variable. The most notable difference is in the lack of significance of the weekly effect in the panel where only the long-run average – picking up the highly persistent level shifts in conditional variance – and the most recent RV are significant, at least in the levels model, as well as the lack of significance of short-term effect at longer horizons is apparent. These differences are likely due to different dependant variable since the ABD specification would tend to pick up the average effect across the next h days while our direct estimation strategy pins down the unique effect.

Levels

$$RV_{i,t+h} = \mu_i + \phi_1^+ RS_{i,t}^+ + \phi_1^- RS_{i,t}^- + \phi_5 \overline{RV}_{5,i,t} + \phi_{22} \overline{RV}_{22,i,t} + \epsilon_{i,t}$$

	SPDR				Panel			
h	ϕ_1^+	ϕ_1^-	ϕ_5	ϕ_{22}	ϕ_1^+	ϕ_1^-	ϕ_5	ϕ_{22}
1	−0.426	1.600	0.153	0.204	0.039	0.717	0.332	0.167
	(.006)	(.000)	(.039)	(.000)	(.801)	(.000)	(.000)	(.007)
5	−0.167	0.391	0.350	0.311	−0.005	0.529	0.233	0.294
	(.007)	(.001)	(.005)	(.003)	(.966)	(.001)	(.120)	(.003)
22	−0.199	0.266	0.280	0.219	−0.060	0.292	0.140	0.391
	(.004)	(.005)	(.085)	(.055)	(.391)	(.001)	(.228)	(.000)
66	−0.076	0.130	0.054	0.193	−0.047	0.175	0.066	0.419
	(.078)	(.028)	(.236)	(.027)	(.281)	(.002)	(.164)	(.000)

Logs

$$\ln RV_{i,t+h} = \mu_i + \phi_1^+ \ln RS_{i,t}^+ + \phi_1^- \ln RS_{i,t}^- + \phi_5 \ln \overline{RV}_{5,i,t} + \phi_{22} \ln \overline{RV}_{22,i,t} + \epsilon_{i,t}$$

	SPDR				Panel			
h	ϕ_1^+	ϕ_1^-	ϕ_5	ϕ_{22}	ϕ_1^+	ϕ_1^-	ϕ_5	ϕ_{22}
1	−0.008	0.401	0.389	0.164	0.098	0.257	0.338	0.255
	(.752)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
5	−0.006	0.225	0.372	0.290	0.042	0.193	0.339	0.345
	(.829)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
22	−0.056	0.174	0.183	0.445	0.018	0.141	0.259	0.437
	(.088)	(.000)	(.055)	(.000)	(.083)	(.000)	(.000)	(.000)
66	−0.020	0.080	0.093	0.477	0.005	0.106	0.162	0.512
	(.563)	(.041)	(.238)	(.000)	(.669)	(.000)	(.000)	(.000)

Table 2: Base model parameter estimates and p-values. The top panel contains results form the model in levels and the bottom contains estimates from the log specification. The coefficients on the positive semi-variance and negative semi-variance are substantially different. In the case of the market the effect of positive semi-variance is either insignificant or significantly negative, while the negative semi-variance has uniformly large, statistically significant effects at any horizon. The volatilities of individual equities exhibit a similar patters although there is some evidence that positive semi-variance has a small positive but statistically significant effect on future volatility, although this effect is small relative to the magnitude of the coefficients on negative semi-variance.

4.1 Decomposing Recent Quadratic Variation

Given the exact decomposition of RV into RS^+ and RS^+, extending eqs. () – () is leads to a direct test of whether signed information has further information than RV for future volatility. Applying this decomposition produces a base specification model in levels

$$RV_{t+h} = \mu + \phi_1^+ RS_t^+ + \phi_1^- RS_t^- + \phi_5 \overline{RV}_{5,t} + \phi_{22} \overline{RV}_{22,t} + \epsilon_t \tag{23}$$

and in logs

$$\ln RV_{t+h} = \mu + \phi_1^+ \ln RS_t^+ + \phi_1^- \ln RS_t^- + \phi_5 \ln \overline{RV}_{5,t} + \phi_{22} \ln \overline{RV}_{22,t} + \epsilon_t. \tag{24}$$

The pooled panel models are virtually identical, only modified to permit fixed effects, and are omitted for brevity.

Results from these models for both the S&P 500 SPDR as well as the pooled panel of individual firms are reported in table 2. Both the large number of negative coefficients, as well as the large differences in the magnitude of the coefficients on the positive and negative Realized Semivariance, differ from the reference results in table 1. The effect of the RV can be constructed by adding $\phi_1 \approx \phi_1^+ + \phi_1^-$ where the approximation would be exact if the coefficient on the weekly and monthly RV were restricted to be identical in the two tables. In contrast to the predictive power of RV, both Realized Semivariances appear to have significant effects over many horizons. This is especially true for the S&P 500 SPDR results where the daily RV was insignificant at all leads except 1 day. Using the decomposed results shows that so called "good volatility" – positive Realized Semivariance – has a long lasting effect on future volatility by *lowering* it. In contrast, "bad volatility" – negative Realized Semivariance – leads to long lasting increases in future variance.

The results for the panel of individual firms are similar with the important caveat that good volatility does not generally lead to lower future volatility. In the log specification of the panel, positive Realized Semivariance leads to larger future variance out to a month, although the magnitude of the change is small relative to the effect of negative Realized Semivariance.

Figures 2 and 3 present the complete set of 66 estimated parameters on the S&P 500 SPDR and the panel of individual firms, respectively, where both represent the levels version of the model. The S&P 500 shows a very large effect over short horizons for negative Realized Semivariance which degrades into a persistent and generally statistically significantly different from zero effect out to the maximum horizon. The positive Realized Semivariance has uniformly negative coefficients although these are both lower in absolute magnitude than those on the negative RS and only significantly different from zero for the first half of the leads. It appears likely that all of the coefficients on the negative Semivariance could be parsimoniously fit using an exponentially decaying function, possibly with two scales (one would be needed to generate the fast decay over a short horizon, while the other would be needed to fit the longer horizons).

The coefficients on RS^+ and RS^- for the panel has a slightly different patters with some evidence of a short term increase subsequent to positive Realized Semivariance, although all coefficients on RS^+ are not statistically different from zero. The coefficient on negative Realized Semivariance, on the other hand, are large in magnitude and uniformly statistically significant.

The smoothness indicated in both curves is a feature of the estimated parameters and additional smoothing was used to produce these figures.

4.2 Completely Decomposing Quadratic Variance

The specification in eq. (23) restricts the coefficients on the weekly and monthly realized semivariances to be identical by parameterizing the weekly and monthly terms using Realized Variances. This restriction can be relaxed by decomposing the RV terms at all lags. With this modification, the levels model is

$$RV_{t+h} = \mu + \phi_1^+ RS_t^+ + \phi_1^- RS_t^- + \phi_5^+ \overline{RS}_t^{5+} + \phi_5^- \overline{RS}^5-_t + \phi_{22}^+ \overline{RS}_t^{22+} + \phi_{22}^- \overline{RS}_t^{22-} + \epsilon_t \quad (25)$$

where $\overline{RS}^j-$ is the j-day average of negative semivariance and $\overline{RS}^j+$ is the j-day average of positive semivariance, and the log specification is

$$\ln RV_{t+h} = \mu + \phi_1^+ \ln RS_t^+ + \phi_1^- \ln RS_t^- + \phi_5^+ \ln \overline{RS}_t^{5+} + \phi_5^- \ln \overline{RS}^5-_t + \phi_{22}^+ \ln \overline{RS}_t^{22+} + \phi_{22}^- \ln \overline{RS}_t^{22-} + \epsilon_t \quad (26)$$

where the convention that the weekly and monthly terms in the log specification are logs of averages and not averages of logs was continued. The panel version of these two models is identical except for the fixed effect allowed to capture the different long-run level of each asset's variance.

Results of these extended specifications where all terms are decomposed are presented in table 3. It is not surprising that many of the coefficients are now insignificant. Two features are present in the models estimated in levels (top row). First, the terms based on the positive Realized Semivariance are generally insignificant and in the three cases where a positive RS is significant, it has a negative sign. The three occurrences are all for the SPDR in the 1-day lead model for both daily and monthly positive Realized Semivariance and monthly positive RS for the 5-day lead model. All other statistically significant coefficients in the levels models have are for negative RS and have positive signs. Interestingly the weekly negative Realized Semivariance for the SPDR is significant and large in magnitude at all leads except 66-days ahead which contrasts to the results using the non-decomposed RV in table 2 where only the 1-day and 5-day leads had significant coefficients and the values were much smaller.

