CARDIFF BUSINESS SCHOOL
WORKING PAPER SERIES

Forecasting volatility using drift burst information

Kefu Liao [1], Kevin P. Evans, and Dudley Gilder

December 2022

Cardiff Business School
Cardiff University
Colum Drive Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
www.cardiff.ac.uk/carbs

[1] Contact author. Email: LiaoK2@cardiff.ac.uk

Forecasting volatility using drift burst information

Abstract

Using high-frequency data for the S&P 500 Index, this paper investigates whether the information of drift burst intensity predicts future volatility. We find that drift burst intensity increases future volatility. Moreover, we find the intensity of drift burst signs is also important: the intensity of negative drift bursts strongly increases the future volatility while the intensity of positive drift bursts weakly decreases the future volatility. And the models exploiting these findings lead to significantly better out-of-sample forecast performance.

JEL: G12, E44, E32

1 Introduction

Recent years see a growing strand of literature that emphasized the economic value of drift bursts. Research tends to use this model to identify flash crashes, which are then used for case studies for financial stability. Relying on the flash crashes informed by the drift burst model, Bellia et al. (2020) find that High-Frequency Traders (HFTs) consume liquidity during a flash crash, which exacerbates the downtrend. Using the crash events identified by the drift bust model, other studies find that well-capitalized standby liquidity providers such as short-term traders (Getmansky et al., 2017, Jagannathan et al., 2019) and mutual funds (Jagannathan et al., 2021) help the market recover from crashes. Flora and Renò (2020) apply the drift burst model to detect mini-flash crashes and estimate the transient crash loss for some security markets. Christensen et al. (2022) apply drift bursts to forecast return dynamics during flash crash periods.

Although researchers find drift bursts are economically important, whether drift bursts also benefit volatility forecasting is new in the literature. Recently financial econometricians argue that reducing the drift burst bias has volatility forecasting advantages. For example, Andersen et al. (2021) find alleviating drift burst bias for the historical volatility predictors improves their predictive ability. Only very few papers focus on the possible usefulness of drift burst itself in forecasting volatility (Laurent et al., 2022).

We contribute to the literature by investigating whether drift burst intensity predicts volatility. We identify the drift bursts based on the method by Christensen et al. (2022) and estimate the intensity by the approaches by Hawkes (1971) and Tauchen and Zhou (2011). The in-sample results from S&P 500 index reveal that drift burst intensity strongly increases future volatility. Moreover, we also show the importance of decomposing the signs of the drift bursts: the negative drift burst intensity strongly increases volatility and the positive drift burst intensity weakly decreases volatility, and this effect is more prominent and robust than that of

(unsigned) drift bursts. In addition, our findings from the in-sample are robust to the intensity measures by both Hawkes (1971) and Tauchen and Zhou (2011).

More importantly, we find the drift bursts also greatly contribute to the out-of-sample forecast. The model associated with drift burst intensities is able to provide significantly better out-of-sample volatility forecasts, for both short-term and long-term horizons. Further, we show that it is also essential to separate the signs of drift bursts for out-of-sample: the model associated with the intensities of signed drift burst provide more accurate forecasts than that with the intensities of (unsigned) drift bursts. Besides, these out-of-sample results are robust to the intensity measures by Hawkes (1971) and Tauchen and Zhou (2011), and the drift burst intensity measure by Tauchen and Zhou (2011) leads to better out-of-sample than that by Hawkes (1971).

The work by Laurent et al. (2022) is closely related to our study. They also find the usefulness of drift bursts for predicting volatility. However, there are several aspects which differ from (or extend) their works. First, we measure the drift bursts from a different perspective. Laurent et al. (2022) measure the drift burst magnitude while we gauge the intensity of drift bursts. Second, Laurent et al. (2022) have not yet investigated the information on the sign of drift bursts. As discussed in Christensen et al. (2022) and Andersen et al. (2021), drift bursts commonly lead to substantial short price trends, with positive (negative) drifts contributing to the upside (downside) price trend. Motivated by this, we expect that drift burst intensity should have a leverage effect on future volatility: positive drift burst intensity increases volatility while negative drift burst intensity decreases volatility.

Third, Laurent et al. (2022) identify drift bursts at the daily level while we detect the drift bursts at the intraday level by the method of Christensen et al. (2022). Drift bursts at the intraday level may provide richer and possibly more useful information. Fourth, for forecasting volatility, Laurent et al. (2022) use general drift bursts (not pretested) while we use statistically

significant (or pre-tested) drift bursts. The use of significant drift bursts benefits less bias driven by the measurement error. Fifth, Laurent et al. (2022) focus on short horizon forecast (tomorrow) while we include both short- and long-term volatility forecasts. The long-term volatility forecast has important implications for option pricing models and longer-term value-at-risk models (Ederington and Guan, 2010).

The remainder of the paper is organized as follows. Section 2 introduces the detection methods for drift bursts and the approaches for estimating drift burst intensity. Section 3 shows the data and descriptive analysis. Section 4 reports the in-sample and out-of-sample analysis of the drift burst intensity. Section 5 is the robustness analysis. Section 6 concludes.

2 Volatility estimation, drift burst detection and intensity measurement

2.1 Measuring volatility

Assuming log prices p_{t_i} are observed at $t_0, t_1, \dots, t_n$ at day t. Andersen and Bollerslev (1998) show that the full variation QV_t for day t is estimated by the Realized Variance (RV_t), which sums up all n squared returns computed from these prices,

$$RV_t = \sum_{i=1}^{n} r_{t_i}^2 \xrightarrow{p} QV_t,$$

Based on a continuous-time stochastic process for log prices, the full variation QV can be decomposed into the drift, integrated volatility and jump component,

$$QV_t = \int_0^t \mu_{t_s}^2 + \int_0^t \sigma_{t_s}^2 + \sum_{1<s\le t} \kappa_s^2 I$$

As integrated volatility is of primary interest in finance, Mancini (2009) estimates the volatility by assuming the drift component is small and using a threshold-based estimator to remove the jump component. However, recent literature assumes drift could be not small but large and explosive during some short intraday periods (or drift bursts). Therefore, the threshold-based estimator may be biased in estimating integrated volatility. To reduce this bias in the threshold-based estimator, Andersen et al. (2021) proposed a so-called Differenced-return Volatility (DV) estimator [2],

$$DV_{m,t} = \frac{1}{2}\sum_{i=2}^{n} \left(r_{t_i} - r_{t_{i-m}}\right)^2 I\left(\left|r_{t_i} - r_{t_{i-m}}\right| \le 3\sqrt{2}\hat{\sigma}_t^{med}/\sqrt{n}\right),$$

with $\left(\hat{\sigma}_t^{med}\right)^2 = \frac{\pi}{\pi+6-4\sqrt{3}}\left(\frac{n}{n-2}\right)\sum_{i=3}^{n} \mathrm{med}\left(\left|r_{t_{i-2}}\right|, \left|r_{t_{i-1}}\right|, \left|r_{t,i}\right|\right)^2$.

[2] For generalization, we only report the results for DV as opposed to DV_{1-3} in Andersen et al. (2021). But we confirm that the results for DV_{1-3} are qualitatively consistent.

2.2 Detecting intraday drift bursts

The drift burst component is trackable. In the literature, there are three detection methods developed by Christensen et al. (2022), Laurent et al. (2022) and Andersen et al. (2021). The methods by Laurent et al. (2022) and Andersen et al. (2021) detect whether drift bursts occur over a long period (e.g., one day) while the method by Christensen et al. (2022) can locate both drift bursts and their signs at the intraday level. In this paper, we choose the method by Christensen et al. (2022) as it is the only method that identifies the sign of drift bursts, which is our primary interest.

The method by Christensen et al. (2022) detects the existence of intraday drift bursts at t_i if

$$\left|\hat{T}_{t_i}^n\right| = \sqrt{h_n}\frac{|\hat{\mu}_t^n|}{\sqrt{\hat{\sigma}_t^n}} > \Phi,$$

where Φ is a positive critical value,

$$\hat{\mu}_{t_i}^n = \frac{1}{h_n}\sum_{j=1}^{n-k_n+2} K\left(\frac{t_{j-1}-t_i}{h_n}\right) r_{j-1,k_n}^*,$$

and

$$\hat{\sigma}_{t_i}^n = \frac{1}{h_n'}\sum_{j=1}^{n-k_n+2}\left(K\left(\frac{t_{j-1}-t_i}{h_n'}\right) r_{j-1,k_n}^*\right)^2 + 2\sum_{L=1}^{L_n}\omega\left(\frac{L}{L_n}\right)\sum_{j=1}^{n-k_n-L+2} K\left(\frac{t_{j-1}-t_i}{h_n'}\right) K\left(\frac{t_{j+L-1}-t_i}{h_n'}\right) r_{j-1,k_n}^* r_{j-1+L,k_n}^* Y.$$

The remaining parameters are defined as follows. h_n and h_n' are the bandwidths for estimating the spot drift $\hat{\mu}_{t_i}^n$ and spot volatility $\hat{\sigma}_{t_i}^n$, respectively. And L_n is the lag number for the Heteroscedasticity and Autocorrelation Consistent (HAC) correction. r_{t_i,k_n}^* is the

preaveraged return with $r^*_{t_i,k_n} = \frac{1}{k_n}\left(\sum_{j=k_n/2}^{k_n-1} P_{(i+j)/N} - \sum_{j=0}^{k_n/2-1} P_{(i+j)/N}\right)$, where P denotes the price. $\omega(\cdot)$ is a Parzen kernel,

$$\omega(x) = \begin{cases} 1 - 6x^2 + 6|x|^3, \text{for } 0 \le |x| < 1/2, \\ 2(1-|x|)^3, \text{for } 1/2 < |x| \le 1, \\ 0, \text{otherwise.} \end{cases}$$

And $k(\cdot)$ is a backwards-looking exponential kernel, $K(x) = \exp(-|x|)$, for $x \le 0$. The sign of drift burst at t_i is consistent with the sign of $\hat{T}^n_{t_i}$. After identifying all drift bursts for day t, we obtain the total number of drift bursts N_t, positive drift bursts N_t^+, and negative drift bursts N_t^-.

2.3 Hawkes drift burst intensity

We utilize the information on the drift burst occurrence, through a parametric intensity measure by Hawkes (1971). Such an approach is widely applied to deal with event studies in financial markets (Bauwens and Hautsch, 2009, Bowsher, 2007, Large, 2007) and jump studies (Aït-Sahalia et al., 2015, Ma et al., 2019, Clements and Liao, 2017). To begin with, several definitions are required. Let $\{t_q\}_{q\in 1,\dots,N}$ be a random sequence of increasing event times $0 \ge t_1 > \cdots > t_N$ that describe a simple point process (Index t represents the daily frequency).

