Measuring downside risk — realised semivariance

Neil Shephard

Oxford-Man Institute of Quantitative Finance
and
Department of Economics
University of Oxford

Joint with
Ole Barndorff-Nielsen
Aarhus University

and

Silja Kinnebrock
Maths and OMI, Oxford

January, 2008

Introduction

Economic risk is the risk of prices falling
Tradition of doing this just using negative returns

- semivariance,
- value at risk
- expected shortfall

— estimated using daily returns.
New measure of the variation of asset prices based on high frequency data.
It is called realised semivariance (RS).
Derive its limiting properties, relating it to quadratic variation and, in particular, negative jumps.

Realised terms

ABDL(2001) and BNS(2002) formalised realised variances (RV), establishing links these commonly used statistics to the quadratic variation process.
BNS(2004,2006) went inside the quadratic variation process and separate out components of the variation of prices into that due to jumps and that due to the continuous evolution. Bipower variation.
This work has prompted papers by, for example, ABD(2003), Huang and Tauchen (2005) and Lee and Mykland (2007) on the importance of this decomposition empirically in economics.
Surveys of this kind of thinking are provided by ABD(06) and BNS(2007), while a lengthy discussion of the relevant probability theory is given in Jacod(2007).

Quadratic variation

Realised variance (RV) estimates the ex-post variance of asset prices over a fixed time period.
We will suppose that this period is 0 to 1.
In our applied work it can be thought of as any individual day of interest.
Then RV is defined as

$$RV = \sum_{j=1}^{n} \left(Y_{t_j} - Y_{t_{j-1}} \right)^2 .$$

where $0 = t_0 < t_1 < \ldots < t_n = 1$ are the times at which (trade or quote) prices are available.

$$RV = \sum_{j=1}^{n} \left(Y_{t_j} - Y_{t_{j-1}} \right)^2 .$$

For arbitrage free-markets, Y must follow a semimartingale. This estimator converges as we have more and more data in that interval to the quadratic variation at time one,

$$[Y]_1 = \mathrm{p}\!-\!\lim_{n\to\infty} \sum_{j=1}^{n} \left(Y_{t_j} - Y_{t_{j-1}} \right)^2 ,$$

(e.g. Protter (2004)) for any sequence of deterministic partitions $0 = t_0 < t_1 < ... < t_n = 1$ with $\sup_j \{t_{j+1} - t_j\} \to 0$ for $n \to \infty$. This limiting operation is often referred to as "in-fill asymptotics" in statistics and econometrics

Being blind to signs

RV solely uses squares of the data, while the research of, for example, Black (1976), Nelson (1991), Glosten, Jagannathan and Runkle (1993) and Engle and Ng (1993) has indicated the importance of falls in prices as a driver of conditional variance.

The reason for this is clear, as the high frequency data becomes dense, the extra information in the sign of the data can fall to zero for some models.

Elegant framework in which to see this is where Y is a Brownian semimartingale

$$Y_t = \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s, \quad t \geq 0,$$

where a is a locally bounded predictable drift process and σ is a càdlàg volatility process – all adapted to some common filtration $\mathcal{F}_t$, implying the model can allow for classic leverage effects.

QV

$$Y_t = \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s, \quad t \geq 0,$$

For such a process

$$[Y]_t = \int_0^t \sigma_s^2 \mathrm{d}s,$$

and so

$$\mathrm{d}[Y]_t = \sigma_t^2 \mathrm{d}t,$$

which means for a Brownian semimartingale the QV process tells us everything we can know about the ex-post variation of Y. The signs of the returns are irrelevant in the limit — this is true whether there is leverage or not.

Sign of jumps

If there are jumps in the process there are additional things to learn than just the QV process. Let

$$Y_t = \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + J_t,$$

where J is a pure jump process. Then, writing jumps in Y as $\Delta Y_t = Y_t - Y_{t-}$, then

$$[Y]_t = \int_0^t \sigma_s^2 \mathrm{d}s + \sum_{s \leq t} (\Delta Y_s)^2,$$

and so QV aggregates two sources of risk. Even when we employ bipower variation (BNS(2004, 2006)), which allows us to estimate $\int_0^t \sigma_s^2 \mathrm{d}s$ robustly to jumps, this still leaves us with estimates of $\sum_{s \leq t} (\Delta J_s)^2$. This tells us nothing about the asymmetric behaviour of the jumps — which is important if we wish to understand downside risk.