More parameters in the log specification were statistically significant although the signs were not as clearly aligned between the positive and negative Realized Semivariances. For the SPDR the positive RS are only significant for the weekly term in the 1-day lead models. The remainder are insignificant with mixed signs. Negative RS is generally significant with large in magnitude coefficients. A similar pattern is evidenced in the panel although the coefficient on the daily positive RS for the 1-day lead model is both statistically significant and large in magnitude relative to the coefficient on the negative RS.

5 Signed Jump Information

All of the models estimated thus far are examined the role that decomposing Realized Variances into positive and negative Realized Semivariance can play in explaining future volatility. These results consistently demonstrated that the information content of negative Realized Semivariance was substantially larger than that of positive Realized Semivariance. While the theory of BNKS shows that the difference in these two can be attributed to differences in jump variation, the direct effect of jumps is diluted since both Realized Semivariances contain half of the integrated variance.

We use the previously introduced *signed jump variation*, $\Delta J_t \equiv RS_t^+ - RS_t^-$, as a method to isolate the jumps. This difference should eliminate the common integrated variance term in each and produce a measure that is positive when a day is dominated by an upward jump and negative when a day is dominated by a downward jump. This difference has the added advantage that it isn't necessary to introduce a jump robust estimator such as Bipower Variation or MedRV when computing this signed measure. If jumps are rare as is often advocated in the stochastic volatility literature, then this measure should broadly correspond to the jump variation when there a jump occurs and be mean zero noise otherwise.

The directly explore the role that signed jumps play in future variance we formulate a model which contains signed jump variation as well as an estimator of the variation due to the continuous part using Bipower Variance. In levels this specification is

$$RV_{t+h} = \mu + \phi^J \Delta J_t + \phi^C BV_t^- + \phi_5 \overline{RV}_{5,t} + \phi_{22} \overline{RV}_{22,t} + \epsilon_t. \tag{27}$$

Since signed jump variation is often negative, the log specification requires a slight modification. Since log-log specifications given rise to a natural elasticity interpretation, a modified signed jump variation which measures the signed percentage of the total variation due to jumps is constructed by

$$\%\Delta J^2 = \left(1 + \frac{\Delta J_t}{RV_t}\right)$$

Using this measure the log specification can be specified where the other terms are as before,

$$\ln RV_{t+h} = \mu + \phi^J \ln\left(1 + \frac{\Delta J_t}{RV_t}\right) + \phi^C \ln BV_t^- + \phi_5 \ln \overline{RV}_{5,t} + \phi_{22} \ln \overline{RV}_{22,t} + \epsilon_t. \tag{28}$$

The panel specification are sufficiently similar to not warrant separate presentation. It is also worth noting that while this specification is very similar to the baseline model (eq. 23) that it is not nested, even if the estimators on the RHS are replaced by their population values since it is not possibly to construct a measure of the variation due to the continuous part from the two Realized Semivariance alone.

Results of the signed jump specification are presented in table 4 for both the SPDR and the panel of individual firm variances. These results are remarkably consistent, although there are important differences in magnitude between the SPDR and the individual firms. In the top row – the models estimated in levels – signed jump variation has a uniformly negative sign and is significant at all leads. Variation due to the diffusive part of volatility is has a large effect over short horizons but becomes less relevant for monthly and quarterly leads. Moreover, in the SPDR series, there is no evidence of persistence of diffusive volatility at the longer horizons.

The log models confirm the findings in the level specifications where the sign on percentage signed jump variation is also uniformly negative, although more evidence is present for the persistence of the variation due to the diffusive component of asset prices. The other coefficients, as well as their significance, are virtually unchanged from the baseline models reported in table 2. Figures 3 and 4 contain the entire series of values for the signed jump variation component and the diffusive volatility component for the SPDR and Panel, respectively. Figure 3 shows that the SPDR has nearly equal and opposite signs on the diffusive part of volatility and signed jump variation. Figure 4 documents a similar effect but shows that individual firms have a more pronounced diffusive component of volatility and that the effect of signed jump variation is less precise at the shortest horizon. This may be due to heterogeneity in the panel where the short term reaction of some assets differs from others.

Adding in signed information paints a very different picture that what was found by ABD who documented in their HAR-RV-J model that jumps leads to a slight decrease in future variance in the S&P 500, although some caution is warranted. Our model does not pretest for jumps while ABD does, and so on days where there no jump component is detected ABD's jump measure is zero. Since we do not pretest, we may have a noisier jump measure, although it is consistent for the object of interest.

The final specification used will allow for the coefficient on positive excess jump variation to differ from that of negative jump variation, and so it allows the restriction which is imposed in eqs. (27) and (28) to be tested to see if the effect is uniform across the sign, or is being driven more by positive or negative jump variation. Since the sign is no longer restricted it is not possible to eliminate the variation due to the continuous part of the price process simply by differencing the Realized Semivariances. One option would be to subtract a consistent estimator of the IV, for example to use $RS_t^+ - \frac{1}{2}BV_t$. We opted for a simpler specification which inserts a 0 for the signed jump variation depending on which Realized Semivariance was larger, which avoids the introduction of a noisy IV estimator. With these changes the linear specification is

$$RV_{t+h} = \mu + \phi^{J+}\left(RS_t^+ - RS_t^-\right)I_t + \phi^{J-}\left(RS_t^+ - RS_t^-\right)(1 - I_t) + \phi^C BV_t^- + \phi_5 \overline{RV}_{5,t} + \phi_{22}\overline{RV}_{22,t} + \epsilon_t. \tag{29}$$

where $I_t \equiv (RS_t^+ - RS_t^-) > 0$ is an indicator variable for whether the positive Realized Semivariance was larger than the negative Realized Semivariance. The log specification requires further modification since one or both of the jump measures may be negative. We used a similar change to compute the percentage jump variation, and so the log specification is

$$\begin{aligned}\ln RV_{t+h} =& \mu + \phi^{J+} \ln\left(1 + I_t \frac{RS_t^+ - RS_t^-}{RV_t}\right) + \phi^{J-} \ln\left(1 + (1 - I_t) \frac{RS_t^+ - RS_t^-}{RV_t}\right) \\ & + \phi^C \ln BV_t^- + \phi_5 \ln \overline{RV}_{5,t} + \phi_{22} \ln \overline{RV}_{22,t} + \epsilon_t.\end{aligned} \quad (30)$$

Table 5 contains estimates and p-values for the extended jump specification. These estimates confirm that the findings in the restricted model are robust to the imposition of the shared coefficient. With the exception of an insignificant coefficient at the longest lead for the SPDR, all coefficients on the signed jump variation measures were uniformly negative. Moreover, the coefficients on the negative signed jump variation were more larger in magnitude than on the positive and were consistently statistically significant.

Figure 5 contains a plot of the coefficients for all 66 leads in the panel model estimated in levels. Aside from some mixed evidence for very short term effects, both sets of coefficients are negative although the coefficients on the positive jump variation are not significantly different from zero.

5.1 Individual Significance

While the pooled panel results were highly significant and supportive of the idea that jump variation – sign included – is an important determinant of future volatility, some caution is needed. It may be the case that all of these series have some exposure to a market effect and in pooling this common effect is highlighted where it wouldn't be if the models were estimated separately. To test whether the findings were robust to running individual models the three main models, the baseline specifications (eqs. (23) and (24)), the jump specifications (eqs. (27) and (28)), and the extended jump specification were all fit to the 105 individual firm variances.

To facilitate presentation of these results, only the percent significant for the main effect, along with the sign, are reported in table 6 using a test with a 5% size. In the baseline specifications in levels, negative Realized Semivariance was significantly negative for many of the series and shad a significantly negative coefficient only once. Similarly positive Realized Semivariance was only significantly positive (above size) for the shortest horizon, where 16% of series rejected the null. These results are inline with the panel and provide some insight into the mixed results for positive Realized Semivariance over very short horizons: a minority appear to have a positive loading on positive Realized Semivariance and virtually none exhibit a decrease in the short-term. The results for the baseline log specification are similar only with more evidence of significance for long horizon positive effects of negative Realized Semivariance.

The jump and the extended jump specifications both confirm that the findings in the pooled panel model are pervasive in the individual variance series. At most two series have significantly positive coefficients on signed jump variation in the jump specification – either levels or logs – which corresponds to a rejection rate well below size. On the other hand 30% and up have a significantly negative sign on signed jump variation indicating that negative jumps have a significant and long lasting affect on future variance. Finally the extended jump specification confirms that the sign restriction in the jump specification is not overly restrictive where the rejection rates that are above size are in favor of negative loadings on the decomposed components of signed jump variation.