Given that $\Upsilon_t \coloneqq \sum_{q\ge 1} \mathbf{1}_{t_q \ge t}$ is a counting function, the conditional intensity λ_t can be viewed as the expected change in Υ_t (as a reflection of the probability of an event occurring) over a short time horizon.

$$\lambda_t = \lim_{s\to t} \frac{1}{s-t} E[\Upsilon_s - \Upsilon_t].$$

A common specification for λ_t is the self-exciting Hawkes process (Hawkes, 1971)

$$\lambda_t = \mu + \int_0^t w(t-u)d\Upsilon_u$$

$$= \mu + \sum_{t_q<t} w(t-t_q)$$

where μ is a constant and $w(\cdot)$ is a non-negative weight function. This process is self-exciting in the sense that $\mathrm{Cov}[\Upsilon(a,b),\Upsilon(b,c)] > 0$, where $N(a,b)$ represents the number of events within time interval $[a,b]$, and $0 > a \geq b < c$. The weight function $w(\cdot)$ is a decreasing function of $t-u$, meaning that the intensity decays following a spike. The common approach to implementing the Hawkes model in the above equation is to replace the integral with a discrete sum over past events such that the intensity is given by,

$$\lambda_t = \mu + \sum_{t_q<t} \alpha e^{-\beta(t-t_q)} \tag{1}$$

where α captures the immediate impact on the intensity after a drift burst occurs, and β controls the rate of decay in the exponential weighting function as $t-t_q$ grows. Estimates of the parameters in the above equation are obtained by maximum likelihood. Ogata (1981) shows that the log-likelihood function can be defined recursively as,

$$lnL\left(\{t_q\}_{q=1,\dots,n}\right) = -\mu t_n - \frac{\alpha}{\beta}\sum_{q=1}^{n} 1 - e^{-\beta(t_n-t_q)} + \sum_{q=1}^{n} \ln[\mu + \alpha R_q], \tag{2}$$

where $R_0 = 0$ and $R_q = e^{-\beta(t_q-t_{q-1})}(1+R_{q-1})$, with R_q corresponding to the time of the qth event.

The evaluation of the likelihood begins by setting $R_0 = 0$, with the recursive series R_q then being constructed from Equation 1, given the timing of each jump event. The values for R_q are then used to evaluate the loglikelihood in Equation 2. We estimate α and β by the maximum likelihood estimation (MLE) method, minimising the negative of the log-likelihood

function. The variance of parameter estimates is the inverse of the diagonal of the hessian matrix of the loglikelihood function (Hessian is the equivalent of the observed Fisher information matrix of the log-likelihood function). To estimate the subsequent forecasting models, estimates of the intensity for all points in day t (including those when no events occur) can be constructed recursively from Equation 1 using the estimated α and β and the occurrence time t_q of all past events, $t > t_q$. The intensities estimated on the drift burst, positive, and negative drift burst series are indicated as λ_t, λ_t^+, and λ_t^-, respectively.

2.4 Drift burst intensity by Tauchen and Zhou (2011)

The parametric nature of the Hawkes intensity measure may lead to model specification bias. Moreover, the Hawkes measure may require a high computational cost (due to iterations for the Maximum likelihood estimation) when applied to a high volume of assets. In contrast, the Tauchen and Zhou (2011) (T&Z) measure only requires simple mean calculation, which also requires lower computational costs. Specifically, T&Z proposed an intensity measure by simply averaging the previous point process. The drift burst intensity λ is defined by,

$$\lambda_t = \frac{\sum_{t-k}^{t} I(N_t > 0)}{k}, \tag{3}$$

where $k = 504$ is a 2-year window (Note N_t is previously defined in section 2.2)[3]. Analogously, the intensity of positive and negative drift bursts is defined by,

$$\lambda_t^+ = \frac{\sum_{t-k}^{t} I(N_t^+ > 0)}{k}, \tag{4}$$

and

$$\lambda_t^- = \frac{\sum_{t-k}^{t} I(N_t^- > 0)}{k}. \tag{5}$$

[3] When there are not enough elements to fill the window, we substitute nonexisting elements with the average of the full sample. We apply this method for the rest of the paper.

3 Data

We retrieve one-second E-mini futures (ES) trade prices from Tick Data Inc. This sample is for 17 years from June 2, 2003, to September 30, 2020. To ensures that the sample consists of regular trading days, we remove all non-business days and days in which the exchange closed earlier. After cleaning, we obtain $T = 4310$ observations. As in Andersen et al. (2021), we focus on the most active intraday session (09:30–16:00 EST) and eliminate days with reduced trading hours, to avoid idiosyncratic overnight and weekend effects.

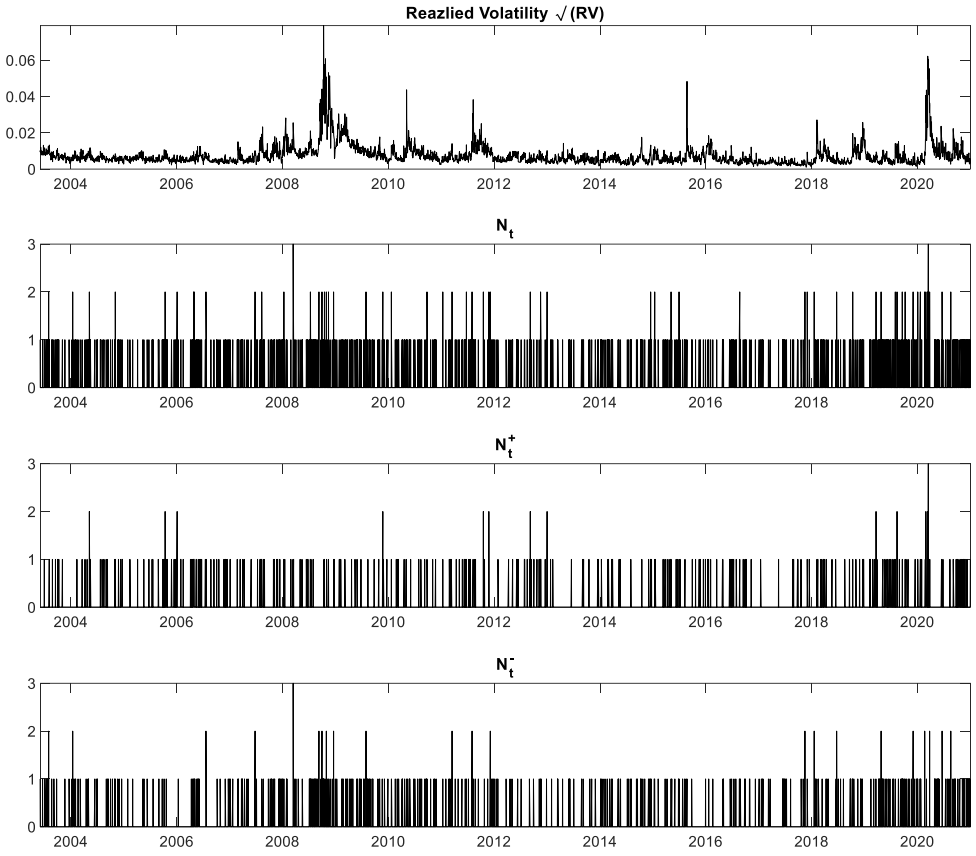
We sample 5-minute prices for previously defined volatility measures. Typically, the five-minute frequency is to alleviate the distortion from market microstructure noise (Ait-Sahalia et al., 2005, Bandi and Russell, 2006, Hansen and Lunde, 2006, Andersen et al., 2007, Bandi and Russell, 2008) [4]. In addition to the realized measures, we obtain an option-based S&P 500 index volatility measure (labelled SV) by Todorov (2019). The SV can be downloaded on www.tailindex.com website, ranging from January 2008 until the end of December 2020. The SV is originally the percentage of annualized volatility, and we transform it to daily level variance to be consistent with the RV and DV measures. As discussed by Andersen et al. (2021), the SV is constructed exclusively from option prices, and thus is void of the specific form for noise structure present in the high-frequency asset prices.

Following Christensen et al. (2022) [5], the implementation of the drift burst test depends on a 5-minute bandwidth ($h_n = 300$ second) for the spot drift and 25-minute ($h_n' = 1500$ second) bandwidth for the spot volatility, with $k_n = 3$ and $L_n = 2(k_n - 1) + 10$ lags for the HAC robust estimate. Using 1-second returns ($n = 23400$), the drift burst test statistic $\hat{T}_{t_i}^n$ is calculated across a regular five-second grid ($i = 1,6,11, \dots, 23400$). Following Christensen et

[4] An alternative strategy to guard against microstructure noise uses the 3-minute prices as in Andersen et al. (2021). We confirm that under the 3-minute frequency, the results hold true.
[5] We thank the authors for sharing the code for the estimation procedure.

al. (2022), we identify an intraday drift burst at t_i if $\left|\widehat{T}_{t_i}^{n}\right|$ exceed $\Phi = 4$, 4.5, or 5, with the sign of this drift burst consistent with the sign of $\widehat{T}_{t_i}^{n}$ [6].

Figure 1. daily realized volatility and drift burst intensities for $|T| > 4$



As in Christensen et al. (2022), we find an unignorable number of intraday drift bursts, with a greater number of negative drift bursts. For example, for $|T| > 4$, we find overall $\sum_1^T N_t = 799$ intraday drift bursts [7], with 349 positive drift bursts and 450 negative drift bursts. The distribution of these intraday drift burst quantities across sample days is depicted in Figure 1. As the figure shows, most days only have one intraday drift burst. There appear to be both more daily and intraday drift bursts during the 2008 financial crisis and 2020 pandemic sessions,

[6] Following Christensen et al. (2022), we allow at most one drift burst to be established over any 5-minute window at which the test statistic attains a local extremum and exceeds the critical value. This cleaning method accounts for the rolling calculation of the test statistic and avoid double counting of events.

[7] This frequency corresponds to about 4 drift bursts one month on average for this ES sample. Table 3 of Christensen et al. (2020) show higher occurrence rate of drift burst for the same E-mini asset (5.7 drift bursts one month). This is because that they detect intraday drift bursts for a longer trading session (02:00–16:15 Eastern Time) and they focus on a more recent subsample of our sample (January 2010–December 2017).

with the realized volatility, also being highest for these periods. It seems that the occurrence rate of drift bursts is positively related to the volatility. To conserve space, the results are not reported for more conservative critical value $|T| > 4.5$ or $|T| > 5$. But we confirm that the patterns of drift bursts in Figure 1 are consistent if those critical values are applied.

To estimate the occurrence of drift bursts, we use the above intensity measures. The T&C intensity is measured based on Equations 3, 4 and 5. To obtain the Hawkes intensity, we first estimate the parameters in Equation (1). Table 1 reports the estimation results for these parameters for $|T| > 4$. As the table shows, the parameters for drift bursts λ_t are overwhelmingly highly significant, with their magnitude in line with those of Clements and Liao (2017), Table 2. Also, the parameters of positive drift burst λ_t^+ and negative drift burst λ_t^- are systematically significant.

