3113F

Modelling Volatility During a Flash Crash

Abstract

Using data for the E-Mini S&P 500 futures contract, this paper investigates various aspects of intraday volatility in the days surrounding the 6 May 2010 Flash Crash in US equity markets. First, I calculate robust measures of realised volatility, which show that volatility was unusually high across the whole of 6 May. Next, I explore the lead-lag relationship between volatility and volume and find that there was a structural break in this relationship on the day of the Crash. Finally, I use GARCH and Dynamic Conditional Score models to create out-of-sample volatility forecasts for the Flash Crash.

7500 Words

1. Introduction

May 6 2010 was a day of extreme intraday volatility is US financial markets. In the four and a half minutes from 13:41:00 to 13:45:30 CT, the general markets declined by 5-6% and volumes in US equity instruments surged (CFTC and SEC, 2010). By 14:00:00 this price movement had largely reversed. This extreme event, entailing a rapid price drop and recovery, became known as the Flash Crash.

Motivated by the relative lack of studies that model volatility at the intraday level (cf. Engle and Sokalska, 2012), this paper conducts a high-frequency analysis of volatility in the E-Mini S&P 500 futures contract in the days surrounding the Flash Crash. It also investigates the role of the accompanying surge in trade volume in helping us to understand the dynamics of volatility around the event, and in helping us to obtain better forecasts. Given the evidence that up to 12 "mini flash crashes" occur every day (Farrell, 2013), this analysis is consequential for risk managers operating in the ultra-fast market ecology of today.

First, I compute various measures of realised volatility that are robust to sources of bias inherent in a high frequency analysis of volatility. The series computed here show that volatility was unusually high in the run up to the Crash, and that the spike in volatility from the Crash persisted till the end of the trading day. Additionally, these ex-post, non-parametric measures of volatility serve as ideal benchmarks against which forecasts can be evaluated.

Next, I empirically test theories of the lead-lag relationship between volatility and volume in the days surrounding the Crash. I find that while the volatility-volume relationship was contemporaneous in the three days preceding the Crash, there was a structural break that led to a lagged relationship on 6 May.

Finally, I model volatility using the GARCH (Bollerslev, 1986) and Dynamic Conditional Score (DCS) (Harvey, 2013; Creal et al., 2009, 2011) class of models. Using these models, I create one-step ahead out-of-sample forecasts following a real-time coefficient updating procedure. I find that DCS models outperform traditional GARCH models, both in- and out-of-sample. Meanwhile, the merits of volume in improving forecasting performance are found to be unexceptional.

The paper proceeds as follows: Section 2 provides a literature review; Section 3 describes the data; Section 4 computes measures of realised volatility; Section 5 accounts for seasonality in intraday volatility and volume; Section 6 tests the volatility-volume relationship; Section 7 estimates models of conditional volatility and creates forecasts; Section 8 concludes.

2. Literature review

Given the ubiquity of high-frequency trading today (Foresight, 2012), the ability to accurately forecast intraday volatility is crucial for risk managers. Hence, the relative lack of studies relating to intraday volatility measurement and forecasting (Engle and Sokalska, 2012) leaves a critical gap in the literature.

Nevertheless, there do exist recent studies that model volatility at high frequencies. For instance, Engle and Sokalska (2012) forecast volatility for over 2500 equities in a GARCH framework by using 10-minute returns and decomposing intraday volatility into daily, diurnal and stochastic components. Giot (2005) uses 15 and 30-minute returns and applies GARCH, RiskMetrics (J.P. Morgan, 1996) and high-frequency duration models to quantify market risk for three US stocks. Martens et al (2002) use 30-minute spot exchange rate data to assess how different methods for de-seasonalising returns affect the forecasting performance of the GARCH(1,1). Meanwhile, Nishimura and Sun (2015) estimate an FIEGARCH model using 5-minute returns from Chinese equity and futures markets, and find that volume positively affects volatility.

However, there are some drawbacks to these and similar studies. First, they rely on estimators of realised volatility that are not optimal at high frequencies. For instance, Engle and Sokalska (2012) evaluate their forecasts against 10-minute squared returns, while Martens et al (2002) use 30-minute absolute returns. Such simple estimators are not first-best in high-frequency settings, since observations may be subject to "microstructure noise" (Zhang et al., 2005).

Secondly, given that the sharpest decline in the Flash Crash spanned only four and a half minutes, modelling volatility using even 10-minute returns is likely to miss important dynamics of the Crash. Therefore, I use 3-minute returns for the E-mini. In my view, the reason other papers select lower frequencies is because the bias in simple realised volatility estimates due to microstructure noise is less acute at lower frequencies

(Zhang et al., 2005), meaning a lower frequency analysis facilitates the computation of acceptable realised volatility measures. The realised volatility estimators used here are more robust and are consistent under certain assumptions on the noise.

Furthermore, this paper contributes to literature around the Flash Crash. Although there exist studies that address causality for the Crash (Aldrich et al., 2017; Menkveld and Yueshen, 2015; CFTC and SEC, 2010); measure the level of informed trading around the event (Easley et al., 2011); study the behaviour of market intermediaries on the day (Kirilenko et al., 2017), I am not aware of any studies that attempt to forecast volatility over the event. In fact, excluding Giot (2005), I am not aware of any studies that model intraday volatility on days with extreme price swings.

3. Data

Since the Flash Crash originated in the June 2010 E-Mini S&P 500 futures contract (CFTC and SEC, 2010), I use transactions data for this contract between 3 – 6 May 2010. This data records the price and volume of each completed transaction, accurate to a one second time-stamp. Where there was more than one trade associated with a time-stamp, I took the median price.

The E-mini trades 24/7 electronically on the CME Globex, except for a daily 15-minute maintenance break at 15:15:00 CT. Meanwhile, US equity markets are open from 08:30:00 to 15:00:00 CT. Here, I only use data from the 405-minute interval 08:30:00 – 15:15:00 CT, as the majority of trading in the E-mini occurs in this period.

Also, in the analysis below, I assume that observations are recorded in evenly spaced intervals. For instance, if a given time interval does not contain any trades, the most recent price from a previous interval is used.

To illustrate the data, Figure 3.1 plots the price series on 6 May 2010. The figure also plots volume (in 1-minute bins) for 6 May and average volume for 3 – 5 May.

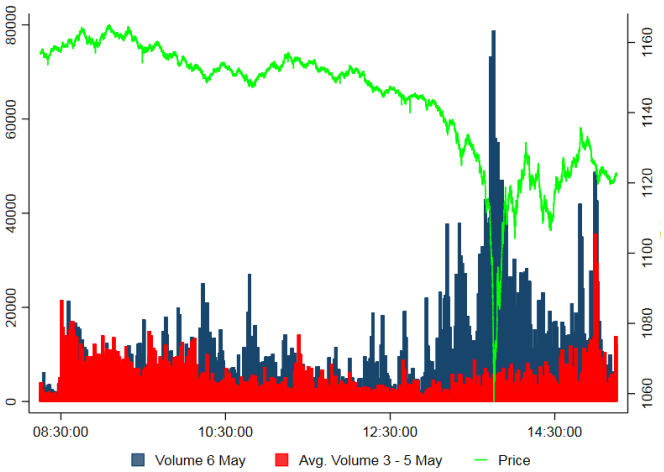

Figure 3.1: Volume measured on left axis, price on right

4. Realised volatility

Theory

Realised volatility is best motivated as a measure of return variation in a continuous time setting.

Suppose that the logarithmic price p_t of a financial security assumes the following continuous time diffusion process of semi-martingale type (McAleer and Medeiros, 2008):

$$p_t = \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s) \tag{4.1}$$

where the drift μ is a predictable and finite variation process, $\sigma(t) > 0$ is a càdlàg volatility process and W is a standard Brownian motion. Then, our object of interest is the integrated volatility (IV_t):

$$IV_t = \int_{t-1}^t \sigma^2(s)ds \tag{4.2}$$

i.e. the amount of price variation at time t accumulated over the interval $[t-1, t]$. Crucially, IV_t is latent and must be estimated.

In practice, the price is observed at n discrete points in the interval $[t-1, t]$, such that we have a sequence of deterministic partitions $t-1 = t_1 < t_2 < \cdots < t_n = t$, with $\sup_j\{t_{j+1} - t_j\} \to 0 \; for \; n \to \infty$ (Barndorff-Nielsen et al., 2008). In this case, the realised volatility, RV_t^n, defined as the sum of squared returns:

$$RV_t^n = \sum_{j=1}^{t_j \le t} r_{t_j}^2 = \sum_{j=1}^{t_j \le t} \left(p_{t_j} - p_{t_{j-1}} \right)^2 \tag{4.3}$$

provides a natural estimator for IV_t.

