The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility

Behrendt, S., & Schmidt, A.

Journal of Banking & Finance, 2018.

Yuwei Bao

2021/01/16

Contents

1. Introduction: Background & Motivation

- Intraday volatility assessment and forecasting have gained importance for highly active investors.
- Not only has the speed of trading increased rapidly, but also the way investors can comment or share on social media platforms.
- Stock prices, reflecting the trading activities of both institutional and retail investors, might reflect retail investor trading activities influenced by sentiment.

1. Introduction: Literature

- Sentiment may influence the market:
  - Kyle (1985) & Black (1986): Define professional or institutional investors as rational informed investors, individual or retail investors as noise traders who have psychological biases.
  - De Long et al. (1990) & Shleifer and Vishny (1997): Propose rational investors could bet against sentiment driven noise traders to make a profit, with caution to the costs and risks.
  - Barber & Odean (2007): Retail investors trade excessively in attention-grabbing stocks.
  - Kumar and Lee (2006) & Barber et al. (2009): Retail investors trade in concert.

1. Introduction: Literature

- Twitter as a proxy for investor sentiment:
  - Bollen et al. (2011): Derive six social mood dimensions from Tweets, indicating that predictions of the DJIA are improved through some of them.
  - Sprenger et al. (2014a): Derive good and bad news from more than 400,000 Tweets related to the S&P 500, find these news have an impact on the market.
  - Sprenger et al. (2014b): Discover a relationship between stock related Twitter sentiment and returns, volume of Tweets and trading volume of the respective stock, disagreement and return volatility.

1. Introduction: Research Objective

- Assess the impact of Twitter investor sentiment and Twitter activity on return volatility.
- Test the performance of intraday volatility forecasts augmented with this additional information.

1. Introduction: Contribution

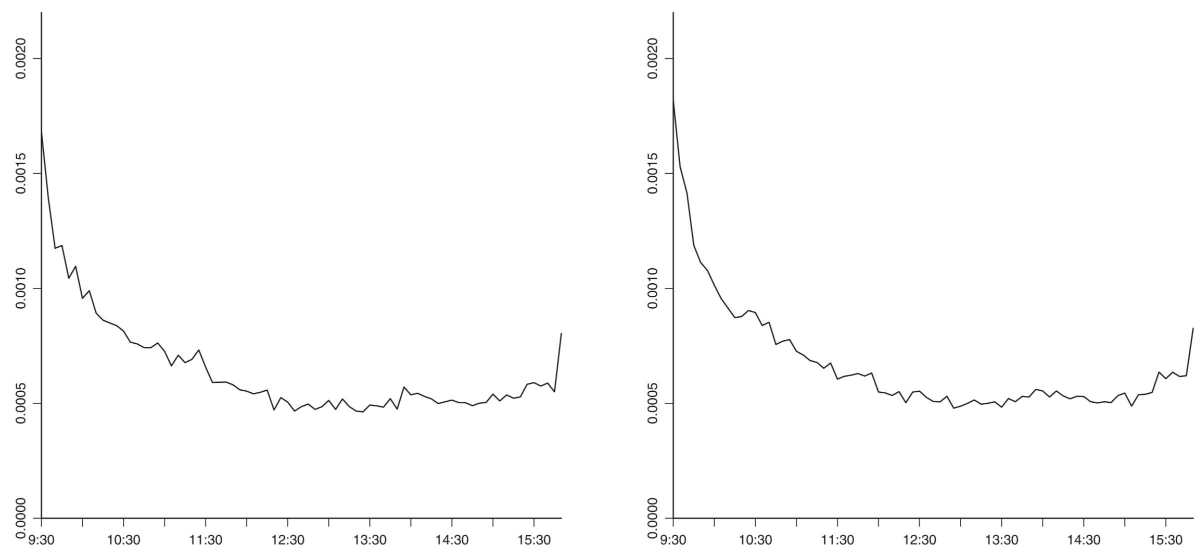
- Focus on intraday periodicity and individual-level stocks, conform to the rapid speed of trading and social media activities.
- Use absolute 5-minute returns as a measure for volatility, which have more persistent autocorrelation patterns, thus can eliminate the long-memory property in the further research.
- The HAR model to the intraday context has not been pursued by other authors so far.