Table 1. Coefficient estimates of Equation (1) ($\lambda_t = \mu + \sum_{t_q<t} \alpha e^{-\beta(t-t_q)}$) for $|T| > 4$. The brackets below the coefficients are t-statistics.

	μ	α	β	Log-like
λ_t	0.006	0.010	0.010	2,062.9
	(1.48)	(25.86)	(25.00)	
λ_t^+	0.013	0.013	0.016	1,203.6
	(3.69)	(14.91)	(14.49)	
λ_t^-	0.042	0.029	0.051	1,431.0
	(10.20)	(11.40)	(10.74)	

We then recursively estimate the Hawkes drift burst intensity, based on the estimated parameters for Equation (1). The upper panel of Figure 2 depicts the time series of the Hawkes drift burst intensity estimates. The black, red, and blue lines denote the drift bursts' intensity, negative intensity, and positive intensity, respectively. As the panel show, both the drift burst intensity and negative drift burst intensity are volatile over the sample periods and reach their highest during the 2008 financial crisis and the 2020 pandemic. In contrast, the positive drift burst intensity is generally smaller, more peaceful, and has a less obvious cyclical pattern during the two recessions.

The lower panel shows the time series for the T&Z intensity estimates. As the figure shows, the cyclical patterns of T&Z intensities are qualitatively very similar to those of the above Hawkes intensity. Besides the similarity, a difference is that Hawkes intensity is rougher than T&Z intensity. This is because Hawkes intensity uses an exponential weighted average, which gives more weighting to recent observations.

Figure 2. Time series of drift burst intensity for $|T| > 4$

Table 2. Descriptive Statistics

	Min	Mean	Q25	Median	Q75	Max
$RV \times 10^4$	0.011	0.881	0.198	0.347	0.722	62.50
$SV \times 10^4$	0.041	0.971	0.240	0.455	1.021	26.48
			Hawkes (1971)			
λ_t	0.089	0.167	0.153	0.165	0.181	0.227
λ_t^+	0.053	0.076	0.067	0.074	0.082	0.116
λ_t^-	0.045	0.096	0.083	0.093	0.108	0.143
			Tauchen and Zhou (2011)			
λ_t	0.109	0.162	0.141	0.157	0.179	0.254
λ_t^+	0.040	0.074	0.063	0.071	0.083	0.115
λ_t^-	0.050	0.093	0.071	0.088	0.109	0.155

4 Predicting volatility using drift burst intensity

The benchmark model is the HAR-DV model of Andersen et al. (2021),

$$\bar{V}_{t,t+h} = \beta_0 + \beta_d\, DV_{t-1,t} + \beta_w\, DV_{t-5,t} + \beta_m\, DV_{t-22,t} + \varepsilon_t, \quad (6)$$

where

$$DV_{t,t+h} = \frac{1}{\mathrm{h}} \sum_{i=1}^{h} DV_{t+i}\,, h = 1,2,\dots,$$

and $\bar{V}_{t+h|t}$ is the h-day average cumulative volatility (following Patton and Sheppard (2015) and Andersen et al. (2021),

$$\bar{V}_{t,\mathrm{t+h}} = \frac{1}{\mathrm{h}} \sum_{i=1}^{h} V_{t+i}.$$

To explore the importance of drift burst intensity in predicting future volatility, we add the drift burst intensity component to the HAR-DV model, resulting in the following HAR-DV-Intensity (HAR-DVI) model,

$$\bar{V}_{t,t+h} = \beta_0 + \beta_\lambda\, \lambda_{t-1} + \beta_d\, DV_{t-1,t} + \beta_w\, DV_{t-5,t} + \beta_m\, DV_{t-22,t} + \varepsilon_t, \quad (7)$$

Analogously, we formulate the HAR-DV -Signed Intensity (HAR-DVI$^{\pm}$) model by including the positive and negative drift burst intensity,

$$\bar{V}_{t,t+h} = \beta_0 + \beta_{\lambda^+}\, \lambda^{+}_{t-1,t} + \beta_{\lambda^-}\, \lambda^{-}_{t-1,t} + \beta_d\, DV_{t-1,t} + \beta_w\, DV_{t-5,t} + \beta_m\, DV_{t-22,t} + \varepsilon_t. \quad (8)$$

4.1 In-sample estimation

Table 3 reports the OLS in-sample results for the HAR-DV model, and the competing drift burst intensity augmented models for $|T| > 4$. The forecasting target is the conventional 5-minute RV, $\bar{V}_{t,t+h} = \overline{RV}_{t,t+h}$. We consider the forecasting horizons $h = 1$, 5, 22, 44, and 66, with the longer horizons covering expiries of the heavily traded options. The brackets are HAC

robust t-statistics by Newey and West (1987). Following Corsi and Renò (2012), the bandwidth used for the HAC is 2(h + 1), where h is the lead length of the left-hand-side variable.

The results reveal the importance of including drift burst information for the in-sample. The drift burst intensity strongly increases future volatility: for both Hawkes and T&Z measures and across all horizons, the coefficient of drift burst intensity is systematically positive and highly significant. Moreover, including the drift burst component leads to a substantial increase in the model goodness of fit. For example, for $h = 66$, the HAR-DVI model based on the Hawkes intensity gains 27.9% R^2 against the HAR-DV model.

The results also evidence the in-sample advantage of separating the drift burst signs. Across all of these different scenarios, drift burst signs show a clear leverage effect on future volatility, with the negative (positive) drift bursts strongly increasing (weakly affecting) volatility. Moreover, the negative sign has a more prominent impact than the unsigned: the t-statistic of negative drift bursts is generally more positive than that of (unsigned) drift bursts. Further, decomposing drift burst signs yield a further increase in the goodness of fit of the drift burst augmented model. In particular, for $h = 66$ for the Hawkes measure, the HAR-DVI$^{\Delta}$ model has a 5.0% greater R^2 than the HAR-DVI model.

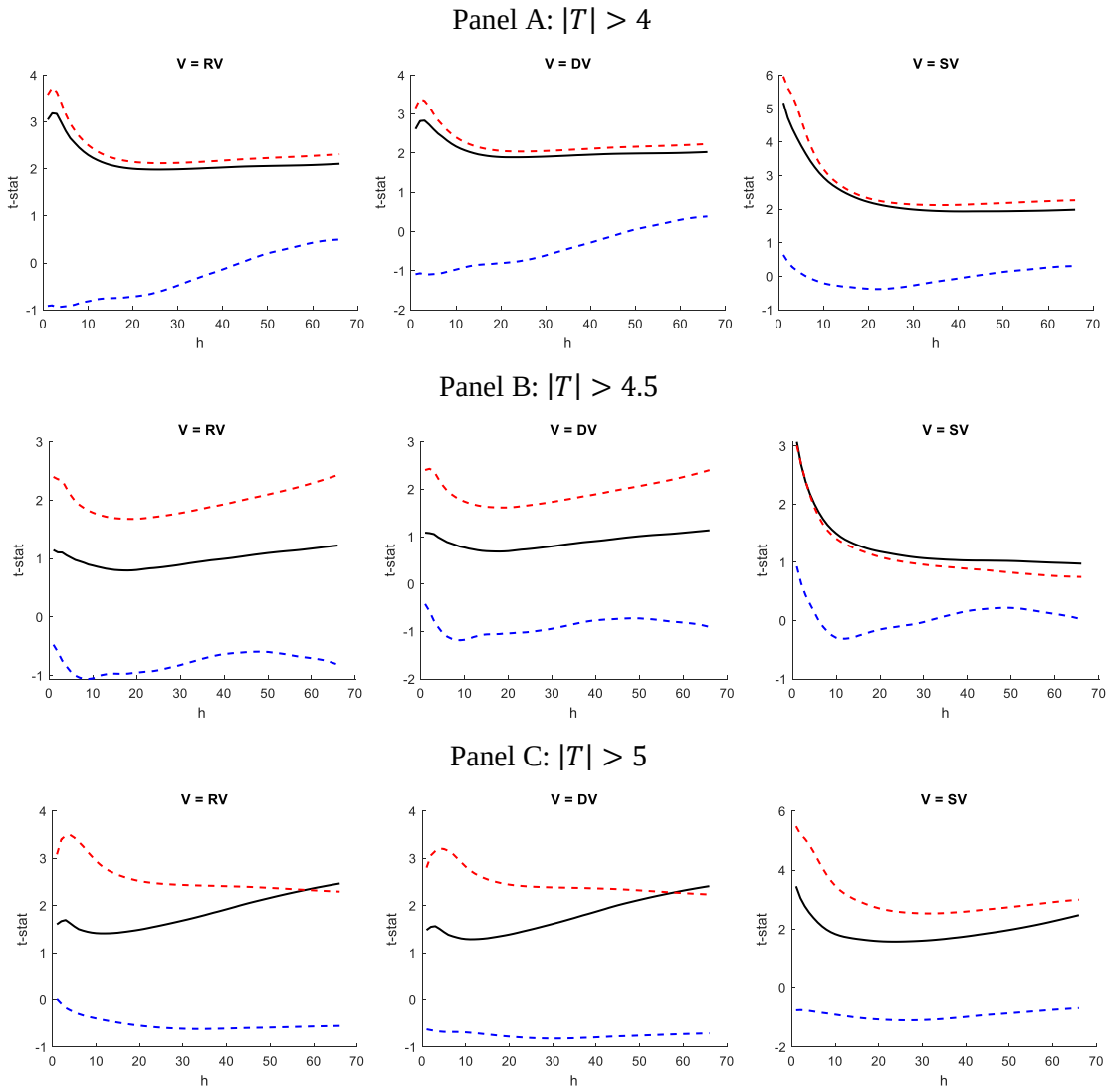
We then extend the in-sample results in Table 3 to a broader range of forecasting targets and more conservative critical values Φ. Figure 3 reports these results for the Hawkes measure. To economize the space, only the t-statistics of drift burst coefficients are reported. Panel A, B and C of this figure reports the results for the $|T| > 4$, $|T| > 4.5$, and $|T| > 5$. In each panel, we report the results for the RV, DV and SV forecasts, with the black, red, and blue lines denoting the t-statistics of drift burst intensity, positive drift burst intensity, and negative drift burst intensity, respectively.

Table 3. In-sample estimation results for $|T| > 4$ for RV forecast

Notes: The brackets are the HAC robust p-values. The intercept result is not reported.