In fact, Andersen et al. (2003) show that given the semi-martingale representation of prices in (4.1), realised volatility is a consistent estimator of integrated volatility in the absence of microstructure noise (defined below). That is:

$$\operatorname*{plim}_{n \to \infty} RV_t^n = \operatorname*{plim}_{n \to \infty} \sum_{j=1}^{t_j \le t} \left(p_{t_j} - p_{t_{j-1}} \right)^2 = \int_{t-1}^{t} \sigma^2(s) ds$$

However, if the price process contains discontinuous "jumps", then realised volatility is not a consistent estimator for IV_t. A jump diffusive process is defined as:

$$p_t = \int_0^t \mu(s) ds + \int_0^t \sigma(s) dW(s) + \sum_{0 \le s < t} \kappa(s) \tag{4.4}$$

where the summation on the right compiles the effect of jumps, such that $\kappa(t)$ is non-zero only at discrete jumps.

Now, an important caveat to the above is that prices sampled at times t_j are usually observed with an error, termed market microstructure noise[1]. Explicitly, observed prices take the form:

$$\tilde{p}_{t_j} = p_{t_j} + u_{t_j} \tag{4.5}$$

where p_{t_j} is the true, latent price in (4.1), and u_{t_j} is microstructure noise. One consequence of (4.5) is that the presence of microstructure noise may lead to a large bias in realised volatility estimates and non-robustness to changes in sampling frequency n (Zhang et al., 2005). Zhang et al. also show that this bias is larger the higher the sampling frequency. Sampling sparsely would help to reduce this bias, but would increase variance (Barndorff-Nielsen et al., 2002); creating a bias-variance trade-off.

Given these drawbacks to the standard realised volatility estimator, I consider three alternatives.

[1] Madhavan (2000) discusses potential sources of such noise, such as price discreteness or bid-ask bounce effects

Alternative estimators

Here, I compute alternative measures of realised volatility. Since these measures are ex-post and non-parametric, they provide a benchmark against which our forecasts in Section 7 are evaluated (Andersen et al., 2006). They also enable us to document volatility in the E-Mini on 6 May.

First, I compute the Two-Scales Realised Volatility (TSRV) estimator of Zhang et al. (2005). Next, I estimate MedRV, the jump robust estimator of Andersen et al. (2012). Finally, I compute a measure based on realised kernels (KRV), developed by Barndorff-Nielsen et al. (2008). The purpose for computing multiple estimates is robustness.

Throughout this section, volatility is calculated for 3-minute intervals, and prices are sampled every 30 seconds.

I calculate volatility over 3-minute intervals because it was judged, after some experimentation, that this length of interval provided the best trade-off between maximising the number of data points during the Crash and obtaining volatility estimates that are not subject to intervals of zero return observations.

Prices were sampled every 30 seconds because the empirical analyses of Awartani et al. (2009) and Pigorsch et al. (2012) show that for the E-mini, the i.i.d assumption for the noise u_{t_j} is satisfied for sampling frequencies as high as 30 seconds. Since both the TSRV and KRV measures below are consistent under i.i.d noise, sampling every 30 seconds was deemed appropriate.

TSRV

The TSRV estimator is based on the subsampling approach introduced by Zhang et al. (2005). Subsampling involves averaging over several RV series calculated by sampling sparsely over high frequency samples.

More precisely, I compute a total of 30 realised volatility series by sampling prices every 30 seconds and moving forward the time of the first observation in each series by 1 second increments. The left term in (4.6) is the average of these estimates. This subsampling procedure ensures that all the data is used regardless of sampling frequency, which improves efficiency.

The term on the right in (4.6) is obtained by calculating RV using all available observations. This serves as a bias correction term.

The TSRV estimator is:

$$TSRV_t^{(n,n_1,\dots n_K)} = \frac{1}{K}\sum_{k=1}^{K} RV_t^{(k,n_k)} - \frac{\bar{n}}{n} RV_t^{(all)} \tag{4.6}$$

Where K is the number of RV series generated by the subsampling procedure, n_k is the sampling frequency used in the RV calculation for subsample k, n is the sampling frequency used for $RV^{(all)}$, and $\bar{n} = (1/K)\ \sum_{k=1}^{K} n_k$. Here, we set $K = 30$, $n_k = 6\ \forall k$ and $n = 180$.

This TSRV estimator is approximately unbiased under the conditions that u_{t_j} is an i.i.d white noise process and that u_{t_j} is independent of the latent price p_{t_j} in (4.5) (Aït-Sahalia et al., 2012).

MedRV

The jump robust MedRV estimator of Andersen et al. (2012) sums the squares of returns obtained from a 3-period rolling median. The median operator removes returns affected by a large jump. Furthermore, MedRV is consistent for integrated volatility (in the absence of microstructure noise) if prices follow the jump diffusion process in (4.4), which is not true for TSRV and KRV.

The MedRV estimator is:

$$MedRV_t = \frac{\pi}{6 - 4\sqrt{3} + \pi}\left(\frac{n}{n-2}\right)\sum_{i=2}^{n-1} med\left(\left|r_{t_{j-1}}\right|, \left|r_{t_j}\right|, \left|r_{t_{j+1}}\right|\right)^2 \tag{4.7}$$

The subsampling procedure was also applied to MedRV to increase efficiency.

Kernel RV

Barndorff-Nielsen et al. (2008) introduced a class of consistent, kernel based estimators for IV_t, known as realised kernels. For use in practice, Barndorff-Nielsen et al. (2009) suggest an estimator based on non-flat-top realised kernels (Pigorsch et al., 2012):

$$KRV_t^{(n,H)} = RV_t^n + \sum_{h=1}^{H} k\left(\frac{h}{H}\right)(\gamma_h + \gamma_{-h}) \qquad where\ \gamma_h = \sum_{j=1}^{n} r_{t_j} r_{t_{j-h}} \tag{4.8}$$

Where $k(x)$ is a kernel weight function, γ_h is the hth realised autocovariance, and H is the bandwidth.

This estimator can accommodate serial correlation in the noise u_{t_j}, and non-zero correlation between u_{t_j} and p_{t_j} in (4.5). However, these properties come at the expense of a small asymptotic bias and a sub-optimal convergence rate (Pigorsch et al., 2012).

Barndorff-Nielsen et al. (2009) suggest using the Parzen kernel, since it is smooth and non-negative:

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & for\ 0 \leq x < 1/2 \\ 2(1-x)^3 & for\ 1/2 \leq x \leq 1 \\ 0 & for\ x > 1 \end{cases}$$

Now, the optimal bandwith, H^*, is calculated as follows:

$$H^* = c^* \xi^{4/5} n^{3/5} \quad with \quad c^* = \left(\frac{k''(0)^2}{k_0} \right)^{1/5} \quad and\ \xi^2 = \frac{\omega^2}{\sqrt{t \int_0^t \sigma^4(s) ds}}$$

where $\omega^2 = Var\left(u_{t_j}\right)$ and $\int_0^t \sigma^4(s) ds$ is known as the Integrated Quarticity. For the Parzen Kernel, $c^* = 3.5134$.

Barndorff-Nielsen et al. (2009) mention that it is not crucial to obtain a consistent estimator for ξ^2, and therefore recommend:

$$\hat{\xi}^2 = \frac{\widehat{\omega}^2}{RV_t^n} \tag{4.9}$$

I calculate RV_t^n using 30-second returns and subsampling.

To estimate ω^2, I use the estimator of Zhang et al. (2005):

$$\widehat{\omega}^2 = \frac{1}{2n} RV^{(all)} \tag{4.10}$$

Furthermore, there do exist subsampling procedures for kernel based estimators, but these are harmful when smooth kernels, such as the Parzen, are used (Barndorff-Nielsen et al., 2007).

Results

Figure 4.1 superimposes the three estimated volatility series. Panel (a) plots the series for 3 – 5 May, while panel (b) plots the series for 6 May. In general, the three series agree well, except that MedRV is noisier.

To avoid carrying three separate series and for robustness, I average them into one, labelled ARV (Averaged Realised Volatility), which is used for the remaining analysis.