2. Model Design: Data

- Intraday prices, Twitter sentiment and count:
  - ➢Source: Bloomberg.
  - ➢Sample: 30 most liquid blue-chip individual-level stocks, DJIA constituents.
  - ➢Period (t): 639 trading days, from June 18, 2015 to December 29, 2017.
  - ➢Frequency (n): 1-minute, then turn into 5-minute, so 77 intervals a day.
- Then take two companies: International Business Machines Corporation (IBM) and Walmart Inc. (WMT) for example.

2. Model Design: Variable

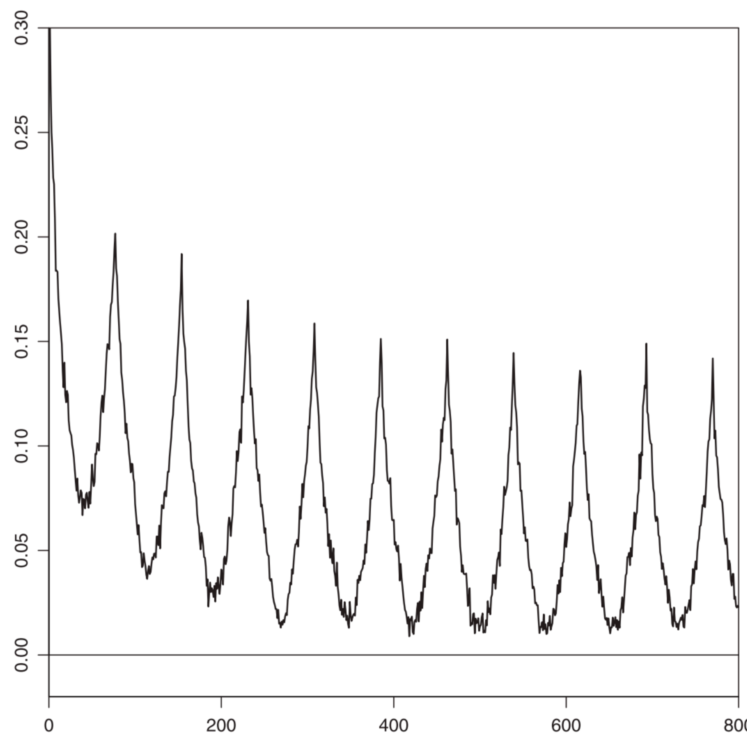
- Intraday Volatility
  - ➢Absolute 5-minute returns $(R_{t,n})$: Sum of five 1-minute returns, the log-price changes from one minute to the next.
    - Average of absolute 5-minute returns:



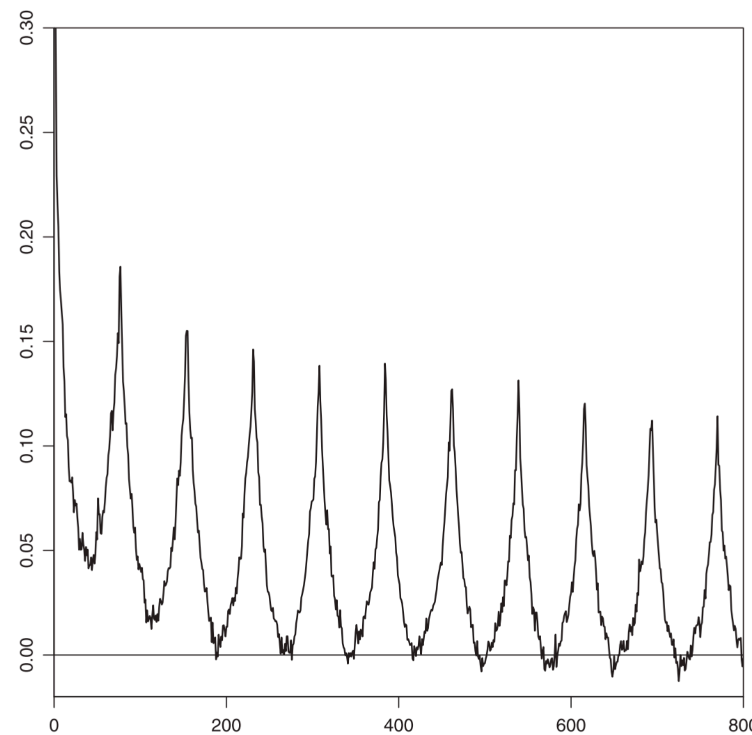
(a) IBM **(b)** WMT

2. Model Design: Variable

- Intraday Volatility
  - ➢Absolute 5-minute returns $(R_{t,n})$
    - ACFs of absolute 5-minute returns:



(a) IBM



(b) WMT

2. Model Design: Variable

- Intraday Volatility
  - Filtered 5-minute returns $(R^*_{t,n})$: Apply a two-step procedure based on an FFF (Fourier Flexible Form) estimation to purge the long-memory property.