		β_{λ^+}	β_{λ^-}	β_{λ}	β_d	β_w	β_m	R^2
h=1					0.390	0.596	0.030	0.593
					(3.22)	(3.95)	(0.38)	
	Hawkes			0.340	0.389	0.595	0.016	0.593
				(3.09)	(3.22)	(3.95)	(0.20)	
	T&Z			0.233	0.388	0.596	0.016	0.593
				(2.39)	(3.22)	(3.95)	(0.20)	
	Hawkes	-0.198	0.530		0.387	0.596	0.011	0.594
		(-0.92)	(3.58)		(3.21)	(3.95)	(0.13)	
	T&Z	0.181	0.235		0.388	0.596	0.016	0.593
		(1.11)	(2.75)		(3.22)	(3.95)	(0.19)	
h=5					0.380	0.455	0.100	0.646
					(3.60)	(2.96)	(0.94)	
	Hawkes			0.575	0.378	0.454	0.076	0.649
				(2.85)	(3.61)	(2.95)	(0.69)	
	T&Z			0.376	0.377	0.456	0.077	0.648
				(2.32)	(3.61)	(2.94)	(0.69)	
	Hawkes	-0.281	0.890		0.375	0.455	0.067	0.651
		(-0.92)	(3.16)		(3.62)	(2.98)	(0.60)	
	T&Z	0.296	0.378		0.377	0.456	0.076	0.648
		(1.12)	(2.86)		(3.61)	(2.94)	(0.67)	
h=22					0.229	0.333	0.147	0.462
					(3.49)	(2.62)	(1.23)	
	Hawkes			1.226	0.225	0.329	0.096	0.479
				(1.99)	(3.61)	(2.60)	(0.71)	
	T&Z			0.775	0.224	0.334	0.100	0.476
				(1.84)	(3.66)	(2.60)	(0.73)	
	Hawkes	-0.403	1.815		0.220	0.331	0.078	0.488
		(-0.69)	(2.13)		(3.68)	(2.65)	(0.57)	
	T&Z	0.696	0.745		0.224	0.334	0.098	0.476
		(1.19)	(2.06)		(3.66)	(2.60)	(0.70)	
h=44					0.158	0.257	0.114	0.323
					(4.51)	(2.11)	(1.67)	
	Hawkes			1.871	0.151	0.252	0.037	0.370
				(2.02)	(4.85)	(2.07)	(0.41)	
	T&Z			1.143	0.150	0.259	0.046	0.360
				(2.10)	(4.96)	(2.10)	(0.52)	
	Hawkes	0.002	2.547		0.146	0.253	0.014	0.386
		(0.00)	(2.20)		(5.02)	(2.10)	(0.15)	
	T&Z	1.134	1.051		0.149	0.260	0.044	0.360
		(1.58)	(2.20)		(4.97)	(2.10)	(0.49)	
h=66					0.118	0.168	0.152	0.251
					(5.03)	(2.08)	(2.41)	
	Hawkes			2.106	0.111	0.161	0.066	0.321
				(2.06)	(5.53)	(2.02)	(0.80)	
	T&Z			1.286	0.109	0.169	0.077	0.305
				(2.31)	(5.70)	(2.07)	(0.93)	
	Hawkes	0.378	2.722		0.106	0.161	0.043	0.337
		(0.50)	(2.31)		(5.74)	(2.05)	(0.54)	
	T&Z	1.383	1.129		0.109	0.169	0.075	0.304
		(1.94)	(2.27)		(5.72)	(2.06)	(0.90)	

Figure 3. HAC robust t-statistics of Hawkes drift burst intensity as a function of forecasting horizons. Black, red, and blue lines denote the t-statistics of drift burst intensity, positive drift burst intensity, and negative drift burst intensity, respectively.



The results show that the positive effect of drift burst is consistent for all of these cases but is sometimes not highly significant for a more conservative Φ. For example, the coefficients of drift bursts are not highly significant for RV and DV targets for $|T| > 4.5$ and $|T| > 5$. The leverage effect of drift bursts is more robust: for all cases of the three panels, the coefficient of negative drift burst is systematically negative and significant. The more consistent impact of negative drift bursts on future volatility again indicates the importance of decomposing drift burst signs.

Figure 4. HAC robust t-statistics of drift burst intensity by Tauchen and Zhou (2011) as a function of forecasting horizons. Black, red, and blue lines denote the t-statistics of drift burst intensity, positive drift burst intensity, and negative drift burst intensity, respectively.

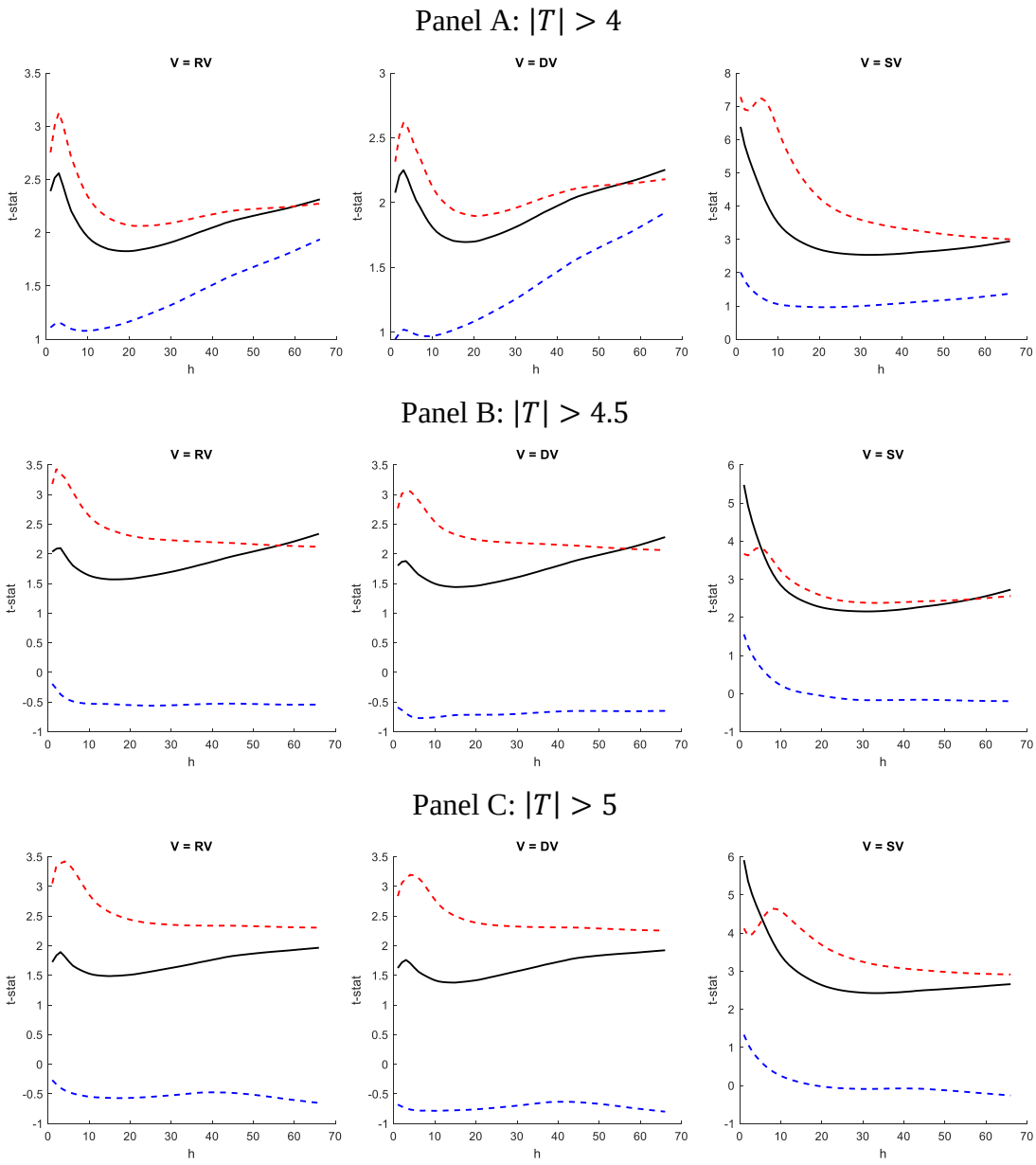


Figure 4 depicts the in-sample t-statistics for T&Z drift burst intensities for extended critical values and forecast targets. The pattern of Figure 4 is very similar to that of Figure 3: across all of these different scenarios, the drift burst intensity has a positive effect on volatility while the intensity of drift burst signs has a leverage effect on volatility.

Overall, including information on drift bursts and drift burst signs strongly predict future volatility and improves the model's in-sample performance. However, whether the

improvements for in-sample are also of practical importance remains an empirical question to be answered in the out-of-sample forecasting exercise in the following section 4.2.

4.2 Out-of-sample forecasting

This section compares the drift burst intensity augmented HAR-DV models with the HAR-DV model, in terms of the out-of-sample performance. Motivated by the above limited in-sample evidence of positive drift burst intensity, we formulate a new HAR-DVI^{-} model by including the negative drift burst intensity only,

$$\bar{V}_{t,t+h} = \beta_0 + \beta_{\lambda^-}\,\lambda^-_{t-1} + \beta_d\,DV_{t-1,\mathrm{t}} + \beta_w\,DV_{t-5,\mathrm{t}} + \beta_m\,DV_{t-22,\mathrm{t}} + \varepsilon_t. \quad (6)$$

For in-sample estimation, we rely on OLS to generate both a rolling window forecast using the prior 1000 trading days and an increasing window forecast using all prior observations, starting from an initial set of 1000 trading days. The intensities and models are all estimated only within each window, to ensure that only past/current information is used for forecasting. These estimated parameters are then used to construct h-step-ahead out-of-sample forecasts which incorporate new information as it becomes available.

The following two loss functions are applied for evaluating the forecasting performance:

(a) mean square error (MSE),

$$\mathrm{MSE}\left(\bar{V}_{t,t+h}, F_{t,t+h}\right) = \left(\bar{V}_{t,t+h} - F_{t,t+h}\right)^2,$$

(b) gaussian Quasi-likelihood (QLIKE) loss function:

$$\mathrm{QLIKE}\left(\bar{V}_{t+h|t}, F_{t,t+h}\right) = \frac{\bar{V}_{t,t+h}}{F_{t,t+h}} - \ln\frac{\bar{V}_{t,t+h}}{F_{t,t+h}} - 1,$$

where $F_{t+h|t}$ denotes the h-day-ahead forecast. The MSE and QLIKE are both unbiased loss functions (Patton, 2011) and are widely applied by key research (Bollerslev et al., 2016,

Andersen et al., 2021). The statistical significance difference is evaluated via the Diebold–Mariano-West (DMW) statistic [8] developed by Diebold and Mariano (1995) and West (1996), with adjustment to the Newey–West Heteroskedasticity and Autocorrelation Corrected (HAC) standard errors.

4.2.1 Drift burst information

The aim of this section is to explore whether including drift burst intensity leads to an improvement in predictive accuracy over the HAR-DVI models. Table 4 presents DM statistics for comparing the augmented HAR-DV models with the benchmark HAR-DV model for the case of the Hawkes intensity measure. A positive (negative) statistic indicates the competing model is superior (inferior) to the benchmark model, with the significant and positive statistics (at 5% level) indicated in bold.