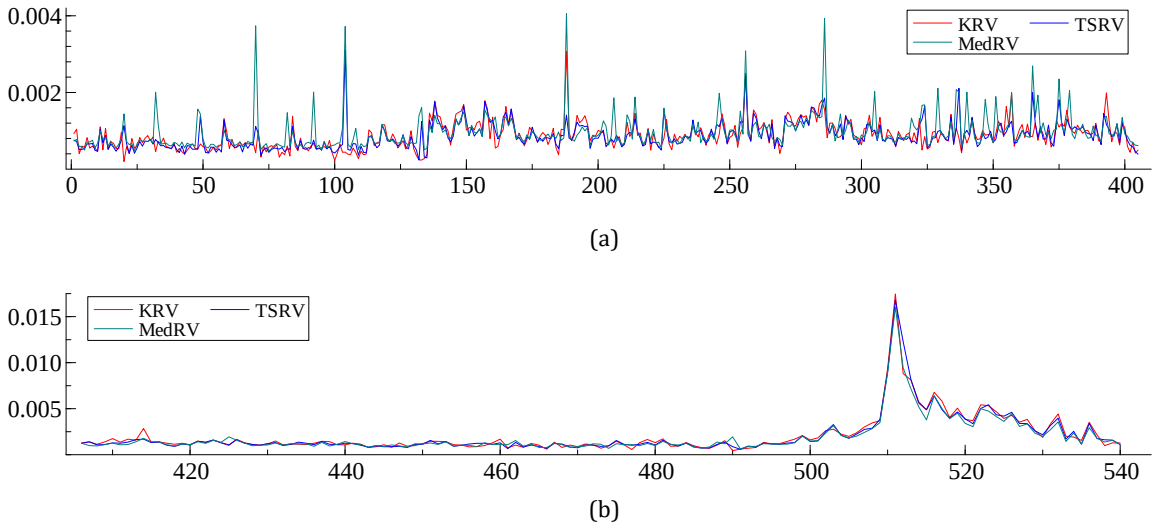

Figure 4.1: Y-axis measured in standard deviations, X-axis in 3-minute intervals

Figure 4.2 superimposes the *ARV* series for 6 May onto the *ARV* series obtained by averaging across 3 to 5 May in each 3-minute interval. The figure splits the trading day such that panel (a) plots the series from the start of the trading day to just before the onset of the Crash (08:30:00 – 13:30:00) and panel (b) plots the remaining day (13:30:00 – 15:15:00).

Panel (a) shows that volatility was higher in the run up to the Flash Crash than the average for 3 – 5 May, while panel (b) shows that the spike in volatility from the Crash persisted till the end of the trading day.

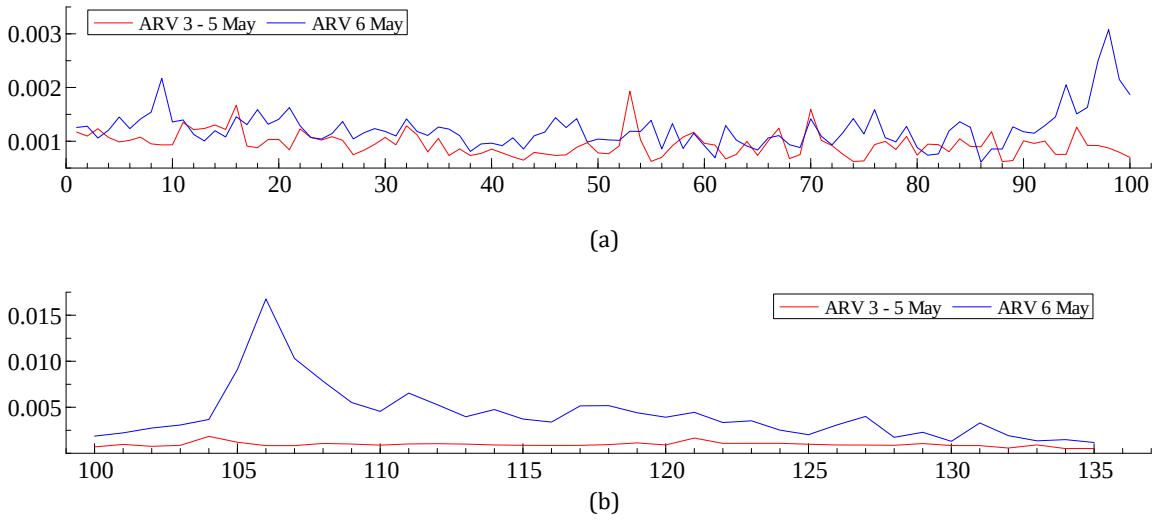

Figure 4.2: Y-axis measured in standard deviations, X-axis in 3-minute intervals

Having obtained robust realised volatility estimates, I now account for seasonality in intraday volatility and volume before proceeding with the econometric analyses of Sections 6 and 7.

5. Seasonality

Here, I estimate the seasonal component in intraday volatility and volume.

Volatility

I use the two-step method of Andersen and Bollerslev (1997, 1998) to compute the seasonal component of intraday volatility in the E-mini.

Assume the following model for returns:

$$r_{t,m} - E(r_{t,m}) = \sigma_{t,m}.s_{t,m}.z_{t,m} \qquad \sigma_{t,m}, s_{t,m} > 0 \tag{5.1}$$

where $r_{t,m}$ is now the return in 3-minute interval m on day t, $z_{t,m}$ is an i.i.d error term with mean zero and variance one, $s_{t,m}$ is the seasonal component, and $\sigma_{t,m}$ represents the underlying volatility component.

Note, the ARV series in Section 4 is an estimate of the total volatility $\sigma_{t,m}.s_{t,m}$. The additional model assumption in (5.1) means that the seasonal component $s_{t,m}$ and underlying volatility $\sigma_{t,m}$ can be separately identified.

Now, squaring, taking logs and rearranging (5.1), define:

$$x_{t,m} \equiv 2\log|r_{t,m} - E(r_{t,m})| - \log(\sigma_{t,m}^2) = c + 2\log(s_{t,m}) + \eta_{t,m} \tag{5.2}$$

where $\eta_{t,m} = \log(z_{t,m}^2) - E(\log(z_{t,m}^2))$ and $c = E(\log(z_{t,m}^2))$.

The first step is to obtain an estimate for $x_{t,m}$. This is done by replacing $E(r_{t,m})$ by the sample mean of returns, $\bar{r}$, and $\sigma_{t,m}^2$ by a reasonable estimate $\hat{\sigma}_{t,m}^2$.

The second step is to estimate the following Flexible Fourier Form (FFF) regression by OLS:

$$\hat{x}_{t,m} = f(\theta, t, m) + u_{t,m} \tag{5.3}$$

with

$$f(\theta,t,m) = \sum_{j=0}^{J} \sigma_t^j \left[\mu_{0j} + \mu_{1j}\frac{m}{M_1} + \mu_{2j}\frac{m^2}{M_2} + \sum_{i=1}^{D} \lambda_{ij} I_{m=d_i} + \sum_{i=1}^{P} \left(\gamma_{pj}\cos\left(\frac{pm2\pi}{M}\right) + \delta_{pj}\sin\left(\frac{pm2\pi}{M}\right)\right)\right]$$

where $M_1 \equiv \frac{1}{M}\sum_{i=1}^{M} i = (M+1)/2$; $M_2 \equiv \frac{1}{M}\sum_{i=1}^{M} i^2 = (M+1)(M+2)/6$; $\theta = (\mu_{0j}, \mu_{1j}, \lambda_{ij}, \gamma_{pj}, \delta_{pj})$ is the vector of parameters to be estimated and σ_t is daily volatility.

In (5.3), the $J+1$ flexible Fourier forms are parametrised by a polynomial and P sinusoids. The time specific dummies, $I_{m=d_i}$, accommodate intraday intervals that do not conform to a smooth pattern.

Here, I estimate the FFF using only 3 – 5 May. I do not include the day of the Crash in the estimation sample because this exogenous extreme event would dominate the underlying seasonality. I therefore assume that the seasonal pattern was constant across all days. Another reason for excluding 6 May is that I aim to create out-of-sample forecasts on this day in Section 7, meaning that we cannot use future information for estimation. An implication is that we must set $J = 0$ in (5.3), a point made by Martens et al. (2002). Omitting σ_t from the FFF also means that $\ln(\sigma_t^2)$ is withdrawn from the definition of $x_{t,m}$. Andersen and Bollerslev (1998) show that this is inconsequential.

For parsimony, I set $P = 1$. Additionally, a dummy which equals one for the last 15 minutes of the daily sample (15:00:00 – 15:15:00) is included to account for the close of equity markets.

Now, the intraday seasonal component is estimated as:

$$\hat{s}_{t,m} = \exp\left(\frac{\hat{\hat{x}}_{t,m}}{2}\right) \tag{5.4}$$

Then, the average of $\hat{s}_{t,m}$ is taken across 3 – 5 May to give $\hat{s}_m$, which is the estimated seasonality that we assume to be constant across all days.

Also, the FFF can be estimated directly in terms of $\left(r_{t,m} - \bar{r}\right)^2$ instead of $\hat{x}_{t,m}$, in which case $\hat{s}_{t,m} = \hat{f}(\theta, t, m)^{1/2}$. This specification is also estimated.

Volume

To estimate the seasonality in intraday volume, I use a "volume factors" approach inspired by Taylor and Xu (1997)

To define the intraday volume factors v_m, first suppose that

$$V_{t,m} = v_m . V_t \qquad with \quad \frac{1}{M}\sum_{m=1}^{M} v_m = 1 \tag{5.5}$$

where $V_{t,m}$ is volume in 3-minute interval m on day t, V_t is total volume on day t and the intraday volume factors v_m have mean one. Thus, $M^{-1}.v_m$ is the proportion of a day's trading volume that occurs in interval m, assuming the factors are constant across all days t.