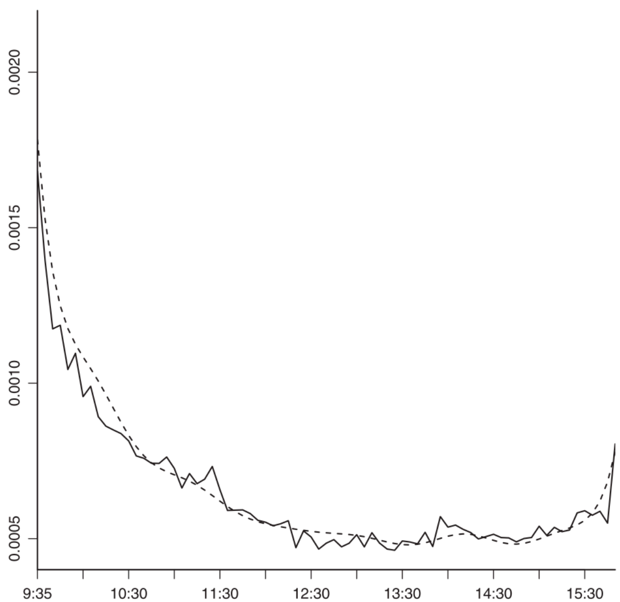
$$R_{t,n} - \mathbb{E}(R_{t,n}) = \varepsilon_{t,n} = s_{t,n}\sigma_{t,n}Z_{t,n}$$

$$R^*_{t,n} = {|R_{t,n}|}/{\hat{s}_{t,n}}$$

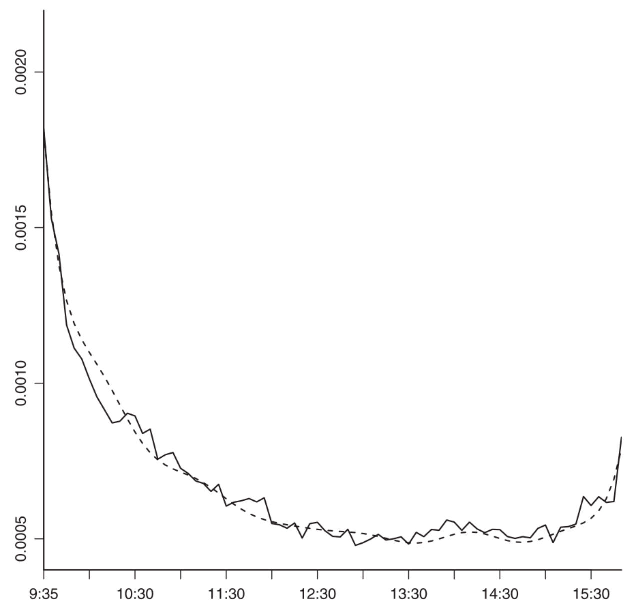
Where $s_{t,n}$ is the intraday periodic component, estimated in an FFF regression, $\sigma_{t,n}$ denotes a 5-minute volatility factor for trading day t and $Z_{t,n}$ is an $i.i.d.$ zero mean and unit variance innovation.

2. Model Design: Variable

- Intraday Volatility
  - ➢Filtered 5-minute returns $\left(R^*_{t,n}\right)$
    - Average of absolute 5-minute returns and intraday periodic volatility component $s_{t,n}$:



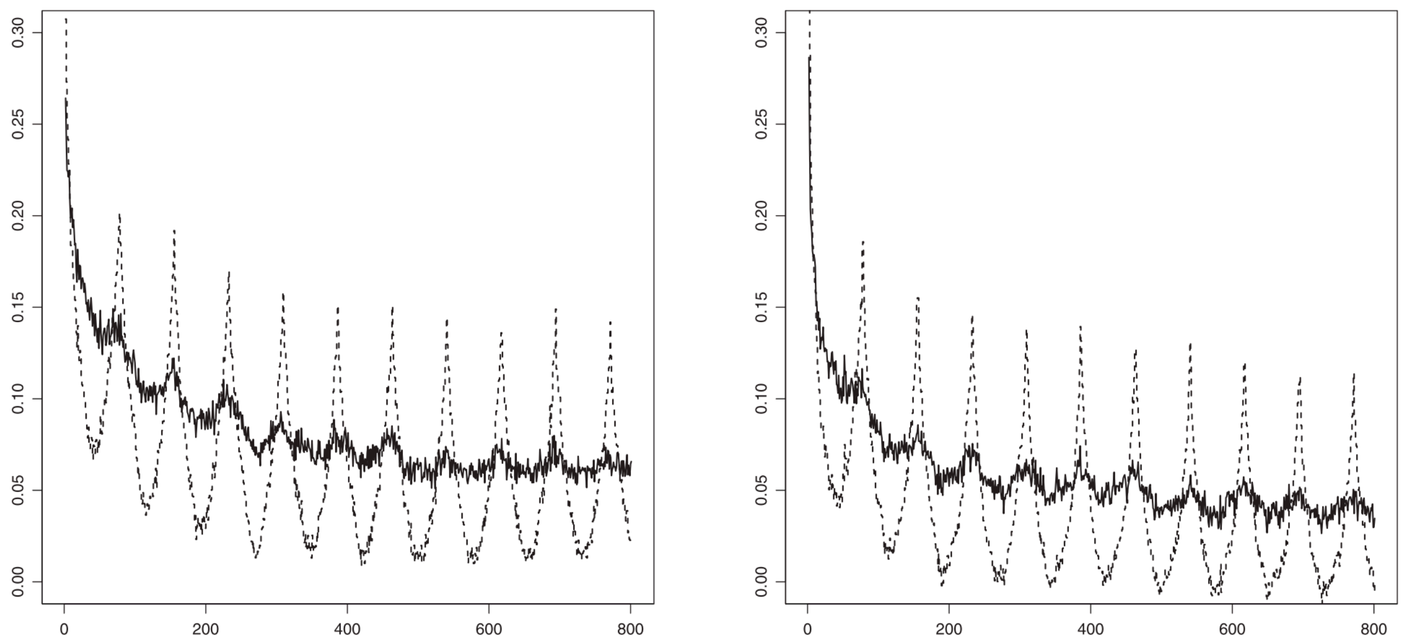
(a) IBM



(b) WMT

2. Model Design: Variable

- Intraday Volatility
  - ➢Filtered 5-minute returns $\left(R^{*}_{t,n}\right)$
    - ACFs of raw and filtered absolute 5-minute returns:



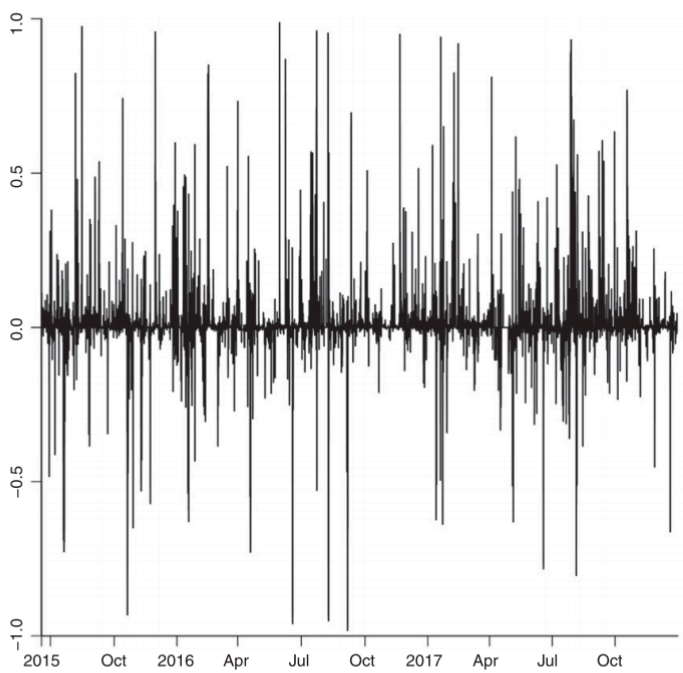
(a) IBM **(b)** WMT

2. Model Design: Variable

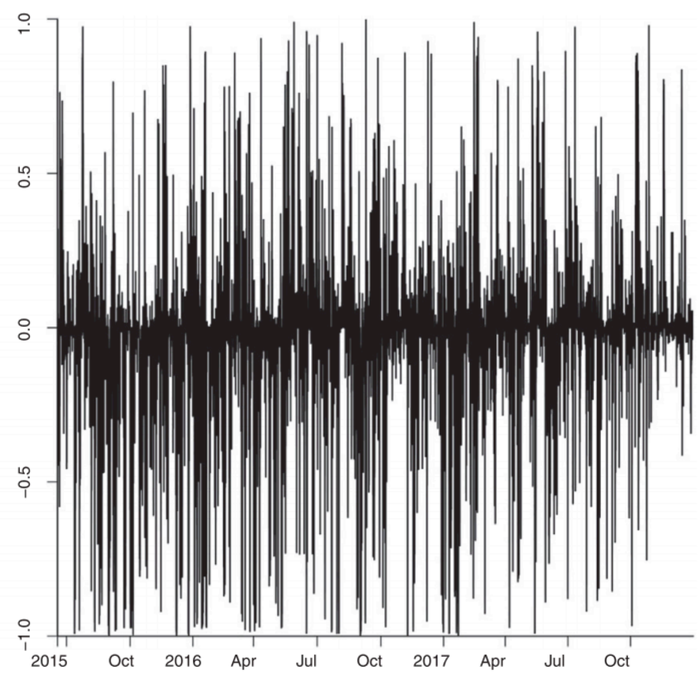
- Twitter Sentiment
  - Ranges continuously from 1 (positive investor sentiment) to −1 (negative investor sentiment).
  - The average sentiment over five minutes.
  - Only update the value if a change is observed. If no, fill in with the previously observed change.

2. Model Design: Variable

- Twitter Sentiment
  - Twitter sentiment time series:



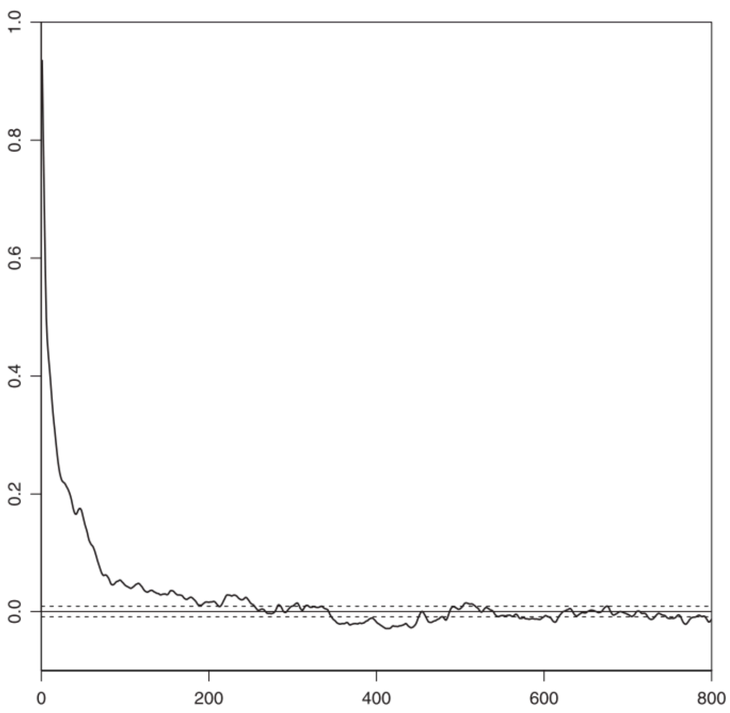
(a) TS: IBM



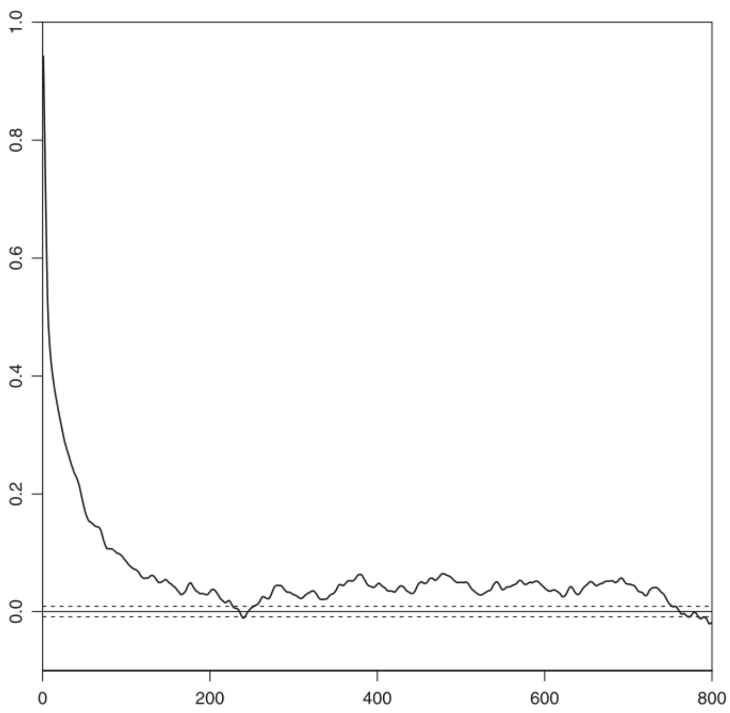
(b) TS: WMT

2. Model Design: Variable

- Twitter Sentiment
  - Twitter sentiment ACFs time series:



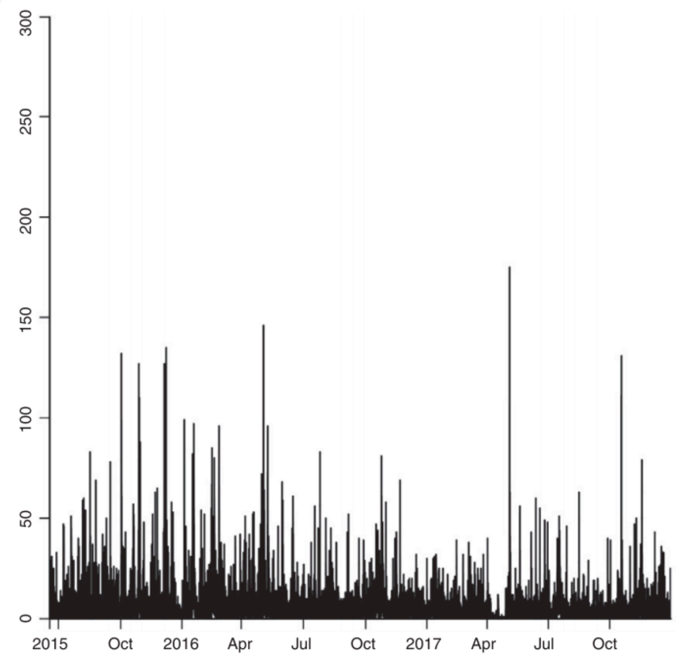
(c) ACF TS: IBM



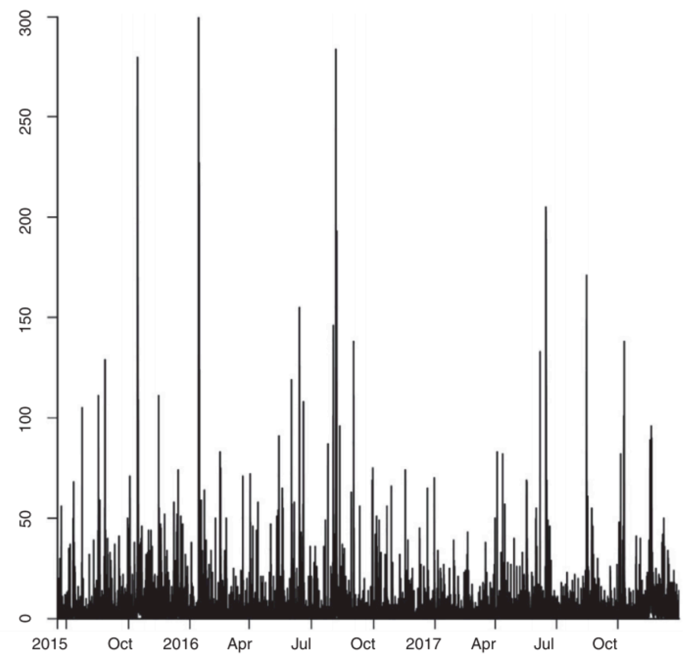
(d) ACF TS: WMT

2. Model Design: Variable

- Twitter Count
  - ➢The sum of count over five minutes.
  - ➢Twitter count time series:



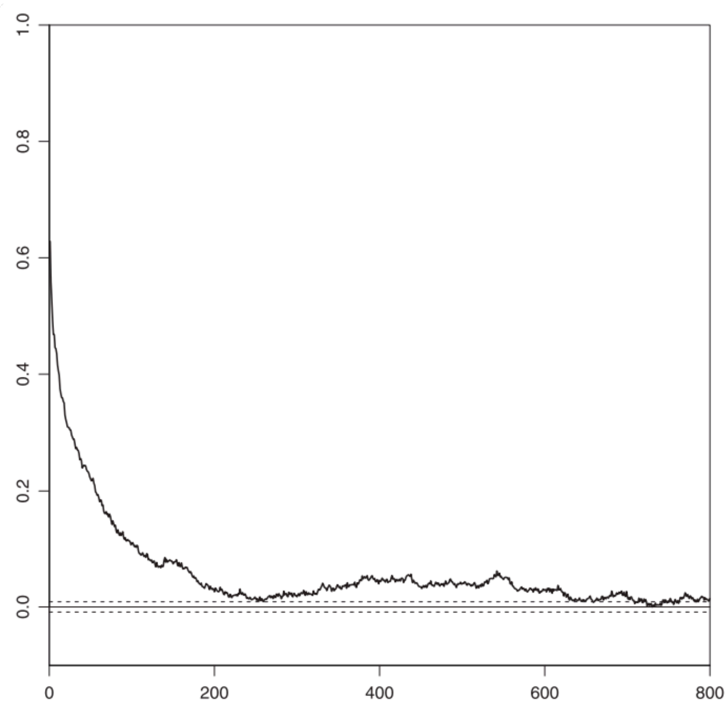
(a) TC: IBM



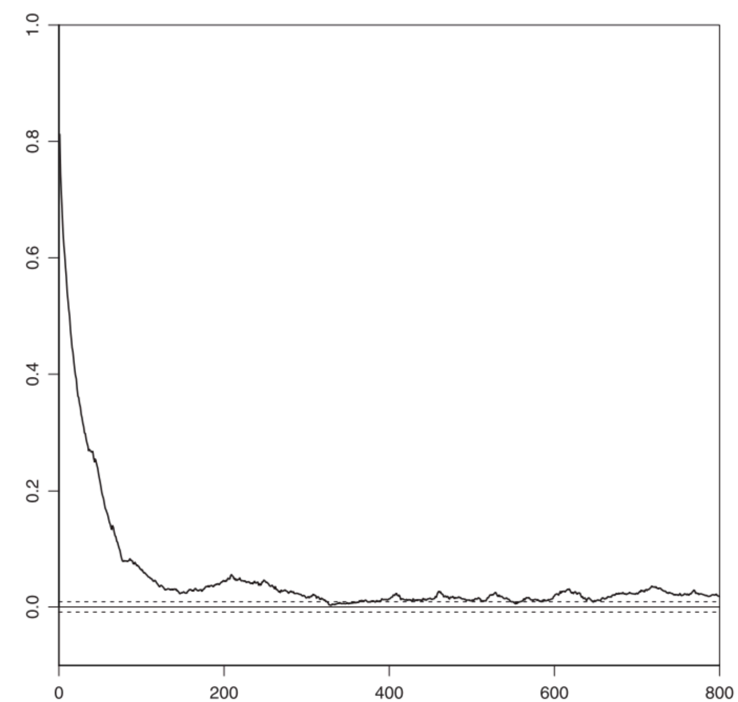
(b) TC: WMT

2. Model Design: Variable

- Twitter Count
  - ➢Twitter count ACFs time series:



(c) ACF TC: IBM



(d) ACF TC: WMT

2. Model Design: Method

- Bivariate VAR model

$$\begin{bmatrix} 1 & 0 \\ b_{21}^{0} & 1 \end{bmatrix} \begin{bmatrix} R_{t,n}^{*} \\ twit_{t,n} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \sum_{j=1}^{p} \begin{bmatrix} b_{11}^{j} & b_{12}^{j} \\ b_{21}^{j} & b_{22}^{j} \end{bmatrix} \begin{bmatrix} R_{t,n-j}^{*} \\ twit_{t,n-j} \end{bmatrix} + \begin{bmatrix} u_{1t,n} \\ u_{2t,n} \end{bmatrix}$$

Where $twit_{t,n}$ denotes Twitter sentiment or count.