Figure 6 contains a plot of the estimated coefficients on the 105 individual firms from the jump specification. The y-value indicated the magnitude of the coefficient where solid bars are statistically significant. The most striking feature of this plot is the pervasiveness of the negative sign on signed jump variation – even in cases where it is insignificant – coupled with the presence on only 1 significantly positive coefficient in both plots.

6 Signed Jump Variation

Finally the times series of signed jump variation, $\Delta J_t^2 = RS_t^+ - RS_t^-$, is also of direct interest. If jumps are rare – one or none a day – then this series contains either jumps or noise. Figure 8 contains the time series plot of the SPDR signed jump variation series. This series is consistent with a heteroskedastic white noise process. The first 5 autocorrelations are -0.111, -0.022, -0.037, 0.015 and 0.0212 and each is individually insignificantly different from zero using heteroskedasticity robust inference. Signed jump variation comprises 15% of total variance, although the percentage due is a persistent time series.

While it is tempting to interpret the heteroskedasticity as evidence of increased jump activity, it cannot be directly interpreted since the asymptotic variance of signed jump variation is proportional to the integrated quarticity which will be high in periods of high volatility. Figure 8 also highlights two days. The first has a large positive signed jump variation and the second has large negative jump variation. These price path of the SPDR on these two is plotted in figure 8. The positive jump on January 3, 2001 and the negative jump on December 11, 2007. Both of these jumps correspond to unexpected cuts by the Fed. On January 3, 2001 the Fed announced a surprise 50 basis point cut which triggered a rally. The December 11 drop also corresponded to an unexpected cut by the Fed – 25 basis points – which was interpreted as an ominous sign of an impending recession.

7 Conclusion

This paper has studied the role that signed information – especially the role of signed jumps – plays in determining future variance. This decomposition was facilitated using Realized Semivariance estimators of Barndorff-Nielsen, Kinnebrock & Shephard (2008) in a framework conceptually similar to Andersen et al. (2007). Beginning with a simple decomposition of the Realized Variance into positive and negative Realized Semivariance, we documented that signed information is important, and that a negative Realized Semivariance captures a modern version of the familiar leverage effect.

This paper has documented that jumps play an important role in determining future price variability although *sign matters*. When signed jump information was introduced into the model it is often found that jump variation has a linear relationship, and so aggregate unsigned jump variation will have a small role in future volatility. However once signed information is incorporated with jumps a very different picture emerges – positive jumps or "good volatility" lead to lower future variance while negative jumps or "bad volatility" lead to higher volatility, often over long horizons.

The most important extension of this work is to reconsider common stochastic volatility models. While many stochastic volatility models contain leverage and/or jumps, to our knowledge only the BNS Lévy driven stochastic volatility model contains leverage *between* the jump process and the volatility (Barndorff-Nielsen & Shepard 2001). Moreover, whether the volatility of market returns contains a diffusive component at all is an open question (Todorov & Tauchen 2008). It is also within the realm of possibility that differing degrees of persistence of positive jump induced volatility, negative jump induced volatility, and continuous part volatility may play a role in the long-memory often documented in volatility.

It would further be interesting to extend this to asset classes outside of equities, and to other markets. The S&P 500 is well known to have a strong leverage effect and it would be interesting to document which assets also exhibit a similarly strong leverage effect. The precision of realized measures allows for easier identification which may allow previously undocumented leverage to be detected. Finally, it is not clear whether the Realized Variance is even required to produce accurate forecasts if Realized Semivariance can be constructed.

A Inference on the Unbalanced Panel

The models will be fit both individually to the S&P 500 SDPR and the individual firms. Separate estimation on the models on the individual firms does not provide a direct method to assess the significance of the average effect, and so a pooled unbalanced panel HAR with a fixed effect for each series will be fit which will allow inference on the average values of the predictive parameters. In the simplest specification,

$$RV_{i,t+h} = \mu_i + \phi_1 RV_{i,t} + \phi_5 \overline{RV}_{5,i,t} + \phi_{22} \overline{RV}_{22,i,t} + \epsilon_{i,t},\ i = 1, \ldots, n_t,\ t = 1, \ldots, T$$

which can be generically expressed as

$$y_{i,t+h} = \mu_i + \boldsymbol{\phi}' Y_{i,t} + \epsilon_{i,t},\ i = 1, \ldots, n_t,\ t = 1, \ldots, T.$$

Define $\tilde{y}_{i,t+h} = y_{i,t+h} - \bar{y}_i$ and $\tilde{Y}_{i,t} = Y_{i,t} - \bar{Y}_i$ as the demeaned regressand and regressors, respectively. The pooled parameters can be estimated by

$$\hat{\boldsymbol{\phi}} = \left(T^{-1} \sum_{t=1}^{T} \left(n_t^{-1} \sum_{i=1}^{n_t} \tilde{Y}_{i,t} \tilde{Y}_{i,t}' \right) \right)^{-1} \left(T^{-1} \sum_{t=1}^{T} \left(n_t^{-1} \sum_{i=1}^{n_t} \tilde{Y}_{i,t} \tilde{y}_{i,t} \right) \right). \tag{31}$$

Inference can be similarly made using the asymptotic distribution

$$\sqrt{T} \left(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0 \right) \xrightarrow{d} N \left(\mathbf{0}, \boldsymbol{\Sigma}^{-1} \boldsymbol{\Omega} \boldsymbol{\Sigma}^{-1} \right)$$

where

$$\boldsymbol{\Sigma}^{-1} = \text{plim}_{\ T \to \infty} T^{-1} \sum_{t=1}^{T} \left(n_t^{-1} \sum_{i=1}^{n_t} \tilde{Y}_{i,t} \tilde{Y}_{i,t}' \right)$$

and

$$\boldsymbol{\Omega}^{-1} = \text{avar} \left(T^{-1/2} \sum_{t=1}^{T} \mathbf{z}_t \right)$$

where

$$\mathbf{z}_t = n_t^{-1} \sum_{i=1}^{n_t} \tilde{Y}_{i,t} \epsilon_{i,t}$$

It should be noted that the cross-section size, even if very large, does not appear in the asymptotic distribution since it should be expected that plim $_{n_t \to \infty} n_t{}^{-1} \sum_{i=1}^{n_t} \tilde{Y}_{i,t} \epsilon_{i,t} \to \tau^2 > 0$ due to the factor structure of returns, and so this inference strategy would be valid under with a fixed cross-section size, or an expanding one. A similar result was found in the context of composite likelihood estimation, and this asymptotic distribution can be seem as a special case of Engle, Shephard & Sheppard (2008).

B Data Cleaning

Only transaction data were taken from the NYSE TAQ. All series were automatically cleaned according to a set of 6 rules:

1. Transactions outside of 9:30:00 AM and 16:00:00 were discarded
2. Transactions with a 0 price or volume were discarded
3. Each day the most active exchange was determined. Only transactions from this exchange were retained.
4. Only trades with conditions E, F or blank were retained.
5. Transaction prices outside of the CRSP high or low were discarded.
6. Trade with immediate reversals more than 5 times a 50-sample moving window - excluding the transaction being tested - were discarded.

These rules are similar to those of Barndorff-Nielsen, Hansen, Lunde & Shephard (2008*b*), and prices were not manually cleaned for problems not addressed by these rules.