The upper panel of Table 4 shows the out-of-sample results for RV forecast [9]. The results support the out-of-sample value of drift burst information: for longer horizons ($h \geq 5$), $|T| > 4.5$, and both forecast windows, the HAR-DVI, HAR-DVI$^{\pm}$, and HAR-DVI^{-} models, with few exceptions, overwhelmingly outperform the benchmark HAR-DV model. For the daily forecast ($h = 1$), although these three augmented models generally perform worse than the benchmark model, the HAR-DVI and HAR-DVI^{-} models are never significantly inferior. The predictive advantage of the drift bursts is somewhat less clear but still qualitatively consistent for $|T| > 4$ and $|T| > 5$, reported in the left and right parts of the upper panel, respectively.

The lower panel of Table 4 is for SV forecast. As suggested by Andersen et al. (2021), the expost SV measure refers directly to a concurrent estimate of the underlying volatility level

[8] The DMW results in this paper were obtained using the robust_loss_1 function from Andrew Patton's Matlab code page, http://public.econ.duke.edu/~ap172/

[9] As in Andersen et al. (2021), we also make the out-of-sample results for DV forecast (unreported). The pattern is very similar to that for RV forecast in Table 4.

thus providing a less noisy benchmark for the assessment of forecast performance. As the panel shows, the out-of-sample evidence of the augmented models appears to be even stronger [10]. For the cases in the panel, the intensity models provide much more forecasts which are significantly better than the HAR-DV model. Moreover, the intensity models are able to provide significantly better forecasts for both short and long horizons (e.g., HAR-DVI^{-} or HAR-DVI for $|T| > 4.5$).

Table 5 reports the same out-of-sample results as Table 4 but for the T&Z intensity measure case. The patterns here are very close to those in Table 4: the augmented models are again superior to the HAR-DV model for longer horizons for forecast RV, or all horizons for forecast SV. The consistent results indicate that the predictability of drift burst intensity holds for the T&Z intensity measure of Tauchen and Zhou (2011).

4.2.2 Drift burst signs

The goal of this section is to investigate whether disentangling drift burst signs leads to an improvement in predictive accuracy over the drift burst intensity models. We begin by comparing the HAR-DV-I$^{\pm}$ and HAR-DV-I^{-}models with the HAR-DV-I model in Table 4. We only discuss the patterns in Table 4 for the Hawkes intensity as the patterns in Table 5 for the T&Z intensity are very similar. The first thing to note is that the HAR-DV-I$^{\pm}$ can perform better than the HAR-DV-I model. For example, for $|T| > 4$ for both SV and RV forecasts, the DM statistics of the HAR-DV-I$^{\pm}$ model is generally more positive than those of the HAR-DV-I model. However, this superiority is not robust for some other cases. For the cases in $|T| > 4$ for both SV and RV, the HAR-DV-I$^{\pm}$ model provides a considerable number of less positive or more negative DM statistics over the HAR-DV-I model.

[10] The DM statistics discrepancies between upper and lower panels also reflect the different sample periods, as the SV forecasts are initiated only during the great financial crisis of 2008-2009. However, we confirm that the difference remains substantial, even if we generate the forecasts over the identical time period.

The results for the HAR-DVI^{-} model reveals a cleaner picture for the advantage of separating drift burst signs. Except for some cases for $|T| > 5$ for RV forecasts, the DM statistics of the HAR-DVI^{-} model is systematically more positive or less negative than the HAR-DVI model. The stronger evidence of the HAR-DVI^{-} model over the HAR-DVI$^{\pm}$ model implies that the positive drift burst intensity does not help forecast volatility, which is consistent with its limited in-sample evidence. Overall, the results show that decomposing the drift burst signs improves the out-of-sample, and this improvement mainly comes from the negative sign of drift bursts.

4.2.3 Comparing two intensity measures

This section aims to compare the forecast value between Hawkes measures and the T&Z measure. The T&Z measure appears to be more beneficial in forecasting accuracy. Comparing the DM statistics for the T&Z measure in Table 4 with those for the Hawkes measure in Table 5, the T&Z measure provides a great number of more positive statistics than the Hawkes measure (e.g., $|T| > 4$ for the HAR-DVI model for SV forecast; $|T| > 5$ for the HAR-DVI^{-} model for RV forecast).

Table 4. Diebold–Mariano statistic for Hawkes drift burst intensity

			$\|T\| > 4$			$\|T\| > 4.5$			$\|T\| > 5$		
			HAR-DVI	HAR-DVI$^{\pm}$	HAR-DVI^{-}	HAR-DVI	HAR-DVI$^{\pm}$	HAR-DVI^{-}	HAR-DVI	HAR-DVI$^{\pm}$	HAR-DVI^{-}
Forecast RV											
h=1	MSE	RW	-1.27	-1.50	-1.06	-0.38	-1.25	-0.95	-1.11	-1.80	-0.97
		IW	0.25	0.60	0.24	-0.18	-0.64	-0.46	-0.88	-1.05	-0.86
	QLIKE	RW	-2.84	-1.11	-1.55	-1.00	-4.44	-1.81	-1.49	-3.85	-3.67
		IW	-3.18	-2.70	-2.89	0.82	-2.77	-0.69	-3.81	-3.76	-3.64
h=5	MSE	RW	0.00	0.01	0.16	0.94	0.35	0.51	0.11	-0.67	0.39
		IW	0.71	**1.98**	1.65	0.15	0.20	0.35	0.44	0.42	0.59
	QLIKE	RW	-1.32	-1.27	-0.64	0.43	-2.48	-0.83	-2.89	-2.83	-2.32
		IW	-3.11	-1.72	-1.80	1.17	-1.37	0.08	-1.00	-3.31	-2.40
h=22	MSE	RW	1.58	1.10	0.82	1.40	1.39	1.67	0.77	0.54	1.05
		IW	0.80	**1.98**	1.92	0.65	0.71	0.87	1.05	1.16	1.22
	QLIKE	RW	-2.23	-1.11	-0.96	1.05	-0.62	1.25	-1.19	-0.91	-0.42
		IW	-2.12	-0.70	-0.02	0.67	-0.30	1.42	1.18	-2.43	0.26
h=44	MSE	RW	1.80	1.10	0.81	1.84	1.67	**2.06**	0.77	1.34	1.13
		IW	1.08	**2.00**	**1.99**	0.14	1.69	1.73	1.49	1.13	1.20
	QLIKE	RW	-1.59	-1.12	-0.35	**2.30**	0.90	**1.97**	-1.01	0.09	-0.51
		IW	-2.22	-1.79	0.29	1.23	**1.99**	**2.05**	**2.06**	-1.40	-1.18
h=66	MSE	RW	1.80	0.88	0.71	**2.17**	1.82	**2.09**	0.69	1.82	1.14
		IW	1.30	**2.24**	**2.21**	-0.02	**2.08**	**2.06**	1.61	0.99	1.05
	QLIKE	RW	0.75	1.43	0.60	**2.73**	-0.57	1.79	-1.28	0.87	-0.44
		IW	-2.26	-1.11	1.64	1.54	**2.58**	**2.50**	**2.61**	-0.87	-2.11
Forecast SV											
h=1	MSE	RW	**3.24**	**3.09**	**3.53**	1.91	**2.15**	**2.40**	**2.88**	1.98	**3.03**
		IW	-0.35	0.83	1.10	1.58	1.64	**2.09**	1.31	1.89	**3.99**
	QLIKE	RW	**2.25**	**2.66**	**2.86**	1.07	1.34	1.01	-1.92	-2.25	0.50
		IW	**2.21**	**2.06**	1.78	**2.12**	**2.83**	1.85	1.65	-1.11	**7.31**
h=5	MSE	RW	**2.31**	**3.62**	**3.28**	**2.14**	**1.97**	**2.37**	**2.37**	1.40	**2.51**
		IW	-0.31	1.32	1.60	1.20	1.53	1.87	1.07	3.65	**2.45**
	QLIKE	RW	1.81	**2.51**	**2.62**	1.43	0.79	0.66	-2.48	-1.81	-0.11
		IW	1.47	2.20	1.95	1.89	**3.09**	**2.40**	**2.46**	-3.17	**6.75**
h=22	MSE	RW	1.21	1.91	1.91	1.63	1.46	1.74	1.31	0.82	1.38
		IW	0.40	1.74	**1.96**	0.93	1.57	1.94	0.83	1.93	1.61
	QLIKE	RW	0.85	1.58	1.57	1.45	1.03	1.09	-1.93	-1.14	0.66
		IW	-1.10	1.79	1.87	1.65	0.72	**2.03**	**2.26**	-3.19	**3.40**
h=44	MSE	RW	1.57	**1.98**	**2.16**	1.70	1.74	**2.05**	1.38	0.86	1.61
		IW	0.57	1.95	**2.37**	0.83	1.88	**2.20**	1.06	**2.11**	**2.01**
	QLIKE	RW	-0.75	1.77	**1.98**	1.95	1.59	1.62	-0.74	-1.11	1.41
		IW	-2.09	1.40	**2.31**	1.53	1.40	0.68	**2.65**	-1.15	**3.26**
h=66	MSE	RW	1.69	**2.21**	**2.41**	**2.18**	1.85	**2.22**	1.62	1.03	**1.97**
		IW	0.66	**2.14**	**3.09**	0.87	1.78	1.90	1.17	**1.96**	**2.20**
	QLIKE	RW	-1.12	1.71	**2.43**	**2.27**	1.68	1.40	-1.57	-1.52	**1.96**
		IW	-1.16	-0.99	**3.34**	1.48	-1.11	-0.04	**2.86**	-2.58	1.12

Table 5. Diebold–Mariano statistic for drift burst intensity of Tauchen and Zhou (2011)