Downside realised semivariance

We introduce the downside realised semivariances (RS^-)

$$RS^- = \sum_{j=1}^{t_j \leq 1} \left(Y_{t_j} - Y_{t_{j-1}} \right)^2 1_{Y_{t_j} - Y_{t_{j-1}} \leq 0},$$

where 1_y is the indicator function taking the value 1 of the argument y is true.
We will study the behaviour of this statistic under in-fill asymptotics.
In particular we will see that

$$RS^- \xrightarrow{p} \frac{1}{2} \int_0^t \sigma_s^2 \mathrm{d}s + \sum_{s \leq 1} (\Delta Y_s)^2 1_{\Delta Y_s \leq 0},$$

under in-fill asymptotics.
Hence RS^- provides a new source of information, one which focuses on squared negative jumps.

Upside realised semivariance

Of course the corresponding upside realised semivariance

$$\begin{aligned} RS^{+} &= \sum_{j=1}^{t_j \le 1} \left(Y_{t_j} - Y_{t_{j-1}} \right)^2 1_{Y_{t_j} - Y_{t_{j-1}} \ge 0} \\ &\xrightarrow{p} \frac{1}{2} \int_0^t \sigma_s^2 \mathrm{d}s + \sum_{s \le 1} (\Delta Y_s)^2 1_{\Delta Y_s \ge 0}, \end{aligned}$$

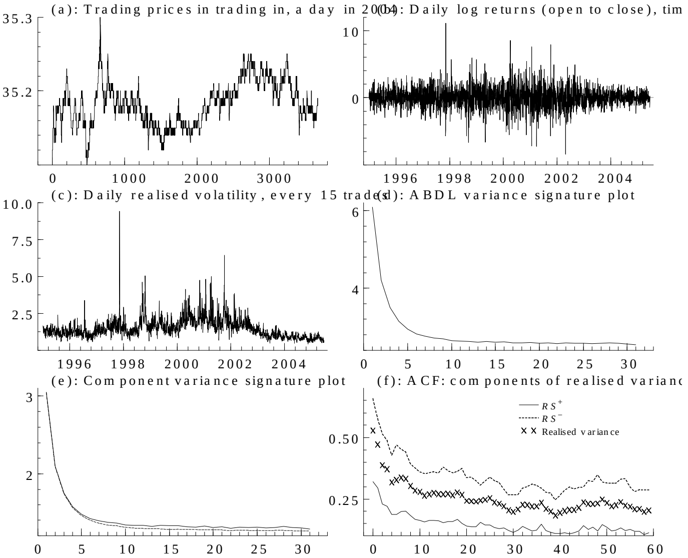
maybe of particular interest to investors who have short positions in the market (hence a fall in price can lead to a positive return and hence is desirable), such as hedge funds.

Of course,

$$RV = RS^{-} + RS^{+}.$$

Initial empirical features

Analysis of trades on General Electric (GE) carried out on the New York Stock Exchange from 1995 to 2005 (giving us 2,616 days of data). Notice by 2004 the tick size has fallen to one cent.


(a): Trading prices in trading in, a day in 2004
35.3
35.2
0
1000
2000
3000
(b): Daily log returns (open to close), tim
10
0
1996
1998
2000
2002
2004
(c): Daily realised volatility, every 15 trades
10.0
7.5
5.0
2.5
1996
1998
2000
2002
2004
(d): ABDL variance signature plot
6
4
0
5
10
15
20
25
30
(e): Component variance signature plot
3
2
0
5
10
15
20
25
30
(f): ACF: components of realised varianc
RS^+
RS^-
Realised variance
0.50
0.25
0
10
20
30
40
50
60

Information content

In the realised volatility literature, authors have typically worked out the impact of using realised volatilities on volatility forecasting using regressions of future realised variance on lagged realised variance and various other explanatory variables. Engle and Gallo (2006) prefers a different route, which is to add lagged realised quantities as variance regressors in Engle (2002) and Bollerslev (1986) GARCH type models of daily returns — the reason for their preference is that it is aimed at a key quantity, a predictive model of future returns, and is more robust to the heteroskedasticity inherent in the data. Typically when Engle generalises to allow for leverage he uses the Glosten, Jagannathan and Runkle (1993) (GJR) extension. This is the method we follow here. Throughout we will use the subscript i to denote discrete time.