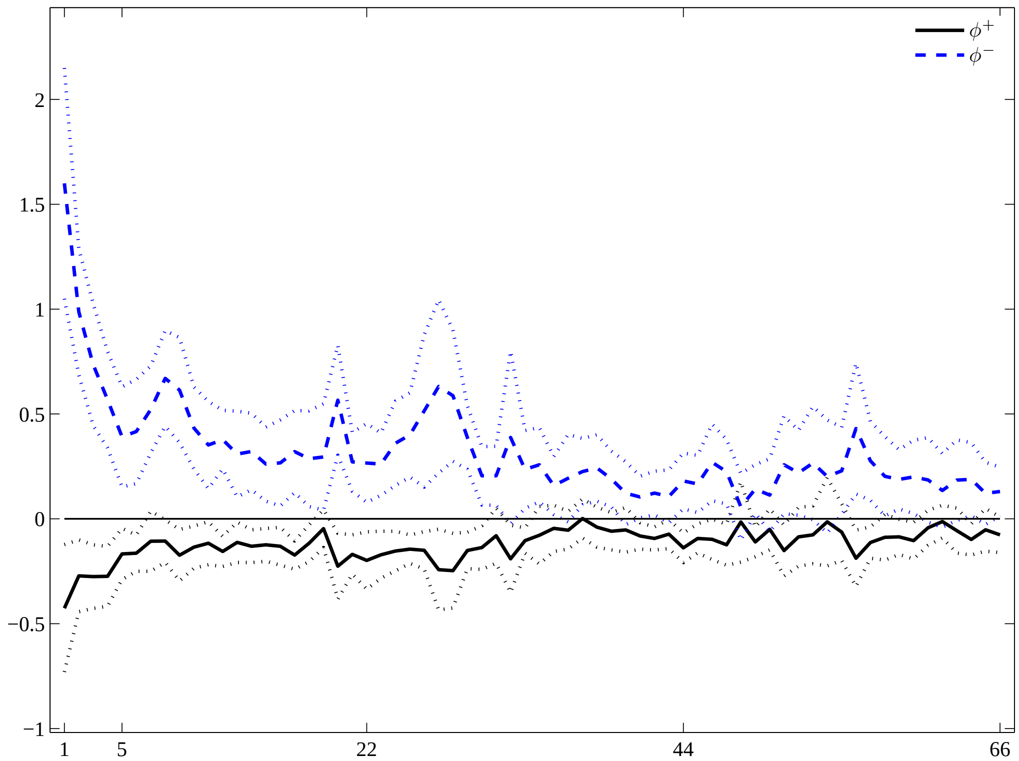

Figure 1: Fit coefficients from the levels model that decomposed realized variance into its signed components, $RV_{t+h} = \mu + \phi_1^+ RS_t^+ + \phi_1^- RS_t^- + \phi_5 \overline{RV}_{5,t} + \phi_{22} \overline{RV}_{22,t} + \epsilon_t$, and the S&P 500 SPDR. The coefficients on positive semi-variance are uniformly negative, and significant over relatively long horizons. The coefficients on negative semi-variance are significant at almost all leads and large in magnitude.

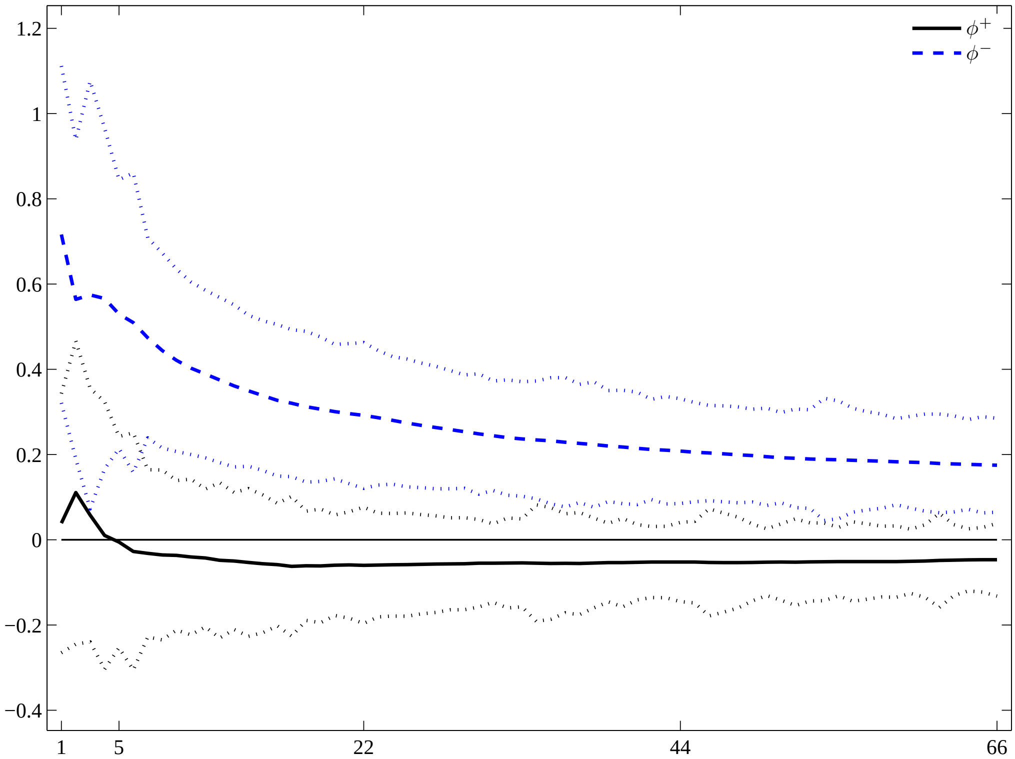

Figure 2: Fit coefficients from the levels model that decomposed realized variance into its signed components, $RV_{i,t+h} = \mu_i + \phi_1^+ RS_{i,t}^+ + \phi_1^- RS_{i,t}^- + \phi_5 \overline{RV}_{5,i,t} + \phi_{22} \overline{RV}_{22,i,t} + \epsilon_{i,t}$, and the panel of individual firm volatilities. The coefficients on positive semi-variance are uniformly negative, and significant over all horizons except the shortest. The coefficients on negative semi-variance are significant at all leads and large in magnitude.

Levels

$$RV_{i,t+h} = \mu_i + \phi_1^+ RS^+_{i,t} + \phi_1^- RS^-_{i,t} + \phi_1^+ \overline{RS}^+_{5,i,t} + \phi_1^- \overline{RS}^-_{5,i,t} + \phi_1^+ \overline{RS}^+_{22,i,t} + \phi_1^- \overline{RS}^-_{22,i,t} + \epsilon_{i,t}$$

	SPDR						Panel					
h	ϕ_1^+	ϕ_1^-	ϕ_5^+	ϕ_5^-	ϕ_{22}^+	ϕ_{22}^-	ϕ_1^+	ϕ_1^-	ϕ_5^+	ϕ_5^-	ϕ_{22}^+	ϕ_{22}^-
1	−0.332	1.411	−0.285	0.657	−0.619	1.136	0.120	0.584	−0.129	0.819	−0.192	0.553
	(.051)	(.000)	(.112)	(.020)	(.035)	(.002)	(.446)	(.004)	(.576)	(.001)	(.136)	(.006)
5	−0.036	0.132	−0.176	0.937	−1.180	1.999	0.090	0.375	−0.274	0.762	−0.283	0.910
	(.585)	(.335)	(.389)	(.006)	(.030)	(.003)	(.435)	(.004)	(.339)	(.000)	(.282)	(.001)
22	−0.049	−0.029	−0.340	0.978	−1.377	2.023	0.013	0.176	−0.204	0.492	−0.213	1.030
	(.337)	(.788)	(.189)	(.011)	(.126)	(.039)	(.842)	(.000)	(.163)	(.011)	(.676)	(.078)
66	−0.032	0.046	−0.077	0.189	−0.484	0.957	0.002	0.096	−0.178	0.320	0.058	0.803
	(.235)	(.312)	(.416)	(.186)	(.505)	(.233)	(.959)	(.000)	(.103)	(.004)	(.862)	(.021)

Logs

$$\ln RV_{i,t+h} = \mu_i + \phi_1^+ \ln RS^+_{i,t} + \phi_1^- \ln RS^-_{i,t} + \phi_1^+ \ln \overline{RS}^+_{5,i,t} + \phi_1^- \ln \overline{RS}^-_{5,i,t} + \phi_1^+ \ln \overline{RS}^+_{22,i,t} + \phi_1^- \ln \overline{RS}^-_{22,i,t} + \epsilon_{i,t}$$

	SPDR						Panel					
h	ϕ_1^+	ϕ_1^-	ϕ_5^+	ϕ_5^-	ϕ_{22}^+	ϕ_{22}^-	ϕ_1^+	ϕ_1^-	ϕ_5^+	ϕ_5^-	ϕ_{22}^+	ϕ_{22}^-
1	0.032	0.324	−0.097	0.479	−0.148	0.356	0.123	0.224	0.056	0.281	0.044	0.219
	(.180)	(.000)	(.048)	(.000)	(.111)	(.000)	(.000)	(.000)	(.000)	(.000)	(.019)	(.000)
5	0.016	0.181	0.081	0.273	−0.248	0.580	0.066	0.161	0.066	0.272	0.062	0.292
	(.570)	(.000)	(.439)	(.019)	(.237)	(.011)	(.000)	(.000)	(.015)	(.000)	(.174)	(.000)
22	−0.025	0.115	−0.070	0.230	−0.226	0.723	0.043	0.109	0.026	0.231	0.096	0.349
	(.414)	(.001)	(.596)	(.059)	(.629)	(.162)	(.000)	(.000)	(.507)	(.000)	(.317)	(.000)
66	0.002	0.033	−0.130	0.228	0.132	0.366	0.027	0.077	−0.016	0.175	0.170	0.350
	(.943)	(.164)	(.437)	(.204)	(.804)	(.521)	(.001)	(.000)	(.680)	(.000)	(.057)	(.000)

Table 3: Extended model where all terms are decomposed into positive and negative semi-variance components parameter estimates and p-values . The top panel contains results form the model in levels and the bottom contains estimates from the log specification. While these coefficients are noisier than in the base specification, the patters of positive loadings on negative semi-variance coefficients and negative loadings on positive semi-variances is preserved at all lags. When the coefficients do not strictly conform to this pattern they are insignificant, although the short term increase in volatility due to positive semi-variance for individual firms remains.