			$\|T\| > 4$			$\|T\| > 4.5$			$\|T\| > 5$		
			HAR-DVI	HAR-DVI$^{\pm}$	HAR-DVI^{-}	HAR-DVI	HAR-DVI$^{\pm}$	HAR-DVI^{-}	HAR-DVI	HAR-DVI$^{\pm}$	HAR-DVI^{-}
Forecast RV											
h=1	MSE	RW	-1.28	-1.45	-1.07	-1.50	-1.97	-1.13	-0.50	-2.39	-1.89
		IW	-1.06	-1.12	-0.70	-0.57	-0.67	-0.21	-0.52	-1.57	-1.01
	QLIKE	RW	-3.36	-2.23	-1.60	-1.40	-2.14	-1.15	-3.02	-3.93	-2.30
		IW	-1.02	-0.88	1.29	-0.50	-1.46	-0.72	-3.02	-1.83	-1.82
h=5	MSE	RW	-0.20	0.55	0.34	-0.40	-0.27	0.24	0.03	-0.94	-1.04
		IW	-0.02	-0.05	0.80	0.54	1.00	1.14	0.22	-0.20	0.59
	QLIKE	RW	-2.11	-1.73	-0.14	-0.56	-1.92	-0.04	-2.27	-3.04	-1.25
		IW	-1.47	-1.98	1.67	-1.11	-1.23	-0.35	-2.31	-2.62	-1.94
h=22	MSE	RW	1.21	1.34	1.09	0.64	0.72	1.07	0.92	1.23	1.46
		IW	1.01	1.06	1.35	1.07	1.76	1.67	1.06	1.28	1.53
	QLIKE	RW	-0.90	-0.43	0.02	0.01	-1.01	0.46	0.49	0.92	1.62
		IW	-1.04	-2.00	1.83	-0.19	-1.55	0.62	0.03	-1.74	-0.67
h=44	MSE	RW	1.36	1.18	1.14	0.93	0.95	1.11	1.30	1.97	1.98
		IW	1.16	1.19	1.48	1.51	**2.27**	**2.10**	1.75	1.71	1.75
	QLIKE	RW	-1.03	-1.15	1.18	0.36	0.70	1.63	1.44	**2.44**	**2.73**
		IW	-1.33	-1.86	**2.62**	-2.13	-1.00	-0.39	1.02	-0.67	**2.26**
h=66	MSE	RW	1.81	1.05	1.29	0.88	1.07	1.11	0.17	**2.10**	1.82
		IW	1.33	1.31	1.56	1.74	**2.52**	**2.34**	1.09	**2.00**	1.87
	QLIKE	RW	-1.73	-1.21	0.00	-0.09	-1.11	**1.99**	1.85	**3.11**	**3.39**
		IW	0.26	-1.20	**3.08**	-2.09	0.93	-0.79	1.43	-0.88	2.91
Forecast SV											
h=1	MSE	RW	**2.08**	**0.91**	1.27	0.60	0.00	1.13	-2.24	-1.21	1.54
		IW	**3.10**	1.36	**5.39**	1.38	0.90	1.34	1.23	1.42	1.66
	QLIKE	RW	**2.99**	**1.16**	1.80	1.80	0.92	1.21	-2.64	0.00	**2.10**
		IW	**6.64**	-1.79	**8.97**	-1.49	1.79	**2.99**	**3.48**	**6.04**	**7.47**
h=5	MSE	RW	1.54	**1.24**	1.20	0.57	1.22	1.01	-0.92	-0.02	1.33
		IW	**2.32**	**1.21**	**3.99**	1.29	0.99	1.22	1.11	1.74	**2.50**
	QLIKE	RW	**2.30**	**1.24**	**2.61**	**2.66**	1.49	**2.08**	-1.91	0.52	**2.19**
		IW	**5.99**	-1.36	**8.50**	1.32	**3.57**	**3.56**	**3.25**	**6.70**	**7.05**
h=22	MSE	RW	1.69	1.66	1.57	1.13	1.29	1.40	0.50	1.64	1.68
		IW	**1.82**	1.06	**2.64**	1.20	1.05	1.23	1.15	1.93	**2.63**
	QLIKE	RW	1.10	1.36	**2.11**	1.92	1.53	1.80	-0.93	1.11	1.66
		IW	**3.71**	-1.65	**4.58**	**3.36**	**2.86**	**3.06**	**2.18**	**3.55**	**2.57**
h=44	MSE	RW	1.91	1.68	1.60	1.42	1.04	1.41	0.99	1.75	1.82
		IW	**2.38**	1.22	**3.00**	1.59	1.39	1.71	1.57	**2.54**	**3.25**
	QLIKE	RW	1.62	1.76	**2.58**	**2.06**	**2.18**	**2.24**	0.17	1.40	**2.19**
		IW	**3.94**	-0.60	**4.29**	**3.56**	**3.04**	**3.55**	**2.29**	-2.35	**4.38**
h=66	MSE	RW	**2.20**	1.43	1.63	1.86	1.00	1.58	1.36	1.43	**1.98**
		IW	**2.73**	0.79	**2.63**	1.83	1.52	**2.11**	1.91	**3.08**	**3.94**
	QLIKE	RW	1.25	-0.42	**3.12**	**2.17**	**2.54**	**2.71**	-0.57	-0.14	**2.75**
		IW	**4.35**	2.64	**4.59**	**3.99**	**3.47**	**4.42**	-0.35	-1.49	**3.66**

5 Robustness analysis

This section aims to check the robustness of our results in section 4. Only robustness results for the critical value $|T| > 4$ are reported but the (unreported) results for more conservative critical value are broadly speaking in line with those reported here.

5.1 Volatility transformations

This section is to examine if the predictive advantages of the drift burst intensity are still observed under some important volatility transformations. As in Andersen et al. (2007), we consider the logarithm and square root transformations. And the logarithm and square root transformed HAR-DV model are defined by,

$$\left(\bar{V}_{t,t+h}\right)^{1/2} = \beta_0 + \beta_d \left(DV_{t-1,t}\right)^{1/2} + \beta_w \left(DV_{t-5,t}\right)^{1/2} + \beta_m \left(DV_{t-22,t}\right)^{1/2} + \varepsilon_t, \quad (3)$$

$$log\left(\bar{V}_{t,t+h}\right) = \beta_0 + \beta_d \, log\left(DV_{t-1,t}\right) + \beta_w \; log\left(DV_{t-5,t}\right) + \beta_m \, log\left(DV_{t-22,t}\right) + \varepsilon_t. \quad (3)$$

We term these two models the sqr-HAR-DV and log-HAR-DV models. We again formulate the drift burst augmented models by adding the drift burst intensity to these two models.

We first explore the in-sample estimation results of the augmented sqr-HAR-DV and log-HAR-DV models. Figure 5 reports the coefficient estimates (t-statistics) for the drift burst intensity components, with the upper and lower panel showing the results for the square root and logarithm transformations, respectively. The general pattern in both the upper and lower panel directly mirrors that of the estimates for the non-transformation case in Figure 3&4: for both Hawkes and T&Z measures for all RV, DV, and SV forecast targets, the drift burst intensity significantly increases the future volatility, with the negative (positive) drift burst intensity strongly increase (weakly decrease) the future volatility. This indicates that the impact of (signed) drift burst intensity on future volatility is robust to the square root and logarithm volatility transformations.

Figure 5. HAC robust t-statistics of drift burst intensity as a function of forecasting horizons. Black, red, and blue lines denote the t-statistics of drift burst intensity, positive drift burst intensity, and negative drift burst intensity, respectively. $|T| > 4$

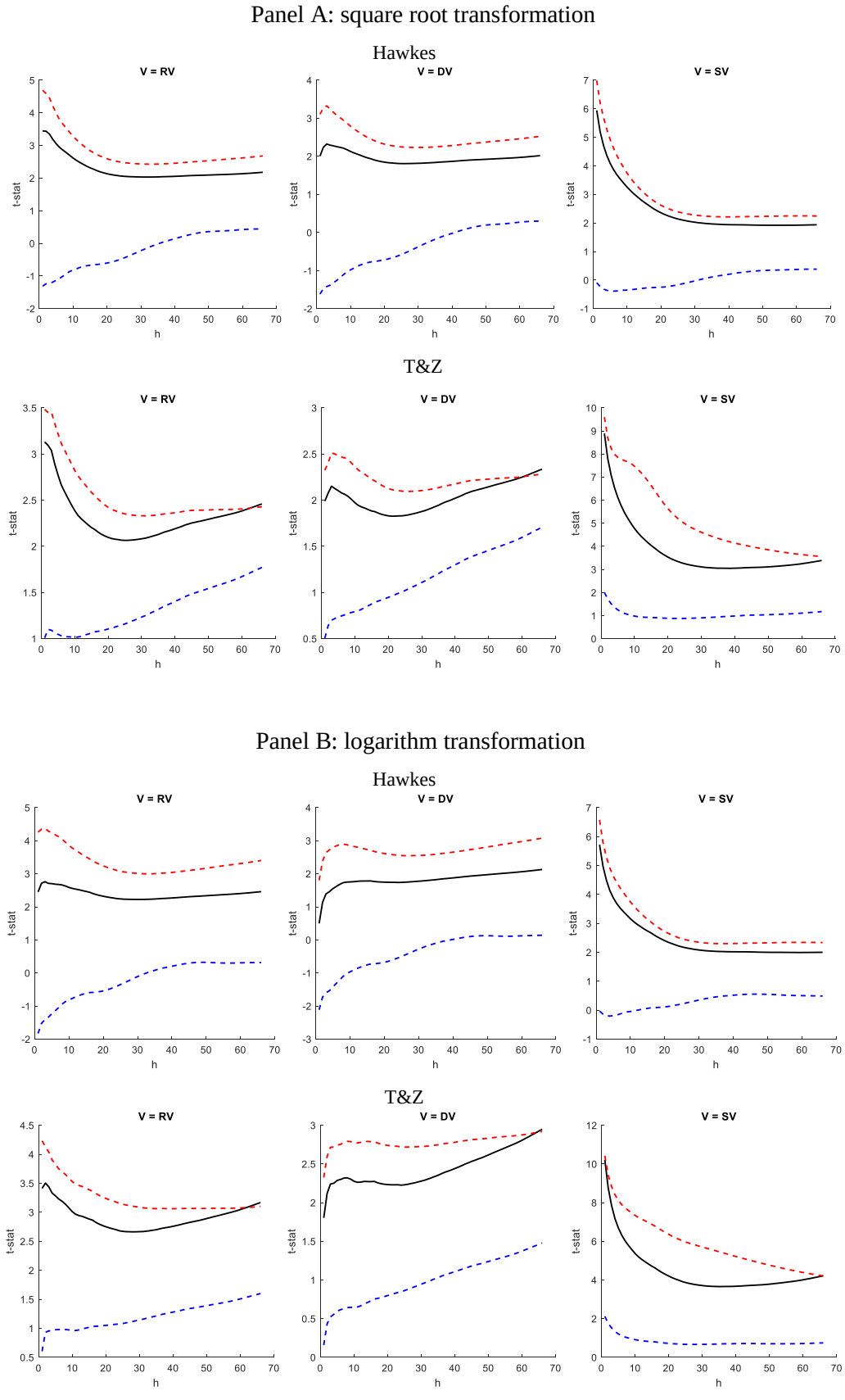


Table 6. Diebold–Mariano statistic for root volatility transformations for $|T| > 4$