Model based analysis

We model daily open to close returns $\{r_i; i = 1, 2, ..., T\}$ as

$$\begin{aligned} \mathrm{E}\left(r_i | \mathcal{G}_{i-1}\right) &= \mu, \\ h_i &= \mathrm{Var}\left(r_i | \mathcal{G}_{i-1}\right) = \omega + \alpha\left(r_{i-1} - \mu\right)^2 + \beta h_{i-1} \\ &\quad + \delta\left(r_{i-1} - \mu\right)^2 I_{r_{i-1} - \mu < 0} + \gamma' z_{i-1}, \end{aligned}$$

and then use a standard Gaussian quasi-likelihood to make inference on the parameters, e.g. Bollerslev & Wooldridge (1992). Here z_{i-1} are the lagged daily realised regressors and $\mathcal{G}_{i-1}$ is the information set generated by discrete time daily statistics available to forecast r_i at time $i - 1$.

Main results

Table 1 shows the fit of the GE trade data from 1995-2005.

Main results

ARCH type models and lagged realised semivariance and variance

	GARCH			GJR		
RS^{-}_{-1}	0.685 (2.78)			0.371 (0.91)		
RV_{-1}	−0.114 (−1.26)		0.228 (3.30)	0.037 (0.18)		0.223 (2.68)
ARCH	0.040 (2.23)	0.046 (2.56)	0.040 (2.11)	0.017 (0.74)	0.016 (1.67)	0.002 (0.12)
GARCH	0.711 (7.79)	0.953 (51.9)	0.711 (9.24)	0.710 (7.28)	0.955 (58.0)	0.708 (7.49)
GJR				0.055 (1.05)	0.052 (2.86)	0.091 (2.27)
Log-L	-4527.3	-4577.6	-4533.5	-4526.2	-4562.2	-4526.9

Table: Gaussian quasi-likelihood fit of GARCH and GJR models fitted to daily open to close returns on General Electric share prices, from 1995 to 2005. We allow lagged daily realised variance (RV) and realised semivariance (RS) to appear in the conditional variance. They are computed using every 15th trade. T-statistics, based on robust standard errors, are reported in small font and in brackets. Code: `GARCH_analysis.ox`

We start this section by repeating some of the theoretical story from Section 1.
Consider a Brownian semimartingale Y given as

$$Y_t = \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s, \tag{1}$$

where a is a locally bounded predictable drift process and σ is a càdlàg volatility process. For such a process

$$[Y]_t = \int_0^t \sigma_s^2 \mathrm{d}s,$$

and so $\mathrm{d}[Y]_t = \sigma_t^2 \mathrm{d}t$, which means that when there are no jumps the QV process tells us everything we can know about the ex-post variation of Y.

When there are jumps this is no longer true, in particular let

$$Y_t = \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + J_t, \tag{2}$$

where J is a pure jump process. Then

$$[Y]_t = \int_0^t \sigma_s^2 \mathrm{d}s + \sum_{s \le t} (\Delta J_s)^2,$$

and $\mathrm{d}[Y]_t = \sigma_t^2 \mathrm{d}t + (\Delta Y_t)^2$. Even when we employ devices like bipower variation (BNS(2004, 2006))

$$\{Y\}_t^{[1,1]} = \mu_1^{-2} \mathrm{p}\!-\!\lim_{n\to\infty} \sum_{j=2}^{t_j \le t} \left|Y_{t_j} - Y_{t_{j-1}}\right| \left|Y_{t_{j-1}} - Y_{t_{j-2}}\right|, \quad \mu_1 = \mathrm{E}\left|U\right|, \quad U \sim$$

we are able to estimate $\int_0^t \sigma_s^2 \mathrm{d}s$ robustly to jumps, but this still leaves us with estimates of $\sum_{s\le t} (\Delta J_s)^2$. This tells us nothing about the asymmetric behaviour of the jumps.