Levels

$$RV_{i,t+h} = \mu_i + \phi_J \Delta J^2_{i,t} + \phi_C BV_{i,t} + \phi_5 \overline{RV}_{5,i,t} + \phi_{22} \overline{RV}_{22,i,t} + \epsilon_{i,t}$$

	SPDR				Panel			
h	ϕ_J	ϕ_C	ϕ_5	ϕ_{22}	ϕ_J	ϕ_C	ϕ_5	ϕ_{22}
1	−0.817	0.622	0.129	0.208	−0.319	0.520	0.296	0.160
	(.000)	(.000)	(.067)	(.000)	(.011)	(.000)	(.001)	(.008)
5	−0.249	0.139	0.328	0.312	−0.256	0.347	0.219	0.288
	(.000)	(.078)	(.012)	(.003)	(.012)	(.000)	(.125)	(.004)
22	−0.227	0.051	0.265	0.219	−0.170	0.156	0.132	0.388
	(.000)	(.388)	(.113)	(.054)	(.005)	(.000)	(.224)	(.000)
66	−0.092	0.023	0.057	0.193	−0.108	0.087	0.062	0.418
	(.053)	(.304)	(.222)	(.027)	(.010)	(.000)	(.178)	(.000)

Logs

$$\ln RV_{i,t+h} = \mu_i + \phi_J \,\%\Delta J^2_{i,t} + \phi_C \ln BV_{i,t} + \phi_1^- \ln RS^-_{i,t} + \phi_5 \ln \overline{RV}_{5,i,t} + \phi_{22} \ln \overline{RV}_{22,i,t} + \epsilon_{i,t}$$

	SPDR				Panel			
h	ϕ_J	ϕ_C	ϕ_5	ϕ_{22}	ϕ_J	ϕ_C	ϕ_5	ϕ_{22}
1	−0.418	0.400	0.376	0.167	−0.176	0.336	0.351	0.264
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
5	−0.224	0.225	0.363	0.292	−0.155	0.215	0.355	0.351
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
22	−0.206	0.124	0.175	0.447	−0.125	0.134	0.281	0.441
	(.000)	(.001)	(.067)	(.000)	(.000)	(.000)	(.000)	(.000)
66	−0.092	0.059	0.093	0.477	−0.104	0.079	0.191	0.514
	(.200)	(.072)	(.241)	(.000)	(.000)	(.000)	(.000)	(.000)

Table 4: Model which includes signed jump information where the asymptotic relationship where quadratic variation has been decomposed into signed jump variation, ΔJ^2, and its continuous component using Bipower Variation, BV. In the log specification the percentage jump variation is defined as $\%\Delta J^2_{i,t} = \ln\left(1 + (\Delta J^2_{i,t})/RV_{i,t}\right)$. P-values are reported below parameter estimated in parentheses. The top panel contains results form the model in levels and the bottom contains estimates from the log specification. The negative coefficients on ΔJ^2 at all horizons indicate that variation due to positive jumps lowers future variance since $\Delta J^2 = RS^+ - RS^- > 0$ while variation due to negative jumps raises future variance. There are also notable differences between the coefficients on Bipower Variation in the S&P 500 SPDR and the individual firms. The continuous component does not persist out to the monthly horizon for the index while the signed jump component does – for the individual firms, both components are significant and most horizons.

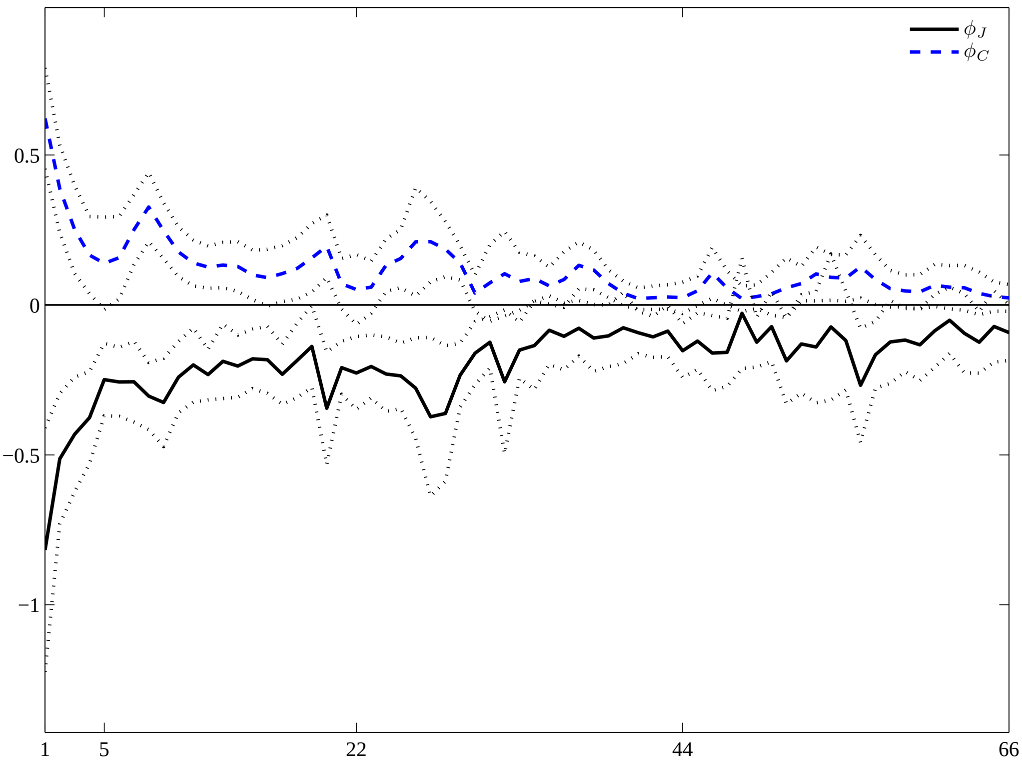

Figure 3: Fit coefficients from the levels model that included both signed jump variation and bipower variation, $RV_{t+h} = \mu + \phi^J \Delta J_t + \phi^C BV_t^- + \phi_5 \overline{RV}_{5,t} + \phi_{22} \overline{RV}_{22,t} + \epsilon_t$, along with 95% confidence intervals for the S&P 500 SPDR.

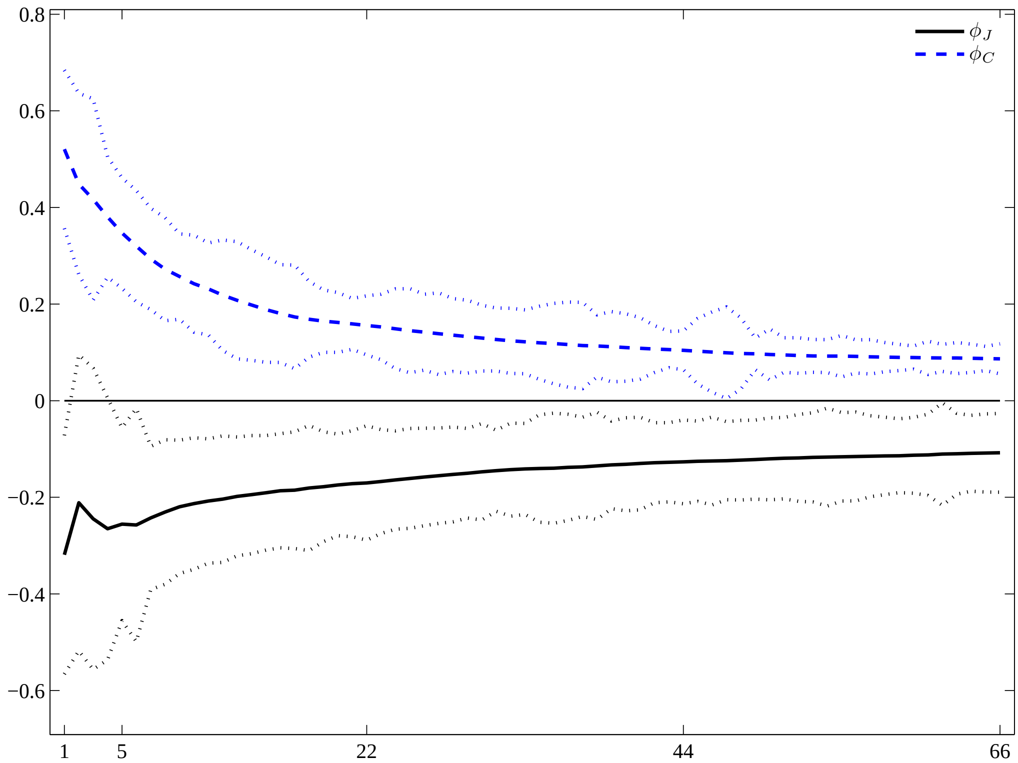

Figure 4: Fit coefficients from the levels model that included both signed jump variation and bipower variation, $RV_{i,t+h} = \mu_i + \phi^J \Delta J_{i,t} + \phi^C BV^-_{i,t} + \phi_5 \overline{RV}_{5,i,t} + \phi_{22} \overline{RV}_{22,i,t} + \epsilon_{i,t}$, along with 95% confidence intervals for the panel of individual volatilities.