The volume factors are then estimated as:

$$\hat{v}_m = M.\frac{\sum_t V_{t,m}}{\sum_t \sum_{m=1}^{M} V_{t,m}} \tag{5.6}$$

As for the FFF, these are estimated using only 3 – 5 May.

Results

Table 5.1: Flexible Fourier Form

FFF	μ_0 (p-val)	μ_1 (p-val)	μ_2 (p-val)	γ_1 (p-val)	γ_2 (p-val)	λ_1 (p-val)
$\hat{x}_{t,m}$ (n=405, R^2 = 0.0105)	-14.86 (0.000)	-1.958 (0.760)	0.256 (0.909)	-0.269 (0.835)	-0.399 (0.478)	1.014 (0.536)
$(r_{t,m} - \bar{r})^2$ (n=405, R^2 = 0.0563)	2.40x10^{-6} (0.000)	-4.29x10^{-6} (0.033)	1.38x10^{-6} (0.049)	-4.56x10^{-7} (0.260)	5.49x10^{-8} (0.755)	-1.07x10^{-6} (0.038)

Table 5.1 reports estimation results for the FFF. The log specification fails to pick up any periodicity and predicts that the best fit is a constant. Hence, I proceed with the $(r_{t,m} - \bar{r})^2$ specification, which picks up a U-shaped seasonality (Figure 5.1a).

The volume factors, graphed in Figure 5.1b, also exhibit a U-shape.

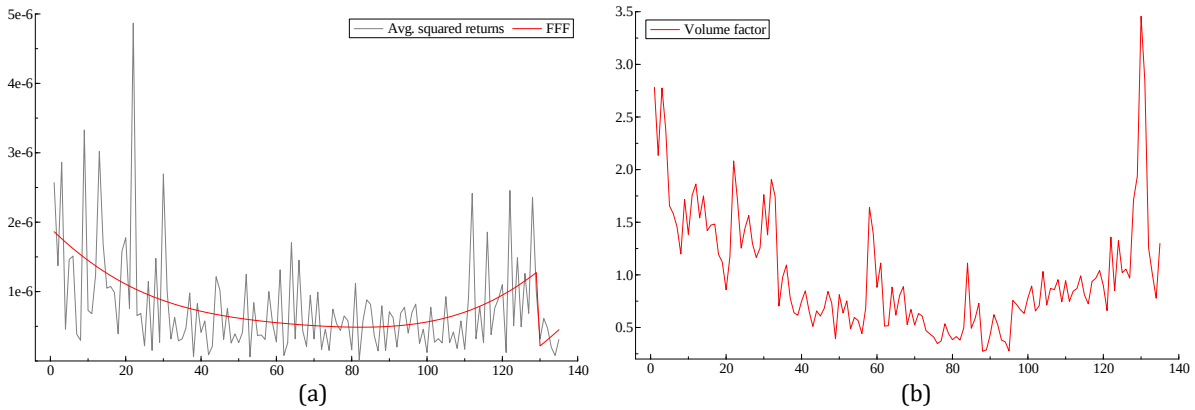

Figure 5.1

De-seasonalised returns are obtained by dividing the raw series by $\hat{s}_m$:

$$r_{t,m}^{ds} = \frac{r_{t,m}}{\hat{s}_m}$$

Similarly, volatility is de-seasonalised by dividing by $\hat{s}_m^2$. Volume is de-seasonalised by dividing by $\hat{v}_m$.

Finally, a point about notation: in the rest of this paper, the m subscript is suppressed, so intraday returns are written as r_t rather than $r_{t,m}$.

6. Volatility-volume relationship

In this section, I test theories of the lead-lag relationship between volatility and volume in the days surrounding the Crash. I also test whether this relationship remained stable in our sample.

This section is motivated by Figure 3.1, which shows that the Flash Crash was accompanied by a surge in trading volume. Hence, it is interesting to examine the role of volume in the dynamics the Crash.

Additionally, this section adds to the existing literature which empirically tests the volatility-volume relationship (e.g. Darrat et al., 2003; Xu et al., 2006; Mougoué and Aggarwal, 2011); the uniqueness of the current analysis is that we test this relationship in the context of an extreme event using a robust measure of realised volatility at the intraday level.

Theory

In a market with asymmetric information, market microstructure theory posits a positive volatility-volume relationship as follows: information arrival into the market generates trading volume as traders with the information advantage look to profit. By observing this trading activity, markets makers infer the new information and update prices (Easley and O'Hara, 1987).

In a high frequency setting, information arrival may be frequent. That is, for market makers who have short holding periods, any information relating to general trading conditions or liquidity demand over that period will be relevant for the price of assets, not just their fundamentals (Easley et al., 2012)

Here, I test the nature of the information arrival process between 3 – 6 May. The two relevant empirical hypotheses are the Sequential Information Arrival Hypothesis (SIAH) (Copeland, 1976) and the Mixture of Distributions Hypothesis (MDH) (Clark, 1973; Andersen, 1996)

In the SIAH, new information reaches agents in a sequential fashion, so that intermediate equilibria are reached before reaching the full information equilibrium. This sequential information assimilation means that lagged volume may have predictive ability for current volatility and vice-versa.

In the MDH, prices and volume are jointly dependent on information arrival. The model predicts that all agents receive new information simultaneously, so transitions to new equilibria are instant. This implies a contemporaneous volatility-volume relationship.

Hence, our analysis of the volatility-volume relationship provides insight into the information assimilation process and market efficiency around the Flash Crash.

Econometric Framework

Here, the de-seasonalised volatility and volume series are used to avoid any spurious correlation.

Summary statistics of the ARV and log volume series (not de-seasonalised) are presented in Table 6.1. Unit-root tests are conducted on de-seasonalised series.

Table 6.1: Summary statistics

	ARV	Log Volume
Mean	2.78×10^{-6}	9.73
Standard deviation	1.41×10^{-5}	0.78
Observations	540	540
Augmented DF[$]	-11.34***	-4.45***
Phillips Perron[$]	-11.22***	-8.28***
DF-GLS[$]	-10.41***	-2.60***

[$]Unit root tests are conducted on de-seasonalised series. Test statistics are reported. Lags in ADF and DF-GLS selected using the Schwarz Information Criterion (SIC). *** p<0.01, ** p<0.05, * p<0.1

Model

To study the volatility-volume relationship, the following Vector Autoregression model is estimated

$$\begin{bmatrix} ARV_t \\ \log(V_t) \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \sum_{i=1}^{q} \begin{bmatrix} \beta_{11,i} & \beta_{12,i} \\ \beta_{21,i} & \beta_{22,i} \end{bmatrix} \begin{bmatrix} ARV_{t-i} \\ \log(V_{t-i}) \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \quad (6.1)$$

and the appropriate Granger and contemporaneous causality tests are conducted, as below. I also conduct a system level Chow breakpoint test to examine whether there was a structural break on the day of the Crash.

The lag order q was set to 20 to allow for a lagged relationship between volatility and volume up to one hour, as in Darrat et al. (2003)[2].

For estimation, the JMulTi software of Lütkepohl and Krätzig (2004) is used.

Granger Causality

To test whether, say, volume Granger causes volatility, a Wald test with the null

$$\beta_{12,i} = 0 \ \ \forall i \tag{6.2}$$

is used. Rejection of this null suggests that lags of volume can help predict current volatility and therefore that volume Granger causes volatility. Similarly, the null for testing whether volatility Granger causes volume is $\beta_{21,i} = 0 \ \ \forall i$.

Hence, evidence of Granger causality supports the SIAH.

Contemporaneous causality

A contemporaneous volatility-volume relationship can be characterised by non-zero covariance between $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ – the null hypothesis is therefore (Lütkepohl, 2005):

$$E(\varepsilon_{1,t}\varepsilon_{2,t}) = 0 \tag{6.3}$$

which is tested against the alternative of non-zero covariance using the Wald test in Lütkepohl (2005), Section 3.6.3. Rejection of the null is evidence for the MDH.

Since the MDH imposes a positive volatility-volume relationship, $Corr(\hat{\varepsilon}_{1,t}, \hat{\varepsilon}_{2,t})$ from the estimated covariance matrix is also reported as a check.

The above tests are conducted following estimation of (6.1) on three separate samples: full sample; 3–5 May only; 6 May only. This allows for the possibility that the volatility-volume relationship changed on the day of the Crash. This is tested formally via the Chow test below.

Also note, neither Granger nor contemporaneous causality imply causality per se. The former implies only that Granger causal variable X temporally leads effect variable Y, while the latter only implies a non-zero correlation (Lütkepohl, 2005).

[2] Similar results were obtained for q between 10 and 20

Chow breakpoint test

Lütkepohl (2005) proposes a Chow breakpoint test that enables us to examine whether there was a structural break in the volatility-volume relationship at a breakpoint date $\tilde{T}$. It works by testing whether there is a significant gain, in terms of estimation residuals, to allowing model parameters to differ about the breakpoint $\tilde{T}$.