			Hawkes			T&Z		
			sqr-HAR-DVI	sqr-HAR-DVI$^{\pm}$	sqr-HAR-DVI^{-}	sqr-HAR-DVI	sqr-HAR-DVI$^{\pm}$	sqr-HAR-DVI^{-}
			Forecast $RV^{1/2}$					
h=1	MSE	RW	-0.34	-0.27	-0.02	-0.65	-0.23	0.26
		IW	-0.18	1.62	1.57	0.12	0.03	1.36
	QLIKE	RW	-0.19	0.48	0.95	0.17	0.09	1.34
		IW	-0.72	-0.01	-0.06	0.21	0.33	1.87
h=5	MSE	RW	0.71	0.81	0.58	0.36	1.06	0.90
		IW	0.74	**2.24**	1.94	0.93	0.97	1.78
	QLIKE	RW	-0.10	1.09	1.22	0.04	1.12	1.92
		IW	-0.68	0.79	0.49	0.54	0.61	**2.15**
h=22	MSE	RW	1.49	1.26	0.67	1.25	1.69	1.52
		IW	0.99	**2.28**	**2.04**	1.44	1.47	**2.26**
	QLIKE	RW	-0.07	1.24	1.18	0.53	**2.16**	**2.57**
		IW	-0.68	1.39	1.19	1.37	1.23	**2.40**
h=44	MSE	RW	**2.32**	1.39	0.79	1.94	**1.98**	**1.96**
		IW	**1.99**	**2.43**	**2.34**	1.92	1.88	**2.69**
	QLIKE	RW	0.61	1.74	1.51	1.08	**2.51**	**2.45**
		IW	0.23	1.91	1.81	**2.01**	1.64	**2.79**
h=66	MSE	RW	**2.42**	1.15	0.75	**2.38**	**2.00**	**2.18**
		IW	**2.28**	**2.75**	**2.69**	**2.34**	**2.21**	**3.08**
	QLIKE	RW	1.55	**2.12**	1.72	1.00	**2.65**	**2.70**
		IW	0.78	**2.44**	**2.34**	**2.47**	**1.98**	**3.12**
			Forecast $SV^{1/2}$					
h=1	MSE	RW	**2.34**	**3.12**	**3.51**	**3.39**	**2.08**	**2.63**
		IW	0.11	0.96	1.14	**3.30**	**2.39**	**5.84**
	QLIKE	RW	**2.42**	**3.13**	**2.93**	**3.36**	1.79	**2.32**
		IW	0.60	0.15	-0.06	**4.10**	**2.65**	**6.99**
h=5	MSE	RW	1.41	3.31	**3.26**	**2.70**	1.77	**2.27**
		IW	-0.06	1.31	1.45	**2.80**	**2.17**	**4.92**
	QLIKE	RW	1.53	**2.65**	**2.44**	**2.83**	1.85	**2.82**
		IW	0.55	0.79	0.57	**3.59**	**2.44**	**6.10**
h=22	MSE	RW	0.93	**2.71**	**2.66**	1.89	1.63	1.87
		IW	0.31	1.69	**1.97**	**2.28**	1.81	**3.48**
	QLIKE	RW	0.71	**2.31**	**2.07**	**2.16**	**2.28**	**3.07**
		IW	0.76	1.71	**1.63**	**2.86**	**2.19**	**3.99**
h=44	MSE	RW	**1.31**	2.31	**2.47**	**2.50**	**2.09**	**2.30**
		IW	0.64	1.83	**2.40**	**2.71**	**1.99**	**3.57**
	QLIKE	RW	1.09	**2.21**	**2.25**	**2.59**	**2.83**	**3.28**
		IW	1.13	**2.15**	**2.21**	**3.10**	**2.57**	**3.74**
h=66	MSE	RW	1.30	**2.46**	**2.60**	**2.83**	**2.22**	**2.54**
		IW	1.03	**1.99**	**3.01**	**2.92**	1.77	**3.30**
	QLIKE	RW	1.30	**2.47**	**2.61**	**2.70**	**3.37**	**3.72**
		IW	1.66	**2.57**	**2.86**	**3.35**	**2.91**	**3.85**

Table 7. Diebold–Mariano statistic for logarithm volatility transformations for $|T| > 4$
Notes: the DM statistics are based on the MSE function.

			Hawkes			T&Z		
			log-HAR-DVI	log-HAR-DVI$^{\pm}$	log-HAR-DVI^{-}	log-HAR-DVI	log-HAR-DVI$^{\pm}$	log-HAR-DVI^{-}
			Forecast log(RV)					
h=1	MSE	RW	-0.69	1.02	0.99	-0.33	0.38	1.22
		IW	-0.44	0.90	1.00	0.98	1.06	1.87
	QLIKE	RW	-0.85	0.93	0.88	-0.15	0.47	1.10
		IW	-0.39	0.98	1.16	0.91	0.97	1.71
h=5	MSE	RW	0.15	1.54	1.22	0.24	1.26	1.59
		IW	0.52	**1.53**	1.38	1.43	1.55	**2.33**
	QLIKE	RW	0.03	1.31	1.01	0.41	1.18	1.38
		IW	0.64	1.64	1.54	1.33	1.43	**2.17**
h=22	MSE	RW	0.86	1.14	0.72	0.79	**2.17**	**2.40**
		IW	0.90	**1.78**	**1.59**	1.92	**2.02**	**2.88**
	QLIKE	RW	0.94	0.72	0.30	0.97	1.88	**1.98**
		IW	1.06	1.80	1.64	1.67	1.79	**2.68**
h=44	MSE	RW	**2.05**	1.59	0.89	1.49	**3.11**	**3.06**
		IW	**2.29**	**2.19**	**2.14**	2.61	**2.61**	**3.36**
	QLIKE	RW	2.24	1.19	0.36	1.80	**2.76**	**2.73**
		IW	2.47	2.15	2.14	**2.39**	**2.44**	**3.30**
h=66	MSE	RW	**1.68**	1.47	1.01	1.26	**3.46**	**3.06**
		IW	**2.88**	**2.83**	**2.82**	**3.12**	**3.02**	**3.72**
	QLIKE	RW	1.85	**0.95**	0.41	1.68	**2.90**	**2.69**
		IW	3.03	**2.73**	**2.77**	**2.98**	**2.94**	**3.71**
			Forecast log(SV)					
h=1	MSE	RW	**2.30**	**3.68**	**3.59**	**3.54**	**2.29**	**2.93**
		IW	0.81	1.05	1.15	**3.39**	**3.09**	**5.92**
	QLIKE	RW	**1.98**	**3.64**	**3.71**	**3.47**	**2.32**	**2.91**
		IW	0.89	1.34	1.47	**3.19**	**2.98**	**5.65**
h=5	MSE	RW	1.16	**2.96**	**2.88**	**2.73**	1.94	**2.73**
		IW	0.53	1.25	1.34	**2.80**	**2.69**	**4.88**
	QLIKE	RW	0.84	**3.06**	**3.07**	**2.70**	1.92	**2.59**
		IW	0.54	1.43	1.56	**2.64**	**2.59**	**4.70**
h=22	MSE	RW	0.54	**2.41**	**2.59**	**2.26**	**2.10**	**2.97**
		IW	0.65	1.63	1.91	**2.61**	**2.67**	**4.15**
	QLIKE	RW	0.36	**2.27**	**2.72**	**2.15**	1.80	**2.55**
		IW	0.52	1.55	1.95	**2.34**	**2.46**	**4.00**
h=44	MSE	RW	**0.97**	2.54	**2.66**	**3.35**	**3.26**	**4.04**
		IW	1.02	1.90	**2.43**	**3.06**	**3.02**	**4.15**
	QLIKE	RW	0.92	**2.48**	**2.73**	**3.37**	**2.98**	**3.67**
		IW	0.86	**1.73**	**2.47**	**2.88**	**2.85**	**4.16**
h=66	MSE	RW	0.99	**2.51**	**2.79**	**3.70**	**3.95**	**4.62**
		IW	1.58	**2.18**	**2.97**	**3.23**	**3.19**	**4.09**
	QLIKE	RW	0.90	**2.50**	**2.81**	**3.83**	**3.55**	**4.26**
		IW	1.39	**1.96**	**3.01**	**3.05**	**2.89**	**4.01**

We next focus on forecast accuracy. Table 6 evaluates the out-of-sample of the augmented sqr-HAR-DV models, in terms of their DM statistics relative to that of the benchmark sqr-HAR-DV model. The upper and lower panel report the results for $\mathrm{RV}^{1/2}$ and $\mathrm{SV}^{1/2}$ forecast, and in each of these two panels results for the Hawkes and T&Z intensity measures are also reported. These results are not only confirmed but strengthened, compared to those in Table 4 and Table 5. In detail, for both $\mathrm{RV}^{1/2}$ and $\mathrm{SV}^{1/2}$ forecasts, the augmented sqr-HAR-DV models overwhelmingly outperform the benchmark model, with few exceptions (mostly from the sqr-HAR-DV model associated with the Hawkes measure for RV forecast). Moreover, the superiority of the sqr-HAR-DVI^{-} model over the sqr-HAR-DV model also holds true for the square root volatility transformations: the DM statistics of the sqr-HAR-DVI^{-} model is substantially more positive across these different cases in the table. Further, comparing the Hawkes and T&Z measure, the latter provides a greater number of significant DM statistics, for these different scenarios. Table 7 compares the augmented log-HAR-DV models with the log-HAR-DV model for the out-of-sample. The pattern is qualitatively in line with those in Table 8, indicating that our findings are also robust to logarithm volatility transformations.

5.2 Alternative benchmark model

This section aims to study whether the drift burst intensity also improves other benchmark models for the out-of-sample. The first benchmark model we choose is the popular HARQ model by Bollerslev et al. (2016),

$$\overline{RV}_{t+h|t} = \beta_0 + \left(\beta_d + \beta_d^Q \sqrt{RQ_{t-1,t}}\right) RV_{t-1,t} + \beta_w\, RV_{t-5,t} + \beta_m\, RV_{t-22,t} + \varepsilon_t, \quad (7)$$

where RQ_t is the realized quarticity estimator, with $RQ_t = \frac{M}{3}\sum_{i=1}^{M} r_{t_i}^2$. The second model we consider is a recent HARDQ developed by Laurent et al. (2022),

$$\overline{RV}_{t,t+h} = \beta_0 + \left(\beta_d + \beta_d^Q \sqrt{RQ_{t-1,t}}\right) RV_{t-1,t} + \beta_w\, RV_{t-5,t} + \beta_m\, RV_{t-22,t}$$
$$+ \left(\beta_d + \beta_d^Q \sqrt{RiceQ_{t-1,t}}\right) RAC^+_{t-1,t} + \beta_w\, RAC^+_{t-5,t} + \beta_m\, RAC^+_{t-22,t} + \varepsilon_t, \quad (7)$$

where RAC is the non-negative realized first-order Auto-Covariance (RAC) which measures the drift burst variation with $RAC_t = RAC_t \cdot I(RAC_t > 0)$ with $\sum_{i=K}^{M} r_{t_i} r_{t_{i-1}}$, and $RiceQ$ is a drift-robust quarticity estimator with $RiceQ_t = \frac{M}{6} \sum_{i=3}^{M} r_{t_i} r_{t_{i-K}}$. It is worthwhile to note that the HARDQ model already contains drift burst information, which, however, is all about the drift burst variation. It is interesting to see whether our intensity information of the drift bursts contributes a distinct value to this model.