Levels

$$RV_{i,t+h} = \mu_i + \phi_{J^+}\Delta J^{2+}_{i,t} + \phi_{J^-}\Delta J^{2-}_{i,t} + \phi_C BV_{i,t} + \phi_5 \overline{RV}_{5,i,t} + \phi_{22}\overline{RV}_{22,i,t} + \epsilon_{i,t}$$

	SPDR					Panel				
h	ϕ_{J+}	ϕ_{J-}	ϕ_C	ϕ_5	ϕ_{22}	ϕ_{J+}	ϕ_{J-}	ϕ_C	ϕ_5	ϕ_{22}
1	−0.151	−0.517	0.763	0.148	0.203	−0.150	−0.224	0.623	0.304	0.158
	(.333)	(.000)	(.000)	(.037)	(.000)	(.264)	(.005)	(.000)	(.000)	(.008)
5	−0.221	−0.144	0.242	0.319	0.314	−0.084	−0.173	0.413	0.224	0.289
	(.055)	(.003)	(.017)	(.017)	(.003)	(.481)	(.002)	(.000)	(.119)	(.004)
22	−0.189	−0.134	0.142	0.258	0.220	−0.068	−0.115	0.205	0.136	0.388
	(.014)	(.004)	(.051)	(.127)	(.052)	(.169)	(.000)	(.000)	(.216)	(.000)
66	0.030	−0.064	0.024	0.063	0.191	−0.025	−0.075	0.111	0.064	0.418
	(.544)	(.030)	(.535)	(.209)	(.028)	(.352)	(.001)	(.000)	(.164)	(.000)

Logs

$$\ln RV_{i,t+h} = \mu_i + \phi_{J^+}\,\%\Delta J^{2+}_{i,t} + \phi_{J^-}\,\%\Delta J^{2-}_{i,t} + \phi_C \ln BV_{i,t} + \phi_1^- \ln RS^-_{i,t} + \phi_5 \ln \overline{RV}_{5,i,t} + \phi_{22} \ln \overline{RV}_{22,i,t} + \epsilon_{i,t}$$

	SPDR					Panel				
h	ϕ_{J+}	ϕ_{J-}	ϕ_C	ϕ_5	ϕ_{22}	ϕ_{J+}	ϕ_{J-}	ϕ_C	ϕ_5	ϕ_{22}
1	−0.246	−0.527	0.399	0.377	0.168	−0.000	−0.284	0.334	0.351	0.266
	(.004)	(.000)	(.000)	(.000)	(.000)	(.979)	(.000)	(.000)	(.000)	(.000)
5	−0.281	−0.187	0.226	0.362	0.291	−0.149	−0.159	0.215	0.355	0.351
	(.009)	(.014)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
22	−0.437	−0.057	0.126	0.173	0.444	−0.135	−0.119	0.134	0.281	0.441
	(.000)	(.593)	(.001)	(.070)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
66	−0.161	−0.048	0.059	0.093	0.476	−0.097	−0.108	0.079	0.191	0.514
	(.317)	(.768)	(.068)	(.248)	(.000)	(.010)	(.001)	(.000)	(.000)	(.000)

Table 5: Model which includes further decomposes signed jump information into positive and negative jump variation where $\Delta J^{2+} = (RS^+ - RS^-)I_{[RS^+ > RS^-]}$, $\Delta J^{2+} = (RS^+ - RS^-)I_{[RS^+ < RS^-]}$, and its continuous component using Bipower Variation, BV. In the log specification the signed percentage jump variation is defined as $\%\Delta J^{2+}_{i,t} = \ln\left(1 + (\Delta J^2_{i,t})/RV_{i,t}\right)$ and $\%\Delta J^{2-}_{i,t} = \ln\left(1 + (\Delta J^2_{i,t})/RV_{i,t}\right)$. P-values are reported below parameter estimated in parentheses. The top panel contains results form the model in levels and the bottom contains estimates from the log specification. The parameter estimates on the decomposed sign variation are similar to the restricted version reported in table 4 with the exception of the lack of significance in the levels model for the S&P 500 SPDR.

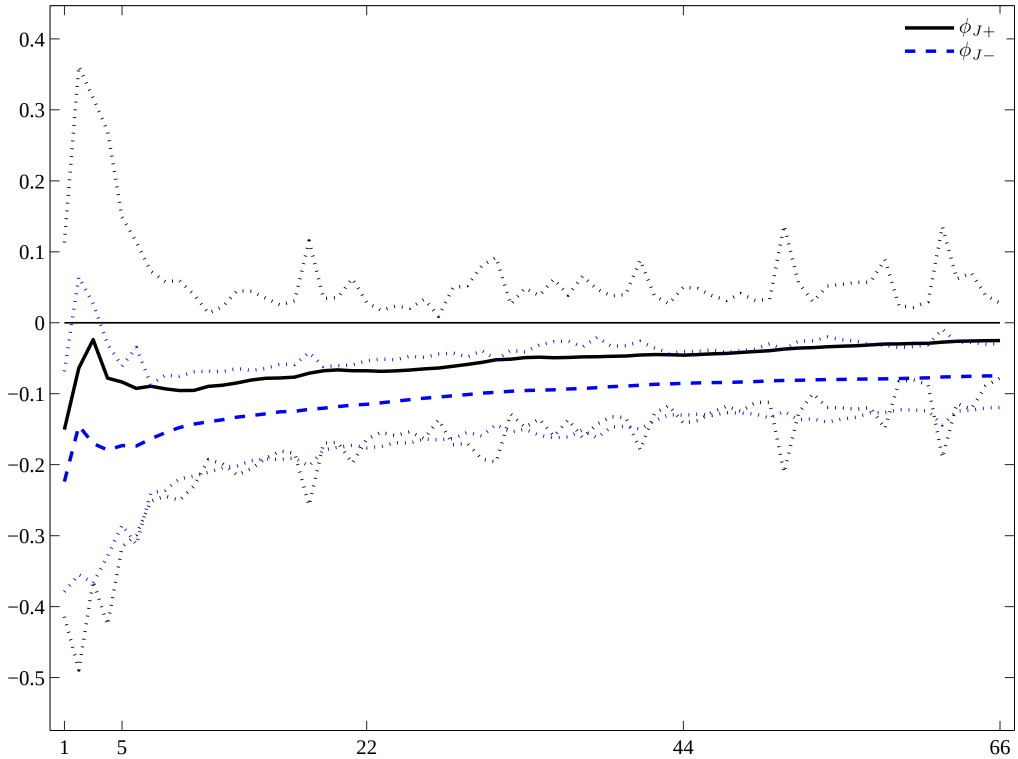

Figure 5: Fit coefficients from the levels model that included both signed jump variation and bipower variation, $RV_{i,t+h} = \mu_i + \phi^{J+}\left(RS^+_{i,t} - RS^-_{i,t}\right) I_{i,t} + \phi^{J-}\left(RS^+_{i,t} - RS^-_{i,t}\right)(1 - I_{i,t}) + \phi^C BV^-_{i,t} + \phi_5 \overline{RV}_{5,i,t} + \phi_{22}\overline{RV}_{22,i,t} + \epsilon_{i,t}.$, along with 95% confidence intervals for the panel of individual volatilities.