The null is of parameter stability. In small samples, the distribution of the test statistic under the null may differ markedly from the asymptotic distribution (Candelon and Lütkepohl, 2000), so bootstrapped p-values are used.

We are interested in testing whether the day of the Crash represented a structural break. Hence, $\tilde{T}$ is set as the first observation on 6 May.

For robustness, the Chow test is also estimated for $\tilde{T}$= 4 May and $\tilde{T}$ = 5 May to see if a potential structural break on 6 May is unique. These Chow tests were estimated for the full sample and for 3 – 5 May only.

Results

Estimation results are displayed in Table 6.2.

Table 6.2

q = 20	Granger		Contemporaneous	
	H_0: $ARV_t \nrightarrow V_t$	H_0: $V_t \nrightarrow ARV_t$	H_0: $E(\varepsilon_{1,t}\varepsilon_{2,t}) = 0$	$Corr(\hat{\varepsilon}_{1,t}, \hat{\varepsilon}_{2,t})$
Full Sample	0.626	2.600	5.304	0.101
(p-val)	(0.896)	(0.000)	(0.021)	
3 – 5 May	1.339	1.409	5.576	0.121
(p-val)	(0.146)	(0.112)	(0.018)	
6 May	0.456	1.621	1.634	0.120
(p-val)	(0.978)	(0.049)	(0.201)	

Chow Breakpoint Test	$\tilde{T}$= 6th May	$\tilde{T}$= 5th May	$\tilde{T}$= 4th May
Full sample			
H_0: Parameters constant	1881.25	1181.73	619.15
(Bootstrapped p-val)	(0.000)	(0.000)	(0.025)
3 – 5 May			
H_0: Parameters constant	-	154.37	144.04
(Bootstrapped p-val)	-	(0.143)	(0.228)

$ARV_t \nrightarrow V_t$ means ARV does not Granger cause V_t etc. Test statistics are reported.

For the full sample, the null hypotheses of no Granger causality (from volume to volatility) and no contemporaneous causality are both rejected at the 5% level. This result is difficult to reconcile with the MDH and SIAH, as these theories are, strictly speaking, mutually exclusive. However, splitting the sample shows that the significance

of Granger causality was being driven by observations from 6 May, while contemporaneous causality was driven by 3 – 5 May.

Furthermore, evidence from the Chow breakpoint tests support the hypothesis that there was a unique break on 6 May in our sample. Although the null of stability is rejected for all $\tilde{T}$ in the full sample, it is possible that this is driven by the presence of observations from 6 May. Removing 6 May from the sample confirms this intuition, as $\tilde{T}$= 4 and 5 May are no longer significant as breakpoints.

Finally, there is no evidence of Granger causality from volatility to volume.

To conclude, while the MDH holds in "normal times", there was a break to the SIAH on 6 May.

Discussion

Our findings are best interpreted in the context of studies that have investigated potential causes for the Flash Crash.

According to Aldrich et al (2017), the primary mechanism behind the Flash Crash was a lack of integrity in mainstream data feeds, such that stale prices were reported alongside live, marketable prices in the lead up to the event. Additionally, Menkveld and Yueshen (2015) show that there were unexploited arbitrage opportunities between the E-mini and the SPY ETF (which also tracks the S&P 500) around the Crash. Meanwhile, Easley et al. (2011) showed, using their "VPIN" metric, that there was an increasing level of informed trading in the hours leading up to the Crash. In each study, the authors claim that the mentioned market anomaly created uncertainty and caused agents to withdraw from the market.

Taken together, these studies help to explain our results. Specifically, a by-product of the market anomalies and subsequent withdrawal of agents was that the information assimilation process became less efficient and new information was not incorporated into prices instantly. Thus, information reached agents in a sequential, rather than contemporaneous fashion on May 6.

7. Conditional volatility models and forecasting

In this section, I assess the in-sample fit and out-of-sample forecasting performance of GARCH and Dynamic Conditional Score models.

I use de-seasonalised 3-minute returns as inputs to these models. However, for obtaining forecasts and fitted values, the volatility estimates are re-seasonalised.

Table 7.1 displays summary statistics for the de-seasonalised returns. It confirms the presence of volatility clustering (by testing for serial correlation in squared returns) and fat tails (excess kurtosis) in the returns. It also confirms the Martingale Difference Hypothesis that the best ex-ante predictor of returns is its unconditional mean. The returns also display a slight positive skew.

Table 7.1: Summary statistics

	De-seasonalised Returns
Mean	-0.0865
Standard deviation	2.607
Excess Kurtosis	89.8
Skewness	0.016
Martingale Diff. (p-val)	0.527 (0.468)
$Q_{20}(r^2)$ (p-val)	110.8 (0.000)

Martingale Diff. is Escanciano and Lobato (2009) test for Martingale Difference Hypothesis. $Q_{20}(r^2)$ is Ljung and Box (1979) test for serial correlation

Econometric Framework

To specify GARCH models, first define the one-step ahead conditional mean and variance(volatility) of returns:

$$\mu_{t|t-1} = E(r_t|\Omega_{t-1}) \text{ and } \sigma^2_{t|t-1} = Var(r_t|\Omega_{t-1})$$

where Ω_{t-1} is the information set containing all information up to time $t-1$.

Then, decompose r_t in terms of its conditional mean and variance

$$r_t = \mu_{t|t-1} + \sigma_{t|t-1} z_t\,, \quad z_t \sim i.i.d.(0,1) \tag{7.1}$$

For now, I set $\mu_{t|t-1} = \mu$, a constant.

The GARCH(1,1) (Bollerslev, 1986) model for conditional volatility is then[3]:

$$\sigma^2_{t|t-1} = \omega + \phi_1 \sigma^2_{t-1|t-2} + \kappa_1 \varepsilon^2_{t-1} \qquad with\ \omega > 0, \phi_1 \geq 0, \kappa_1 \geq 0 \tag{7.2}$$

[3] Notation and structure in this section is similar to Harvey and Sucarrat (2014)

where $\varepsilon_t = r_t - \mu$ is the return innovation. The positivity constraints ensure that the conditional volatility is always positive.

The GARCH(1,1) model may be extended to include a "leverage" effect, which allows $\sigma^2_{t|t-1}$ to respond asymmetrically to positive and negative return innovations. An example is the GJR-GARCH(1,1) of Glosten et al. (1993):

$$\sigma^2_{t|t-1} = \omega + \phi_1 \sigma^2_{t-1|t-2} + \kappa_1 \varepsilon^2_{t-1} + \kappa^* \varepsilon^2_{t-1} I(\varepsilon_{t-1} < 0) \qquad (7.3)$$

where $I(\,.\,)$ is the indicator function. For $\kappa^* > 0$, negative return innovations have a larger effect on future conditional volatility than positive innovations.

Motivated by the results in Section 6, I also include a lag of log de-seasonalised volume in the dynamic equation:

$$\sigma^2_{t|t-1} = \omega + \phi_1 \sigma^2_{t-1|t-2} + \kappa_1 \varepsilon^2_{t-1} + \kappa^* \varepsilon^2_{t-1} I(\varepsilon_{t-1} < 0) + \theta \log(V_{t-1}) \qquad (7.4)$$

Now, to accommodate the fat tails in returns (Table 7.1), Andersen and Bollerslev (2006) recommend assuming that z_t is t_υ distributed with degrees-of-freedom $\upsilon > 2$.

I also consider the skew-t distribution for z_t (as in Giot and Laurent (2003)). Harvey and Sucarrat (2014) show, in context of the Beta-t-EGARCH (presented below), that skewing the t-distribution introduces a slight leverage effect. This reinforces the GJR-type leverage effect in (7.3).

The Fernández and Steel (1998) method of skewing, with skewness parameter ξ, is used. $\xi = 1$ implies symmetry; $\xi < 1$ implies negative skewness; $\xi > 1$ implies positive skewness.

I estimate four specifications from the GARCH class of models. These are the GARCH-t(1,1) (labelled "G1" below), GJR-GARCH-t(1,1) ("G2"), GJR-GARCH-t(1,1)+volume ("G3"), GJR-GARCH-skewed-t(1,1)+volume ("G4").

Now, in context of the Flash Crash, there are some drawbacks to traditional GARCH models. Harvey (2013) mentions that a consequence of specifying conditional volatility as a linear function of squared innovations, ε^2_t, is that it responds too much to extreme values and the effect then dies out only slowly. As an alternative that is more robust to extreme observations, Harvey suggests writing the dynamic equation as a function of the conditional score of the t-distribution.

Here, we rely on the Beta-t-EGARCH class of models (Harvey and Chakravarty, 2008). The primary features of the Beta-t-EGARCH are that the dynamic equation is written in terms of the log of the conditional scale[4] and that an exponential link function is used.