The empirical analysis we carry out throughout this paper is based in trading time, so data arrives into our database at irregular points in time. However, these irregularly spaced observations can be thought of as being equally spaced observations on a new time-changed process, in the same stochastic class, as argued by, for example, BNHLS(2006). Thus there is no intellectual loss in initially considering equally spaced returns

$$y_i = Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}, \quad i = 1, 2, ..., n.$$

$$y_i = Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}, \quad i = 1, 2, ..., n.$$

We study the functional

$$V(Y, n) = \sum_{i=1}^{\lfloor nt \rfloor} \begin{pmatrix} y_i^2 1_{\{y_i \geq 0\}} \\ y_i^2 1_{\{y_i \leq 0\}} \end{pmatrix}. \tag{3}$$

The main results then come from an application of some limit theory of Kinnebrock and Podolskij (2007) for bipower variation. This work can be seen as an important generalisation of Barndorff-Nielsen, Graversen, Jacod and Shephard (2006) who studied bipower type statistics of the form

$$\frac{1}{n} \sum_{j=2}^{n} g(\sqrt{n} y_j) h(\sqrt{n} y_{j-1}),$$

when g and h were assumed to be even functions. Kinnebrock and Podolskij (2007) give the extension to the uneven case, which is essential here.

Suppose BSM holds, then

$$V(Y,n) \xrightarrow{p} \frac{1}{2}\int_0^t \sigma_s^2 \mathrm{d}s \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Proof. Trivial application of Theorem 1 in Kinnebrock and Podolskij (2007).

Corollary Suppose

$$Y_t = \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + J_t,$$

holds, where J is a finite activity jump process then

$$V(Y,n) \xrightarrow{p} \frac{1}{2}\int_0^t \sigma_s^2 \mathrm{d}s \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \sum_{s \le t} \begin{pmatrix} (\Delta Y_s)^2 1_{\{\Delta Y_s \ge 0\}} \\ (\Delta Y_s)^2 1_{\{\Delta Y_s \le 0\}} \end{pmatrix}.$$

The above means that

$$(1,-1)\,V(Y,n) \xrightarrow{p} \sum_{s\leq t} (\Delta Y_s)^2\, 1_{\{\Delta Y_s \geq 0\}} - (\Delta Y_s)^2\, 1_{\{\Delta Y_s \leq 0\}},$$

the difference in the squared jumps. Hence this statistic allows us direct econometric evidence on the importance of the sign of jumps. Of course, by combining with bipower variation

$$V(Y,n) - \frac{1}{2}\int_0^t \sigma_s^2 \mathrm{d}s \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow{p} \sum_{s\leq t} \begin{pmatrix} (\Delta Y_s)^2\, 1_{\{\Delta Y_s \geq 0\}} \\ (\Delta Y_s)^2\, 1_{\{\Delta Y_s \leq 0\}} \end{pmatrix},$$

we can straightforwardly estimate the QV of just positive or negative jumps.

In order to derive a central limit theory we need to make two assumptions on the volatility process.

(H1). If there were no jumps in the volatility then it would be sufficient to employ

$$\sigma_t = \sigma_0 + \int_0^t a_s^* \mathrm{d}s + \int_0^t \sigma_s^* \mathrm{d}W_s + \int_0^t v_s^* \mathrm{d}W_s^*. \tag{4}$$

Here a^*, σ^*, v^* are adapted càdlàg processes, with a^* also being predictable and locally bounded. W^* is a Brownian motion independent of W.

(H2). $\sigma_t^2 > 0$ everywhere.

The assumption (H1) allows for flexible leverage effects, multifactor volatility effects, jumps, non-stationarities, intraday effects, etc. Kinnebrock and Podolskij (2007) also allow jumps in the volatility under the usual (in this context) conditions introduced by Barndorff-Nielsen, Graversen, Jacod, Podolskij and Shephard (2006) and discussed by, for example, Barndorff-Nielsen, Graversen,Jacod and Shephard (2006) but we will not detail this here.

Proposition 1. Suppose BSM, (H1) and (H2) holds, then

$$\sqrt{n}\left\{V(Y,n)-\frac{1}{2}\int_0^t \sigma_s^2 \mathrm{d}s \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\} \xrightarrow{Dst} V_t$$

where

$$\begin{aligned}
V_t &= \int_0^t \alpha_s(1)\,\mathrm{d}s + \int_0^t \alpha_s(2)\,\mathrm{d}W_s + \int_0^t \alpha_s(3)\,\mathrm{d}W_s', \\
\alpha_s(1) &= \frac{1}{\sqrt{2\pi}}\{2a_s\sigma_s + \sigma_s\sigma_s^*\}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \\
\alpha_s(2) &= \frac{2}{\sqrt{2\pi}}\sigma_s^2\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \\
A_s &= \frac{1}{4}\sigma_s^4\begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}, \\
\alpha_s(3)\,\alpha_s(3)' &= A_s - \alpha_s(2)\,\alpha_s(2)',
\end{aligned}$$

where $\alpha_s(3)$ is a 2×2 matrix. Here $W' \perp\!\!\!\perp (W, W^*)$, the Brownian motions which appears in the Brownian semimartingale BSM and (H1).