Individual Regressions in Levels (Percentage Rejection)

	Baseline				Jumps		Extended Jumps			
	RS^+		RS^-				J^+		J^-	
h	Sig. +	Sig -	Sig +	Sig -	Sig +	Sig -	Sig +	Sig -	Sig +	Sig -
1	0.162	0.010	0.857	0.010	0.010	0.486	0.010	0.067	0.019	0.429
5	0.019	0.086	0.571	0.010	0.010	0.390	0.010	0.095	0.010	0.324
22	0.019	0.181	0.467	0.010	0.010	0.352	0.048	0.114	0.010	0.305
66	0.048	0.229	0.324	0.019	0.019	0.267	0.076	0.048	0.029	0.238

Individual Regressions in Logs (Percentage Rejection)

	Baseline				Jumps		Extended Jumps			
	RS^+		RS^-				J^+		J^-	
h	Sig. +	Sig -	Sig +	Sig -	Sig +	Sig -	Sig +	Sig -	Sig +	Sig -
1	0.924	0.000	1.000	0.000	0.000	0.914	0.057	0.048	0.000	0.876
5	0.200	0.010	0.990	0.000	0.000	0.648	0.000	0.238	0.000	0.267
22	0.029	0.067	0.829	0.000	0.000	0.476	0.000	0.219	0.000	0.219
66	0.038	0.048	0.429	0.000	0.010	0.171	0.057	0.048	0.029	0.143

Table 6: This table contains percentage rejects for the Baseline, Jump, and Extended Jump model for individually fit models for the 105 S&P 100 constituents. Two columns are presented for each term, the left of which indicates the percentage of parameters that are significantly positive, the other indicating the percentage significantly negative. For example, in the baseline specification (equation 23) the positive Realized Semivariance was significantly positive in 16.2% of the series while the negative Realized Semivariance was significantly positive 85.7% of the time. These results confirm that the findings on the S&P 500 SPDR and the pooled panel are broadly reproducible in the cross-section of conditional variances.

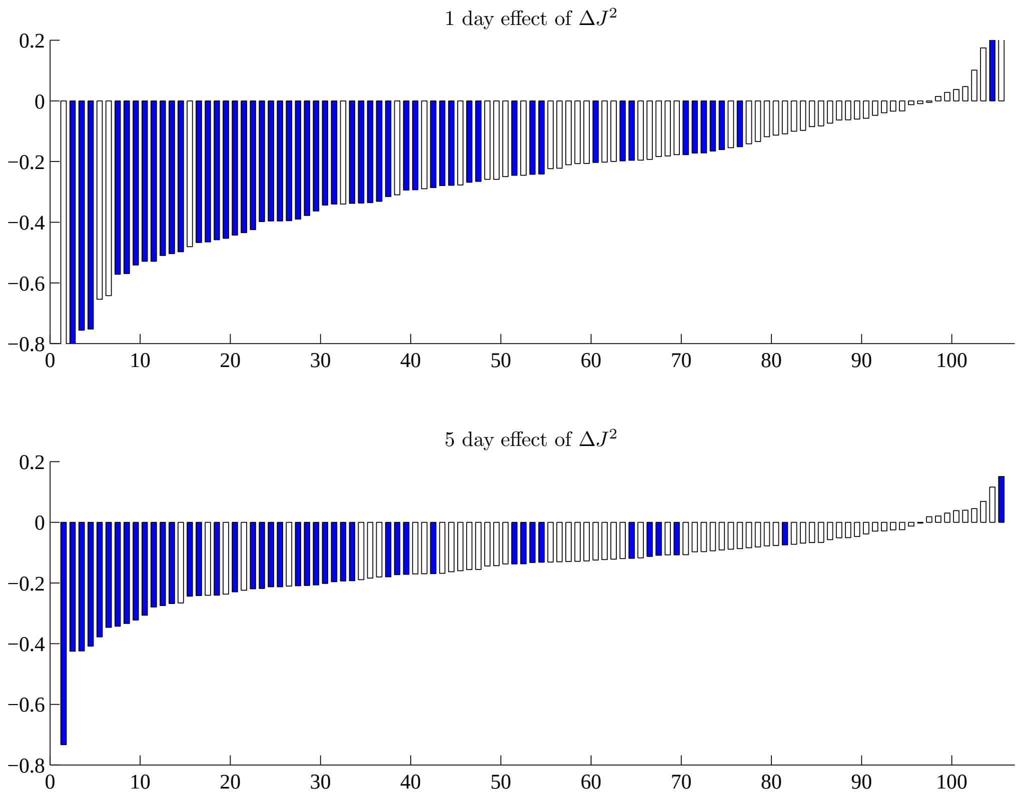

Figure 6: Sorted effects of the signed jump variation in the individual firm volatilities. The magnitude of the coefficient on ΔJ^2 is indicated as distance from the horizontal axis, and solid bars indicate significance at the 5% level. The majority of the parameter estimates are negative and significant.

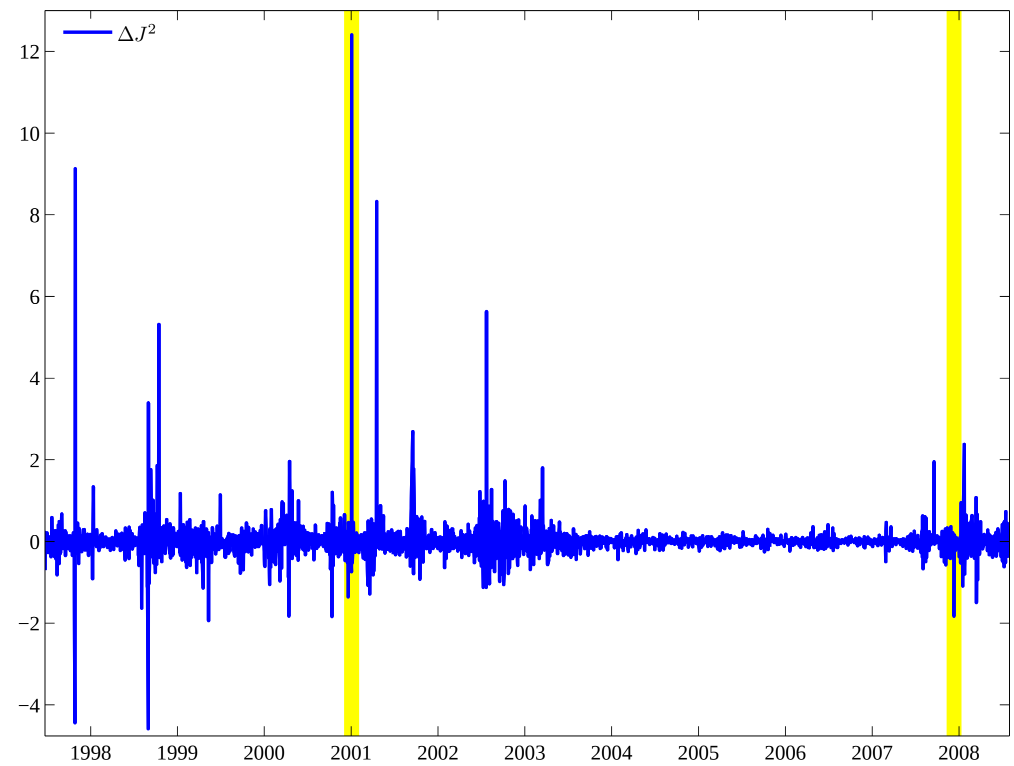

Figure 7: Time series of the signed jump variation ($\Delta J^2 = RS^+ - RS^-$) for the S&P 500 SPDR.