To define the Beta-t-EGARCH, first write

$$r_t = \mu_r + (z_t - \mu_z)\exp\left(\lambda_{t|t-1}\right) \tag{7.5}$$

where μ_r is a constant parameter, which is both the conditional and unconditional mean. Here, we have set $\mu_{t|t-1} = \mu_r - \mu_z \exp\left(\lambda_{t|t-1}\right)$ to be time varying to ensure that r_t is a martingale difference (Harvey and Sucarrat, 2014). z_t is t_v distributed (possibly skewed) as above, while $\lambda_{t|t-1}$, the log of the scale, is a linear function of lags of the conditional score:

$$u_t = \frac{(\nu+1)\left(\mathrm{r_t} - \mu_r + \mu_z \exp\left(\lambda_{t|t-1}\right)\right)(r_t - \mu_r)}{\nu\xi^{2sgn\left(\mathrm{r_t} - \mu_r + \mu_z \exp\left(\lambda_{t|t-1}\right)\right)} \exp\left(2\lambda_{t|t-1}\right) + \left(\mathrm{r_t} - \mu_r + \mu_z \exp\left(\lambda_{t|t-1}\right)\right)^2} - 1$$

The first-order model (denoted Beta-t-EGARCH(1,0) (Harvey, 2013)) is then

$$\lambda_{t|t-1} = \omega + \phi_1 \lambda_{t-1|t-2} + \kappa_1 u_{t-1} \tag{7.6}$$

which is stationary for $|\phi_1| < 1$.

Again, leverage effects dependant on the sign of return innovations can be introduced:

$$\lambda_{t|t-1} = \omega + \phi_1 \lambda_{t-1|t-2} + \kappa_1 u_{t-1} + \kappa^* sgn\left(-(r_{t-1} - \mu_r)\right)(u_{t-1} + 1) \tag{7.7}$$

The negative sign of $r_t - \mu_r$ is taken to keep consistent with the GJR specification. Additionally, the log of de-seasonalised volume can also be included (with coefficient θ) into the dynamic equation.

I estimate four specifications from the Beta-t-EGARCH class of models. These are
Beta-t-EGARCH(1,0) ("B1"), Beta-t-EGARCH(1,0)+leverage ("B2"),
Beta-t-EGARCH(1,0)+leverage+volume ("B3"),
Beta-skewed-t-EGARCH+leverage+volume ("B4").

Harvey (2013) also mentions that the (one component) Beta-t-EGARCH may react too little to short-term increases in volatility. This motivates the construction of models with separate long and short-term components, with the latter capturing transitory spikes in volatility after large shocks.

[4] For t-distribution, conditional scale $\varphi_{t+1|t} = (\nu-2)^{1/2}\sigma_{t+1|t}$

Engle and Lee (1999) propose a two-component GARCH model, which is extended to the Beta-t-EGARCH in Harvey and Sucarrat (2014). Both papers include a leverage effect, but only in the short-term component. This convention is also followed here. I estimate two two-component Beta-t-EGARCH models, one with just leverage ("TB1") and one with leverage and skewness ("TB2").

The two-component Beta-t-EGARCH model is

$$\lambda_{t|t-1} = \omega + \lambda^{\dagger}_{1,t|t-1} + \lambda^{\dagger}_{2,t|t-1} \tag{7.8}$$

where

$$\lambda^{\dagger}_{1,t|t-1} = \phi_1 \lambda^{\dagger}_{1,t-1|t-2} + \kappa_1 u_{t-1}, \qquad |\phi_1| < 1, \;\; \phi_1 \neq \phi_2$$

$$\lambda^{\dagger}_{2,t|t-1} = \phi_2 \lambda^{\dagger}_{2,t-1|t-2} + \kappa_2 u_{t-1} + \kappa^{*} sgn(-(r_{t-1} - \mu_r))(u_{t-1} + 1)$$

$\lambda^{\dagger}_{1,t|t-1}$ and $\lambda^{\dagger}_{2,t|t-1}$ are the long- and short-term components respectively.

Forecasting and evaluation

Here, one-step ahead, out-of-sample volatility forecasts are created for the interval from the onset of the Crash at 13:30:00 to the end of the trading day at 15:15:00. This results in 35 forecasts per model.

To create the forecasts, a real-time coefficient updating algorithm was used as follows. First, the model is estimated using all observations up to 13:30:00 on 6 May. Then, a one-step ahead forecast is created for the 3-minute interval ending at 13:33:00. Next, the model is estimated using all observations to 13:33:00 and a forecast created for the interval ending 13:36:00 etc. I follow this procedure because it reflects what happens in practice.

Recall, the total volatility forecasts are recovered by multiplying the forecasts obtained here by the square of the seasonal component, s^2_m.

These forecasts, denoted f_t, are evaluated against ARV_t. As in Engle and Sokalska (2012), the following loss functions are used: the mean squared error (MSE):

$$MSE = \frac{1}{N}\sum (ARV_t - f_t)^2 \tag{7.9}$$

and the mean out of sample likelihood (LIKE):

$$LIKE = \frac{1}{N}\sum \left(\log(f_t) + \frac{ARV_t}{f_t}\right) \tag{7.10}$$

Using realised volatility in place of the true volatility causes bias in many loss functions (Andersen and Bollerslev, 2005). However, following Patton (2006), MSE and LIKE are unbiased for evaluating optimal forecasts.

The G@RCH Oxmetrics module (Laurent, 2013) and betategarch package in R (Sucarrat, 2016) are used for estimation. The two-component Beta-t-EGARCH model was not augmented with volume because these packages did not allow for such a facility.

Results

In-sample fit

Table 7.2 presents estimation results for the full sample. There is significant evidence (at the 10% level) of fat tails, negative conditional skewness and leverage in all specifications.

Lagged volume is just significant with a positive coefficient (as predicted by market microstructure theory) at the 10% level in the GARCH specifications, but not in the Beta-t-EGARCH, so the overall importance of volume in-sample is weak. However, we saw in Section 6 that lagged volume only had predictive ability for volatility on 6 May, so it may be more helpful in the out-of-sample forecasting below.

Considering overall model fit, both the SIC and log-likelihood favour the one component Beta-t-EGARCH specifications over their respective GARCH counterparts, i.e. B1 is preferred to G1 etc.

Furthermore, the two-component Beta-t-EGARCH with leverage (TB1) is preferred to its one-component counterpart (B2) according to both log likelihood and SIC. Meanwhile, the two-component model with leverage and skew (TB2) is preferred to the one-component model with leverage, skew and volume (B4) according to SIC, but not log likelihood. Hence overall, or at least based on the equivalent comparison between TB1 and B2, the two-component Beta-t-EGARCH is preferred to the one-component model in-sample.

Also note that within each class, SIC prefers the most parsimonious model.

For illustration, Figure 7.1 plots the (re-seasonalised) fitted values for volatility from the fully augmented specification in each class. These are superimposed over the ARV series. Panel (a) plots the series for 3 – 5 May and panel (b) for 6 May.

Panel (a) shows that the fitted values from the models are indistinguishable in "normal times". Differences only arise following extreme observations such as the Flash Crash (Panel b). As expected, the traditional GARCH (G4) reacts the most to the Crash. Meanwhile, the two-component model (TB2) reacts the least. However, the GARCH fails to recover sufficiently quickly after the Crash, meaning that it remains well above the ARV series and the other fitted values until the end of the day.

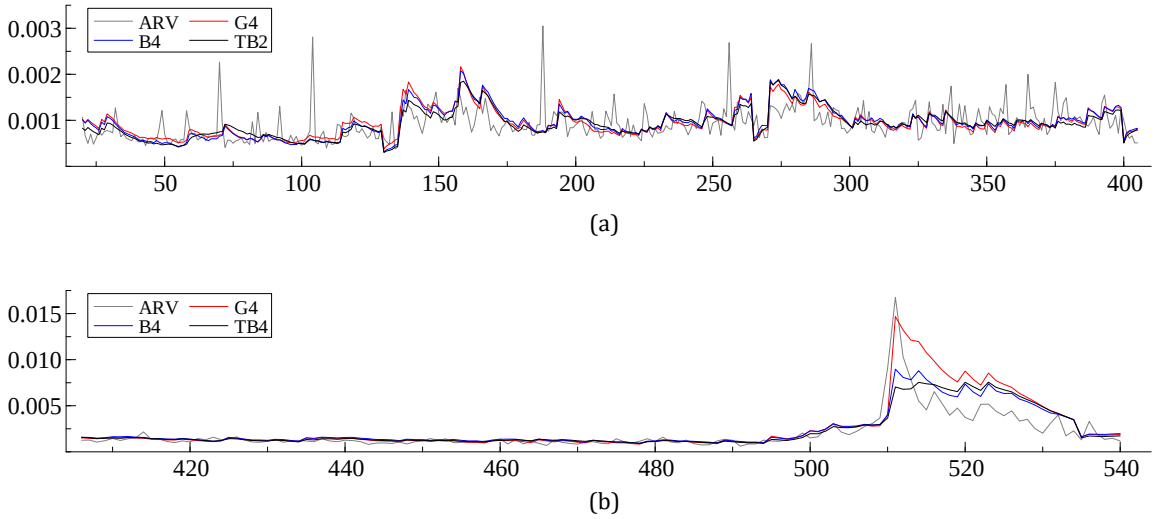

Figure 7.1: Y-axis measured in standard deviations, X-axis in 3-minute intervals

Out-of-sample forecasts

Table 7.3 displays the forecast evaluation results.