When we look at

$$RV = (1,1)\, V(Y,n),$$

then we produce the well known result

$$\sqrt{n}\left\{RV - \int_0^t \sigma_s^2 \mathrm{d}s\right\} \overset{Dst}{\rightarrow} \int_0^t 2\sigma_s^2 \mathrm{d}W_s^{'}$$

which appears in Jacod (1994) and BNS(2002).

Assume $a, \sigma \perp\!\!\!\perp W$ then

$$\sqrt{n}\left\{V(Y,n) - \frac{1}{2}\int_0^t \sigma_s^2 \mathrm{d}s \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$$
$$\stackrel{Dst}{\to} MN\left(\frac{1}{\sqrt{2\pi}}\int_0^t \{2a_s\sigma_s + \sigma_s\sigma_s^*\}\,\mathrm{d}s \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{4}\int_0^t \sigma_s^4 \mathrm{d}s \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}\right)$$

If there is no drift and the volatility of volatility was small then the mean of this mixed Gaussian distribution is zero and we could use this limit result to construct confidence intervals on these quantities. When the drift is not zero we cannot use this result as we do not have a method for estimating the bias which is a scaled version of

$$\frac{1}{\sqrt{n}}\int_0^t \{2a_s\sigma_s + \sigma_s\sigma_s^*\}\,\mathrm{d}s.$$

Of course in practice this bias will be small. The asymptotic variance of $(1,-1)\,V(Y,n)$ is $\frac{3}{n}\int_0^t \sigma_s^4 \mathrm{d}s$, but obviously not mixed Gaussian.

When the $a, \sigma \perp\!\!\!\perp W$ result fails, we do not know how to construct confidence intervals even if the drift is zero. This is because in the limit

$$\sqrt{n}\left\{V(Y,n) - \frac{1}{2}\int_0^t \sigma_s^2 \mathrm{d}s \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$$

depends upon W. All we know is that the asymptotic variance is again

$$\frac{1}{4n}\int_0^t \sigma_s^4 \mathrm{d}s \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}.$$

Notice, throughout the asymptotic variance of RS^- is

$$\frac{5}{4n}\int_0^t \sigma_s^4 \mathrm{d}s$$

so is less than that of the RV (of course it estimates a different quantity so perhaps this observations is not so particularly important). It also means the asymptotic variance of $RS^+ - RS^-$ is

$$\frac{3}{n}\int_0^t \sigma_s^4 \mathrm{d}s.$$

Noise

Suppose instead of seeing Y we see

$$X = Y + U,$$

and think of U as noise. Let us focus entirely on

$$\begin{aligned} \sum_{i=1}^{n} x_i^2 1_{\{x_i \leq 0\}} &= \sum_{i=1}^{n} y_i^2 1_{\{y_i \leq -u_i\}} + \sum_{i=1}^{n} u_i^2 1_{\{y_i \leq -u_i\}} + 2 \sum_{i=1}^{n} y_i u_i 1_{\{y_i \leq -u_i\}} \\ &\simeq \sum_{i=1}^{n} y_i^2 1_{\{u_i \leq 0\}} + \sum_{i=1}^{n} u_i^2 1_{\{u_i \leq 0\}} + 2 \sum_{i=1}^{n} y_i u_i 1_{\{u_i \leq 0\}}. \end{aligned}$$

If we use the framework of Zhou (1996), where U is white noise, uncorrelated with Y, with $\mathrm{E}(U) = 0$ and $\mathrm{Var}(U) = \omega^2$ then it is immediately apparent that the noise will totally dominate this statistic in the limit as $n \to \infty$.
Pre-averaging based statistics of Jacod, Li, Mykland, Podolskij and Vetter (2007) could be used here to reduce the impact of noise on the statistic.

Conclusions

There is much more empirical work in the paper, replicating this result.
Main points

- We can look inside QV in another way
- We can tease out features of negative jumps
- These seem to drive conditional variance of asset returns