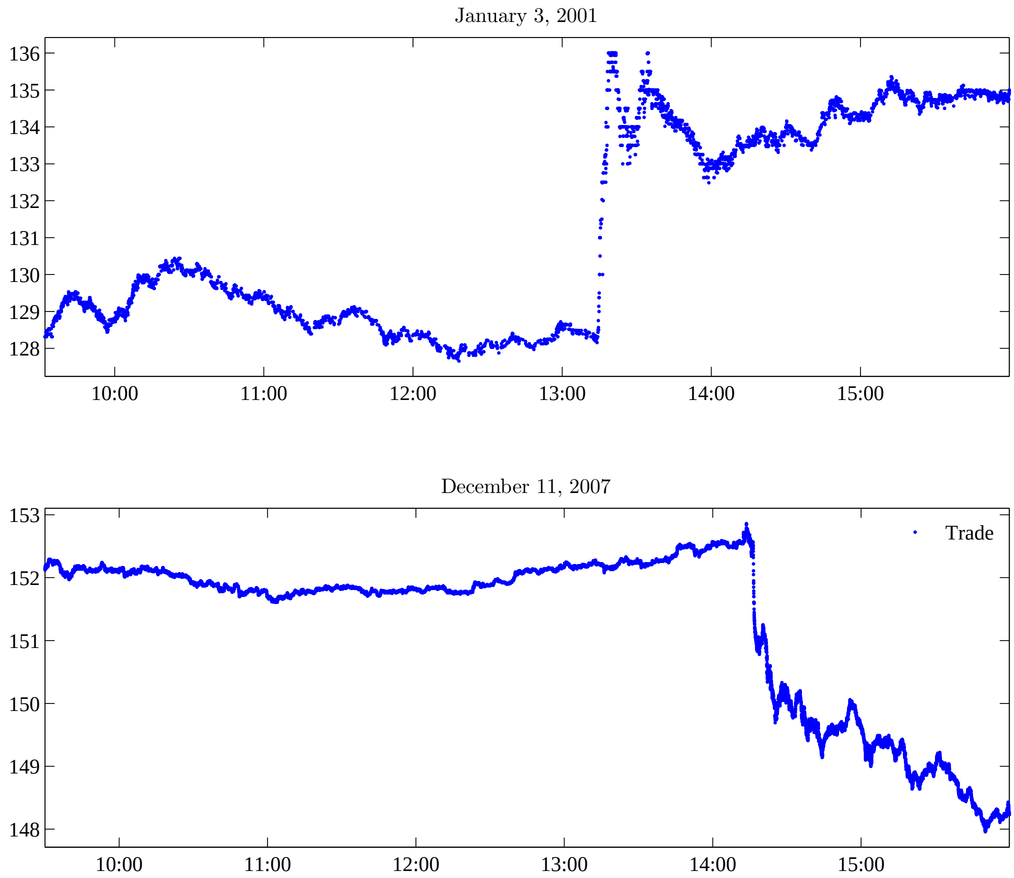

Figure 8: Intradaily prices for the S&P 500 SPDR on the two dates identified in figure 8.

References

Aït-Sahalia, Y. & Yu, J. (2009), 'High frequency market microstructure noise estimates and liquidity measures', *Annals of Applied Statistics . Forthcoming.*

Andersen, T. G. & Bollerslev, T. (1998), 'Answering the skeptics: Yes, standard volatility models do provide accurate forecasts', *International Economic Review* **39**(4), 885–905.

Andersen, T. G., Bollerslev, T. & Diebold, F. X. (2007), 'Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility', *The Review of Economics and Statistics* **89**(4), 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (1999), (understanding, optimizing, using and forecasting) realized volatility and correlation, New York University, Leonard N. Stern School Finance Department Working Paper Series 99-061, Northwestern University.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2001), 'The Distribution of Realized Exchange Rate Volatility', *Journal of the American Statistical Association* **96**(453), 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003), 'Modeling and Forecasting Realized Volatility', *Econometrica* **71**(1), 3–29.

Andersen, T. G., Bollerslev, T. & Meddahi, N. (2005), 'Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities', *Econometrica* **73**(1), 279–296.

Andersen, T. G., Bollerslev, T. & Meddahi, N. (2006), Realized volatility forecasting and market microstructure noise, Technical report, Kellogg School of Management, Northwestern University.

Andersen, T. G., Dobrev, D. & Schaumburg, E. (2008), Jump robust volatility estimation, Technical report, Northwestern Working Paper Series.

Ang, A., Chen, J. & Xing, Y. (2006), 'Downside Risk', *Review of Financial Studies* **19**(4), 1191–1239.

Bandi, F. & Russell, J. (2008), 'Microstructure noise, realized variance, and optimal sampling', *The Review of Economic Studies* **75**(2), 339–369.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2008*a*), 'Designing Realised Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise', *Econometrica . forthcoming.*

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2008*b*), 'Realised Kernels in Practice: Trades and Quotes', *Econometrics Journal . forthcoming.*

Barndorff-Nielsen, O. E., Kinnebrock, S. & Shephard, N. (2008), Measuring downside risk – realised semivariance, Technical report, Univeristy of Oxford.

Barndorff-Nielsen, O. E. & Shepard, N. (2001), 'Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics.', *Journal Of The Royal Statistical Society Series B* **63**, 167241.

Barndorff-Nielsen, O. E. & Shephard (2002*a*), 'Econometric analysis of realized volatility and its use in estimating stochastic volatility models', *Journal Of The Royal Statistical Society Series B* **64**(2), 253–280.

Barndorff-Nielsen, O. E. & Shephard, N. (2002*b*), 'Estimating quadratic variation using realized variance', *Journal of Applied Econometrics* **17**(5), 457–477.

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

Barndorff-Nielsen, O. E. & Shephard, N. (2006), 'Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation', *Journal of Financial Econometrics* **4**(1), 1–30.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**(3), 307–327.

Bollerslev, T., Law, T. H. & Tauchen, G. (2008), 'Risk, jumps, and diversification', *Journal of Econometrics* **144**(1), 234–256.

Chen, X. & Ghysels, E. (2008), News - good or bad - and its impact on volatility predictions over multiple horizons, Technical report, University of North Carolina at Chapel Hill - Department of Economics.

Chiriac, R. & Voev, V. (2007), Long memory modelling of realized covariance matrices, Technical report, University of Konstanz.

Christensen, K., Oomen, R. & Podolskij, M. (2008), Realized quantile-based estimation of integrated variance, Technical report, Working Paper.

Corsi, F. (2004), Simple long memory models of realized volatility, Technical report, University of Lugano.

Engle, R. (2002), 'New frontiers for ARCH models', *Journal of Applied Econometrics* **17**(5), 425–446.

Engle, R. F. (1982), 'Autoregressive conditional heteroskedasticity with esimtates of the variance of U.K. inflation', *Econometrica* **50**(4), 987–1008.

Engle, R. F., Shephard, N. & Sheppard, K. (2008), Fitting vast dimensional time-varying covariance models, Technical report, Oxford-Man Institute of Quantitative Finance.

Fleming, J., Kirby, C. & Ostdiek, B. (2003), 'The economic value of volatility timing using "realized" volatility', *Journal of Financial Economics* **67**(3), 473–509.

Glosten, L., Jagannathan, R. & Runkle, D. (1993), 'On the relationship between the expected value and the volatility of the nominal excess return on stocks', *Journal of Finance* **48**(5), 1779–1801.

Hogan, W. W. & Warren, J. M. (1974), 'Toward the development of an equilibrium capital-market model based on semivariance', *The Journal of Financial and Quantitative Analysis* **9**(1), 1–11.

Huang, X. & Tauchen, G. (2005), 'The Relative Contribution of Jumps to Total Price Variance', *Journal of Financial Econometrics* **3**(4), 456–499.

Lanne, M. (2006*a*), Forecasting realized volatility by decomposition, Technical report, European Univeristy Institute. EUI Working Paper ECO No. 2006/20.

Lanne, M. (2006*b*), A mixture multiplicative error model for realized volatility, Technical report, European Univeristy Institute. ECO No. 2006/3.

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

Lewis, A. L. (1990), 'Semivariance and the performance of portfolios with options', *Financial Analysts Journal* p. 6776.

Liu, C. & Maheu, J. M. (2005), Modeling and forecasting realized volatility: the role of power variation, Technical report, Univeristy of Toronto.

Müller, U., Dacorogna, M., Dav, R., Olsen, R., Pictet, O. & von Weizsacker, J. (1997), 'Volatilities of different time resolutions - analysing the dynamics of market components', *Journal of Empirical Finance* **4**, 213–39.

Nelson, D. (1991), 'Conditional heteroskedasticity in asset returns: a new approach', *Econometrica* **59**(2), 347–370.

Newey, W. K. & West, K. D. (1987), 'A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica* **55**(3), 703–708.

Oomen, R. C. A. (2005), 'Properties of Bias-Corrected Realized Variance Under Alternative Sampling Schemes', *Journal of Financial Econometrics* **3**(4), 555–577.

Patton, A. J. (2008), Data-Based Ranking of Realised Volatility Estimators, Technical report, University of Oxford.

Sortino, F. A. & Satchell, S. (2001), *Managing Downside Risk in Financial Markets*, Butterworth-Heinemann.

Todorov, V. & Tauchen, G. (2008), Volatility jumps, Technical report, Northwestern University.

Visser, M. P. (2008), Forecasting s&p 500 daily volatility using a proxy for downward price pressure, Technical report, Korteweg-de Vries Institute for Mathematics, University of Amsterdam.

Zakoian, J. M. (1994), 'Threshold heteroskedastic models', *Journal of Economic Dynamics and Control* **18**(5), 931–955.

Zhang, L., Mykland, P. & Aït-Sahalia, Y. (2004), A tale of two time scales: Determining integrated volatility with noisy high-frequency data. Forthcoming *Journal of the American Statistical Association.*