Table 7.3: Forecast evaluation

	MSE	LIKE
G1	4.90×10^{-09}	-9.5051
G2	4.42×10^{-09}	-9.5764
G3	3.59×10^{-09}	-9.6207
G4	3.27×10^{-09}	-9.6514
B1	3.52×10^{-09}	-9.5129
B2	2.75×10^{-09}	-9.6551
B3	2.70×10^{-09}	-9.6456
B4	2.45×10^{-09}	-9.7031
TB1	2.34×10^{-09}	-9.7370
TB2	2.15×10^{-09}	-9.7817

The one-component Beta-t-EGARCH specifications out-perform their respective GARCH counterparts according to both MSE and LIKE. Furthermore, both two-component Beta-t-EGARCH specifications outperform all one component models according to both MSE

Table 7.2: In sample fit

	$\widehat{\omega}$ (p-val)	$\hat{\phi}_1$ (p-val)	$\hat{\phi}_2$ (p-val)	$\hat{\kappa}_1$ (p-val)	$\hat{\kappa}_2$ (p-val)	$\hat{\kappa}^*$ (p-val)	$\hat{\theta}$ (p-val)	$\hat{\upsilon}$ (p-val)	$\widehat{\log(\xi)}$ (p-val)	Log Like. (SIC)	$Q_{20}(\hat{z}_t^2)$ (p-val)
G1	0.063	0.806		0.180				7.653		-900.13	13.63
	(0.069)	(0.000)		(0.000)				(0.011)		(3.392)	(0.848)
G2	0.084	0.800		0.090		0.146		8.481		-898.49	14.16
	(0.046)	(0.000)		(0.065)		(0.074)		(0.017)		(3.398)	(0.822)
G3	-0.306	0.797		0.070		0.154	0.043	8.542		-896.26	15.35
	(0.179)	(0.000)		(0.171)		(0.054)	(0.091)	(0.018)		(3.407)	(0.756)
G4	-0.304	0.786		0.064		0.165	0.045	9.318	-0.135	-894.36	15.53
	(0.198)	(0.000)		(0.179)		(0.045)	(0.091)	(0.020)	(0.063)	(3.412)	(0.745)
B1	0.015	0.980		0.168				7.478		-899.07	18.12
	(0.188)	(0.000)		(0.000)				(0.015)		(3.388)	(0.580)
B2	0.014	0.976		0.151		0.042		8.057		-897.50	17.90
	(0.222)	(0.000)		(0.000)		(0.057)		(0.031)		(3.394)	(0.594)
B3	-0.219	0.966		0.140		0.041	0.024	8.143		-895.77	18.34
	(0.232)	(0.000)		(0.000)		(0.059)	(0.203)	(0.034)		(3.406)	(0.565)
B4	-0.227	0.960		0.138		0.042	0.025	9.420	-0.130	-894.09	18.02
	(0.228)	(0.000)		(0.000)		(0.050)	(0.198)	(0.071)	(0.091)	(3.411)	(0.586)
TB1	0.076	0.991	0.933	0.141	0.127	0.038		8.884		-896.65	15.74
	(0.138)	(0.000)	(0.000)	(0.039)	(0.041)	(0.059)		(0.022)		(3.393)	(0.732)
TB2	0.059	0.991	0.941	0.139	0.130	0.039		9.878	-0.126	-895.87	15.90
	(0.168)	(0.000)	(0.000)	(0.048)	(0.045)	(0.054)		(0.043)	(0.075)	(3.398)	(0.723)

G: GARCH specifications. B: One component Beta-t-EGARCH specifications. TB: Two component Beta-t-EGARCH specifications. Log Like.: Log Likelihood. SIC: Schwarz Information Criterion. $Q_{20}(\hat{z}_t^2)$: Ljung and Box (1979) test for 20th order serial correlation in squared standardised residuals.

and LIKE. Overall, the two-component Beta-t-skewed-EGARCH+leverage model performs best.

Adding lagged volume to the traditional GARCH (G2 → G3) significantly improves forecasting performance according to both MSE and LIKE. However, while adding volume reduces the MSE for the Beta-t-EGARCH, it increases LIKE. Hence overall, I conclude that the results here are broadly consistent with those in Section 6, where we found that lagged volume had predictive power for volatility on the day of the Crash.

Also, adding leverage and skewness always improves the forecasting performance for all models.

Figure 7.2 compares the forecasting performance of all specifications. Panel (a) plots all four forecast series from the traditional GARCH class; Panel (b) the four series from the Beta-t-EGARCH class; Panel (c) the two series from the two components class; and panel (d) the best performing series from each class.

The following general pattern is observed: the addition of leverage allows the forecasts to better capture the peak of volatility during the Crash, while the addition of log volume and skewness both help to bring the forecasts down faster after the peak.

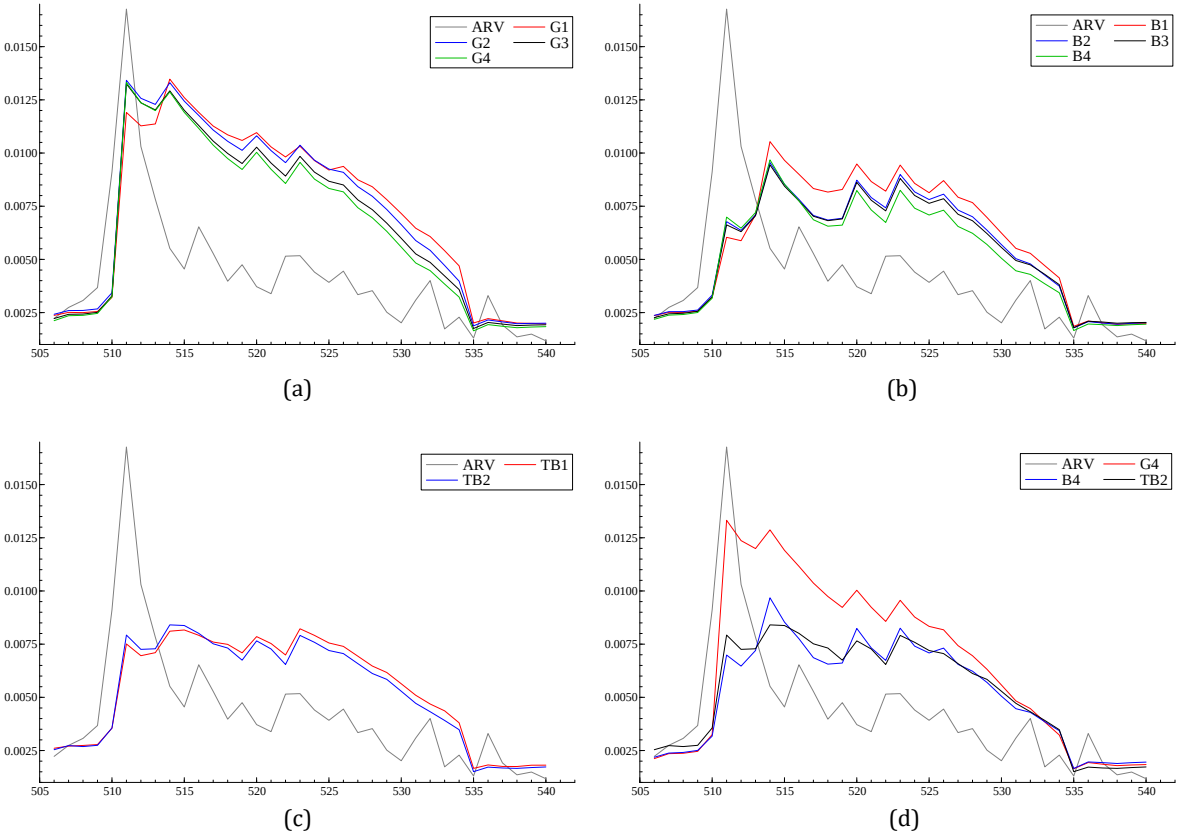

Figure 7.2: Y-axis measured in standard deviations, X-axis in 3-minute intervals

Panel (d), which compares the best performers from each class, shows that the GARCH (G4) best captures the peak, but fails to recover sufficiently quickly. The two-component model (TB2) outperforms the one component Beta-t-EGARCH (B4) because it reaches a higher peak and comes down slightly more quickly.