To study the contribution of drift burst information, we formulate new models by simply adding the different drift burst intensities to the HARQ and the HARDQ model, in the same way, we have done for the HAR-DV model. For example, we create the HARQ-I model and the HARDQ-I model by including the drift burst intensity to the HARQ and the HARDQ model,

$$\overline{RV}_{t+h|t} = \beta_0 + \beta_\lambda\, \lambda_{t-1} + \left(\beta_d + \beta_d^Q \sqrt{RQ_{t-1,t}}\right) RV_{t-1,t} + \beta_w\, RV_{t-5,t} + \beta_m\, RV_{t-22,t} + \varepsilon_t, \quad (8)$$

$$\overline{RV}_{t+h|t} = \beta_0 + \beta_\lambda\, \lambda_{t-1} + \left(\beta_d + \beta_d^Q \sqrt{RQ_{t-1,t}}\right) RV_{t-1,t} + \beta_w\, RV_{t-5,t} + \beta_m\, RV_{t-22,t} + \left(\beta_d + \beta_d^Q \sqrt{RiceQ_{t-1,t}}\right) RAC^+_{t-1,t} + \beta_w\, RAC^+_{t-5,t} + \beta_m\, RAC^+_{t-22,t} + \varepsilon_t.$$

As a result, we will have overall 6 augmented models: HARQ-I, HARQ-I$^{\pm}$, HARQ-I^{-}, HARDQ-I, HARDQ-I$^{\pm}$, and HARDQ-I^{-} models. Table 8 compares these augmented models with their benchmarks, with positive DM statistics indicating that the augmented models perform better. We focus on reporting the results for the T&Z measure and SV forecast, as the main goal of this section is to check the robustness of the predictive power of drift burst intensity. But we confirm that for the HARQ and HARDQ models, the predictive superiority of the T&Z measure over the Hawkes intensity measure still holds, and both intensity measures provide a better forecast for longer horizons for RV forecast.

Table 8. Diebold–Mariano statistic for SV forecast for the augmented-HARQ and the augmented-HARDQ models.

			HARQ	HARQ	HARQ	HARDQ	HARDQ	HARDQ
			-I	-$I^{\pm}$	-I^{-}	-I	-$I^{\pm}$	-I^{-}
h=1	MSE	RW	0.30	-0.25	0.31	-0.71	-0.25	-0.79
		IW	1.72	1.19	**2.19**	1.58	0.97	**2.06**
	QLIKE	RW	**2.28**	0.63	1.76	**2.63**	0.61	**2.38**
		IW	**4.81**	-1.88	**7.49**	**3.76**	-1.92	**7.24**
h=5	MSE	RW	0.95	1.24	1.04	1.58	1.58	1.45
		IW	1.43	1.12	1.86	1.35	0.96	1.78
	QLIKE	RW	1.30	0.50	**2.20**	**2.11**	1.77	**2.76**
		IW	**4.19**	-1.67	**6.99**	**3.03**	-2.02	**6.63**
h=22	MSE	RW	1.73	1.79	1.68	1.60	1.76	1.63
		IW	1.72	1.13	**2.38**	1.70	1.06	**2.36**
	QLIKE	RW	0.98	1.31	**2.20**	1.43	1.66	1.93
		IW	**2.91**	-1.20	**3.81**	**2.83**	-1.08	**3.83**
h=44	MSE	RW	1.74	1.80	1.74	1.67	1.92	1.74
		IW	**2.28**	1.28	**2.82**	**2.28**	1.17	**2.81**
	QLIKE	RW	1.31	1.13	**2.55**	1.10	1.46	**2.58**
		IW	**3.28**	-0.01	**3.68**	**3.19**	-0.80	**3.67**
h=66	MSE	RW	**2.10**	1.72	1.85	1.90	1.66	1.73
		IW	**2.60**	0.83	**2.44**	**2.61**	0.74	**2.44**
	QLIKE	RW	0.64	-0.74	**2.98**	0.11	-0.62	**2.19**
		IW	**3.75**	**2.51**	**4.03**	**3.71**	**2.04**	**4.03**

Table 8 evaluates the out-of-sample of the three intensity models over the benchmark HARQ model, with forecast target SV. Again, the intensity-augmented HARQ model is able to be significantly superior to the HARQ model, and the pattern in the table is very similar to that in Table 6. For example, the HARQ-I^{-} model performs significantly better than the HARQ model across almost all of these different scenarios. And the HARQ-$I^{\pm}$ model generally provides significantly better forecasts than the HARQ model for all cases within $|T| > 4$ and $|T| > 4.5$. In addition, the out-of-sample results for RV and DV as targets are consistent with those in Table 4 and Table 5: the augmented HARQ model performs better for long horizons. For brevity, we report these results in the appendix.

5.3 The insanity filter

To alleviate these abnormal volatility forecasts, Bollerslev et al. (2016) apply the "insanity filter (IF)". The IF substitutes the forecast with the unconditional mean within the estimation window if the forecasts are outside the interval between the minimum and maximum forecasting target $\overline{V}$ of this window. Insanity filters for volatility forecasting have also been used by Patton and Sheppard (2015) and Bollerslev et al. (2018) Even so, the vast majority of studies examining the HAR model, and extensions thereof, do not employ an IF. Examples include Corsi et al. (2010), Corsi and Renò (2012), Andersen et al. (2021), and Caporin (2022). Unreported results show that all our results, including the robustness check, hold for employing an IF.

6 Conclusion

This paper exhibits the importance of drift burst information in predicting future volatility. The in-sample results reveal the positive effect of drift burst intensity and the leverage effect of the intensity of drift burst signs. In addition, including drift burst and drift burst signs intensity leads to a significant increase in the out-of-sample performance of the HAR-DV model by Andersen et al. (2021), the HARQ model by Bollerslev et al. (2016) and the HARDQ model by Laurent et al. (2022). The strong evidence for both in-sample and out-of-sample indicates the importance of drift burst information to volatility forecasting. Our findings are robust to both approaches by Hawkes (1971) and Tauchen and Zhou (2011), the use of insanity filter, and logarithm and square root volatility transformations.

Future research can be on multivariate intensity modelling. For example, one can use the multivariate Hawkes model to model the drift burst intensity across different markets, for investigating the interaction between these markets. In addition, the drift bursts may have implications for asset pricing. Since the drift bursts affect the price volatility, the drift bursts should also influence the beta, which builds on the covariance between the market prices and the asset prices.

7 Bibliography

AÏT-SAHALIA, Y., CACHO-DIAZ, J. & LAEVEN, R. J. 2015. Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics,* 117**,** 585-606.

AIT-SAHALIA, Y., MYKLAND, P. A. & ZHANG, L. 2005. How often to sample a continuous-time process in the presence of market microstructure noise. *The review of financial studies,* 18**,** 351-416.

ANDERSEN, T. G. & BOLLERSLEV, T. 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review***,** 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**,** 701-720.

ANDERSEN, T. G., LI, Y., TODOROV, V. & ZHOU, B. 2021. Volatility measurement with pockets of extreme return persistence. *Journal of Econometrics*.

BANDI, F. M. & RUSSELL, J. R. 2006. Separating microstructure noise from volatility. *Journal of Financial Economics,* 79**,** 655-692.

BANDI, F. M. & RUSSELL, J. R. 2008. Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies,* 75**,** 339-369.

BAUWENS, L. & HAUTSCH, N. 2009. Modelling financial high frequency data using point processes. *Handbook of financial time series.* Springer.

BELLIA, M., CHRISTENSEN, K., KOLOKOLOV, A., PELIZZON, L. & RENÒ, R. 2020. High-frequency trading during flash crashes: Walk of fame or hall of shame? : SAFE Working Paper No. 270.

BOLLERSLEV, T., HOOD, B., HUSS, J. & PEDERSEN, L. H. 2018. Risk everywhere: Modeling and managing volatility. *The Review of Financial Studies,* 31**,** 2729-2773.

BOLLERSLEV, T., PATTON, A. J. & QUAEDVLIEG, R. 2016. Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics,* 192**,** 1-18.

BOWSHER, C. G. 2007. Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics,* 141**,** 876-912.

CAPORIN, M. 2022. The Role of Jumps in Realized Volatility Modeling and Forecasting. *Journal of Financial Econometrics*.

CHRISTENSEN, K., OOMEN, R. & RENÒ, R. 2022. The drift burst hypothesis. *Journal of Econometrics***,** 461-497.

CLEMENTS, A. & LIAO, Y. 2017. Forecasting the variance of stock index returns using jumps and cojumps. *International Journal of Forecasting,* 33**,** 729-742.

CORSI, F., PIRINO, D. & RENO, R. 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics,* 159**,** 276-288.

CORSI, F. & RENÒ, R. 2012. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics,* 30**,** 368-380.

DIEBOLD, F. X. & MARIANO, R. S. 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics,* 13.

EDERINGTON, L. H. & GUAN, W. 2010. Longer-term time-series volatility forecasts. *Journal of Financial and Quantitative Analysis*, 45**,** 1055-1076.

FLORA, M. & RENÒ, R. 2020. V-shapes. *Available at SSRN 3554122*.

GETMANSKY, M., JAGANNATHAN, R., PELIZZON, L., SCHAUMBURG, E. & YUFEROVA, D. 2017. Stock Price Crashes: Role of Slow-Moving Capital. National Bureau of Economic Research.

HANSEN, P. R. & LUNDE, A. 2006. Realized variance and market microstructure noise. *Journal of Business & Economic Statistics,* 24**,** 127-161.

HAWKES, A. G. 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika,* 58**,** 83-90.

JAGANNATHAN, R., PELIZZON, L., SCHAUMBURG, E., GETMANSKY SHERMAN, M. & YUFEROVA, D. 2019. Liquidity provision: Normal times vs Crashes. *Available at SSRN 3079593*.

JAGANNATHAN, R., PELLIZON, L., SCHAUMBURG, E., SHERMAN, M. G. & YUFEROVA, D. 2021. Recovery from fast crashes: Role of mutual funds. *Journal of Financial Markets***,** 100646.

LARGE, J. 2007. Measuring the resiliency of an electronic limit order book. *Journal of Financial Markets,* 10**,** 1-25.

LAURENT, S., RENÒ, R. & SHI, S. 2022. *Realized Drift* [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4084647 [Accessed].

MA, F., LIAO, Y., ZHANG, Y. & CAO, Y. 2019. Harnessing jump component for crude oil volatility forecasting in the presence of extreme shocks. *Journal of Empirical Finance,* 52**,** 40-55.

MANCINI, C. 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics,* 36**,** 270-296.

NEWEY, W. & WEST, K. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica,* 55**,** 703-708.

OGATA, Y. 1981. On Lewis' simulation method for point processes. *IEEE transactions on information theory,* 27**,** 23-31.

PATTON, A. J. 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics,* 160**,** 246-256.

PATTON, A. J. & SHEPPARD, K. 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics,* 97**,** 683-697.

TAUCHEN, G. & ZHOU, H. 2011. Realized jumps on financial markets and predicting credit spreads. *Journal of Econometrics,* 160**,** 102-118.

TODOROV, V. 2019. Nonparametric spot volatility from options. *The Annals of Applied Probability,* 29**,** 3590-3636.