8. Conclusion

In this paper, I explore various aspects of intraday volatility in the E-Mini around the 6 May 2010 Flash Crash. Computing a robust measure of realised volatility, I found that volatility was unusually high in the lead up to the Crash and that high levels of volatility persisted after the Crash. I also found a structural break in the volatility-volume relationship on 6 May, such that the information assimilation process became less efficient. Finally, I created one-step ahead forecasts using a real-time coefficient updating procedure. I found that Dynamic Conditional Score models provide a valuable extension to traditional GARCH models for forecasting volatility over extreme events.

Bibliography

Aït-Sahalia, Y., Mykland, P. A., & Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. Journal of Econometrics, Elsevier, vol. 160(1), 160-175.

Aldrich, Eric M. and Grundfest, J., and Laughlin, G. (2017). The Flash Crash: A New Deconstruction. Available at SSRN: https://ssrn.com/abstract=2721922 or http://dx.doi.org/10.2139/ssrn.2721922.

Andersen, T. (1996). Return volatility and trading volume: an information flow interpretation of stochastic volatility. Journal of Finance 51, 169–204.

Andersen, T., & Bollerslev., T. (1997). Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance, 4, (2-3), 115-158.

Andersen, T., & Bollerslev, T. (1998). Deutsche Mark–Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies. Journal of Finance, 53(1), 219-265.

Andersen, Bollerslev, Christoffersen, & Diebold. (2006). Chapter 15 Volatility and Correlation Forecasting. Handbook of Economic Forecasting, 1, 777-878.

Andersen, T., Bollerslev, T., Diebold, F., & Labys, P. (2003). Modeling and Forecasting Realized Volatility. Econometrica, 71(2), 579-625.

Andersen, T., 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., Dobrev, D., & Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation, Journal of Econometrics, Vol. 169(1), p.75-93.

Awartani, B., Corradi, V., & Distaso, W. (2009). Assessing Market Microstructure Effects via Realized Volatility Measures with an Application to the Dow Jones Industrial Average Stocks. Journal Of Business & Economic Statistics, 27(2), 251-265.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., & Shephard, N. (2007). Subsampling Realised Kernels. SSRN Electronic Journal, SSRN Electronic Journal, 2007.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., & Shephard, N. (2008). Designing Realised Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise. SSRN Electronic Journal, SSRN Electronic Journal, 2008.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: Trades and quotes. Econometrics Journal, 12(3), C1-C32.

Barndorff-Nielsen, O., Shephard, N., Franses, Philip Hans, & Mcaleer, Michael. (2002). Estimating quadratic variation using realized variance. Journal of Applied Econometrics, 17(5), 457-477.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

Candelon, B. and Lütkepohl, H. (2000). On the reliability of Chow type tests for parameter constancy in multivariate dynamic models, *Discussion paper*, Humboldt-Universität Berlin.

CFTC and SEC (2010), "Preliminary Findings Regarding the Market Events of May 6, 2010," *Staff Report*.

Clark, P. (1973). A subordinated stochastic process model with finite variance for speculative prices. Econometrica 41, 135–155.

Copeland, T., (1976). A model of asset trading under the assumption of sequential information arrival. Journal of Finance 31, 1149–1168.

Creal, D., Koopman, S., & Lucas, A. (2009). A General Framework for Observation Driven Time-Varying Parameter Models. IDEAS Working Paper Series from RePEc, IDEAS Working Paper Series from RePEc, 2009.

Creal, D., Koopman, S., & Lucas, A. (2011). A Dynamic Multivariate Heavy-Tailed Model for Time-Varying Volatilities and Correlations. Journal of Business & Economic Statistics, 29(4), 552-563.

Darrat, A., Rahman S., & Zhong, M. (2003). Intraday trading volume and return volatility of the DJIA stocks: A note. Journal of Banking & Finance, 27, (10), 2035-2043.

Easley, David, De Prado, Marcos M. Lopez, & O'Hara, Maureen. (2011). The microstructure of the "flash crash": Flow toxicity, liquidity crashes, and the probability of informed trading. Journal of Portfolio Management, 37(2), 118-128,12,14.

Easley, D., O'Hara, M. (1987). Price, trade size, and information in securities markets. Journal of Financial Economics 19, 69-90.

Easley, D., López de Prado, M., & O'Hara, M. (2012). Flow Toxicity and Liquidity in a High-frequency World. The Review of Financial Studies, 25(5), 1457-1493.

Engle, R.F., Lee, G.G.J., (1999). A long-run and a short-run component model of stock return volatility. In: Engle, R.F., White, H. (Eds.), Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive Granger. Oxford University Press, Oxford.

Engle, R., & Sokalska, M. (2012). Forecasting intraday volatility in the US equity market. Multiplicative component GARCH. Journal of Financial Econometrics, 10(1), 54-83.

Escanciano, J.C., Lobato, I.N., 2009. An automatic portmanteau test for serial correlation. Journal of Econometrics 151, 140–149.

Farrell, M. (2013). Mini flash crashes: A dozen a day. [ONLINE] Available at: http://money.cnn.com/2013/03/20/investing/mini-flash-crash/index.html. [Accessed 24 April 2018].

Fernández, C., Steel, M., (1998). On Bayesian modelling of fat tails and skewness. Journal of the American Statistical Association 93, 359–371.

Foresight, (2012). The Future of Computer Trading in Financial Markets, Final Project Report: Executive Summary. The Government Office for Science, London.

Giot, P. (2005). Market risk models for intraday data. The European Journal of Finance, 11(4), 309-324.

Giot, P., Laurent, S., (2003). Value at risk for long and short trading positions. Journal of Applied Econometrics 18, 641–664.

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

Harvey, A. (2013). Dynamic models for volatility and heavy tails : With applications to financial and economic time series / Andrew C. Harvey, University of Cambridge. (Econometric Society monographs ; 52).

Harvey, A.C., Chakravarty, T., (2008). Cambridge Working paper in Economics, CWPE 0840.

Harvey, A., Sucarrat, G. (2014). EGARCH models with fat tails, skewness and leverage. Computational Statistics and Data Analysis, p. 320-338.

J.P. Morgan (1996). RiskMetrics, Technical Documents, fourth ed. New York.

Kirilenko, A., Kyle, A., Samadi, M., & Tuzun, T. (2017). The Flash Crash: High-Frequency Trading in an Electronic Market. Journal of Finance, 72(3), 967-998.

Laurent, S., 2013. G@RCH7. Timberlake Consultants Ltd., London.

Ljung, G., Box, G., (1979). On a measure of lack of fit in time series models. Biometrika 66, 265–270.

Lütkepohl, H. (2005). New introduction to multiple time series analysis (2nd ed.). Berlin: Springer.

Lütkepohl, H., & Krätzig, M. (2004). Applied time series econometrics (Themes in Modern Econometrics), Cambridge University Press, Cambridge.

Madhavan, A. (2000). Market microstructure: A survey. Journal of Financial Markets, 3(3), 205-258.

Martens, M., Chang, Y., & Taylor, S. (2002). A Comparison of Seasonal Adjustment Methods When Forecasting Intraday Volatility. Journal of Financial Research, 25(2), 283-299.

Mcaleer, M., & Medeiros, M. (2008). Realized Volatility: A Review. Econometric Reviews, 27(1-3), 10-45.

Menkveld, A. J. and Yueshen, B. Z. (2015). The Flash Crash: A Cautionary Tale about Highly Fragmented Markets. Working Paper, VU University Amsterdam.

Mougoué, M., & Aggarwal, R. (2011). Trading volume and exchange rate volatility: Evidence for the sequential arrival of information hypothesis, Journal of Banking & Finance, Elsevier, vol. 35(10), pages 2690-2703.

Nishimura, Y., & Sun, B. (2015). Intraday Volatility and Volume in China's Stock Index and Index Futures Markets. Asia-Pacific Journal of Financial Studies, 44(6), 932-955.

Patton, A. 2006. “Volatility Forecast Evaluation and Comparison Using Imperfect Volatility Proxies.” Working Paper. London School of Economics; forthcoming in Journal of Econometrics.

Pigorsch C., Pigorsch U., Popov I. (2012) Volatility Estimation Based on High-Frequency Data. In: Duan JC., Härdle W., Gentle J. (eds) Handbook of Computational Finance. Springer Handbooks of Computational Statistics. Springer, Berlin, Heidelberg.

Sucarrat, G., 2016. betategarch: estimation and simulation of first-order Beta-t-EGARCH models. R Package Version 3.3. http://cran.r-project.org/web/packages/betategarch/.

Taylor, S., & Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. Journal of Empirical Finance. 4:317-340.

Xu, X., Chen, P., & Wu, C. (2006). Time and dynamic volume–volatility relation. Journal of Banking & Finance. 30. 1535–1558.

Zhang, L., Mykland, P., & Aït-Sahalia, Y. (2005). A Tale of Two Time Scales. Journal of the American Statistical Association, 100(472), 1394-1411.