- The lag order p is the average lag length suggested by the Schwarz information criterion (SC) across all 30 DJIA constituents, which leads to $p = 17$ for the filtered absolute 5-minute returns $R_{t,n}^{*}$ and Twitter sentiment, $p = 18$ for Twitter count.

2. Model Design: Method

- Granger causality test
- The contemporaneous correlation
- Forecast error variance decomposition (FEVD)
- Impulse response analysis

2. Model Design: Method

- The panel HAR model

$$
\begin{aligned}
&R^*_{t,n} \\
&= c + \beta_1 R^*_{t,n-1} + \beta_{12} R^{*12}_{t,n-1} + \beta_{24} R^{*24}_{t,n-1} \\
&+ \delta_1 twit_{t,n-1} + \delta_{12} twit^{12}_{t,n-1} + \delta_{24} twit^{24}_{t,n-1} \\
&+ \gamma_1 \operatorname{sgn}\left(\bar{R}_{t,n-1}\right) + u_{t,n}
\end{aligned}
$$

Where $\operatorname{sgn}\left(\bar{R}_{t,n-1}\right)$ denotes the sign of the average return in the previous 5-minute interval, $R^{*12}_{t,n-1} = \frac{1}{12}\sum_{j=1}^{12} R^*_{t,n-j}$, $R^{*24}_{t,n-1} = \frac{1}{24}\sum_{j=1} R^*_{t,n-j}$ are lagged averages for one and two hours of the filtered returns, $twit^{12}_{t,n-1}$ and $twit^{24}_{t,n-1}$ the same.

2. Model Design: Method

- The panel HAR model
  - Investors can react swiftly to changes, further lags are omitted.
  - The sign variable is for asymmetric effect of returns, as negative returns have a larger effect on volatility than positive returns.
  - 90% of the data to fit the panel HAR model, 10% to forecast.
  - The root mean squared error (RMSE) is used to measure the forecast performance.

3. Empirical Result

- Bivariate VAR model
  - Only estimates of the respective Twitter variable significant at the 10% level are displayed.
  - Coefficient estimates are multiplied by 10^3.

	Panel A: VAR estimation results $\times 10^3$			
	Twitter sentiment		Twitter count	
	IBM	WMT	IBM	WMT
$twit_{t,n-1}$			0.0024 (0.0001)	0.0012 (0.0192)
$twit_{t,n-2}$	0.2504 (0.0938)			
$twit_{t,n-4}$				-0.0017 (0.0043)
$twit_{t,n-5}$				0.0019 (0.0013)
$twit_{t,n-9}$			0.0011 (0.0802)	
$twit_{t,n-10}$	0.2807 (0.0696)			
$twit_{t,n-12}$		-0.2104 (0.0041)		
$twit_{t,n-13}$		0.1922 (0.0088)		
$twit_{t,n-14}$				-0.0017 (0.0039)
$twit_{t,n-18}$			0.0019 (0.0021)	0.0018 (0.0006)

3. Empirical Result

- Granger causality test

Panel B: Granger causality test

	Twitter sentiment		Twitter count	
H_0	IBM	WMT	IBM	WMT
$R^* \nrightarrow twit$	1.2607	0.8019	**5.9968**	**40.4163**
	(0.2076)	(0.6929)	(0.0000)	(0.0000)
$twit \nrightarrow R^*$	1.0957	**1.7018**	**3.2663**	**2.8456**
	(0.3503)	(0.0352)	(0.0000)	(0.0000)

3. Empirical Result

- The contemporaneous correlation

Panel A: Twitter sentiment

$$\begin{array}{c|c}
\text{IBM} & \text{WMT} \\
\hline
\begin{array}{cc} & \begin{array}{cc} R^* & twit \end{array} \\ \begin{array}{c} R^* \\ twit \end{array} & \begin{pmatrix} 1 & -0.0079 \\ -0.0079 & 1 \end{pmatrix} \end{array} &
\begin{array}{c} \begin{array}{cc} R^* & twit \end{array} \\ \begin{pmatrix} 1 & -0.0032 \\ -0.0032 & 1 \end{pmatrix} \end{array}
\end{array}$$

Panel B: Twitter count

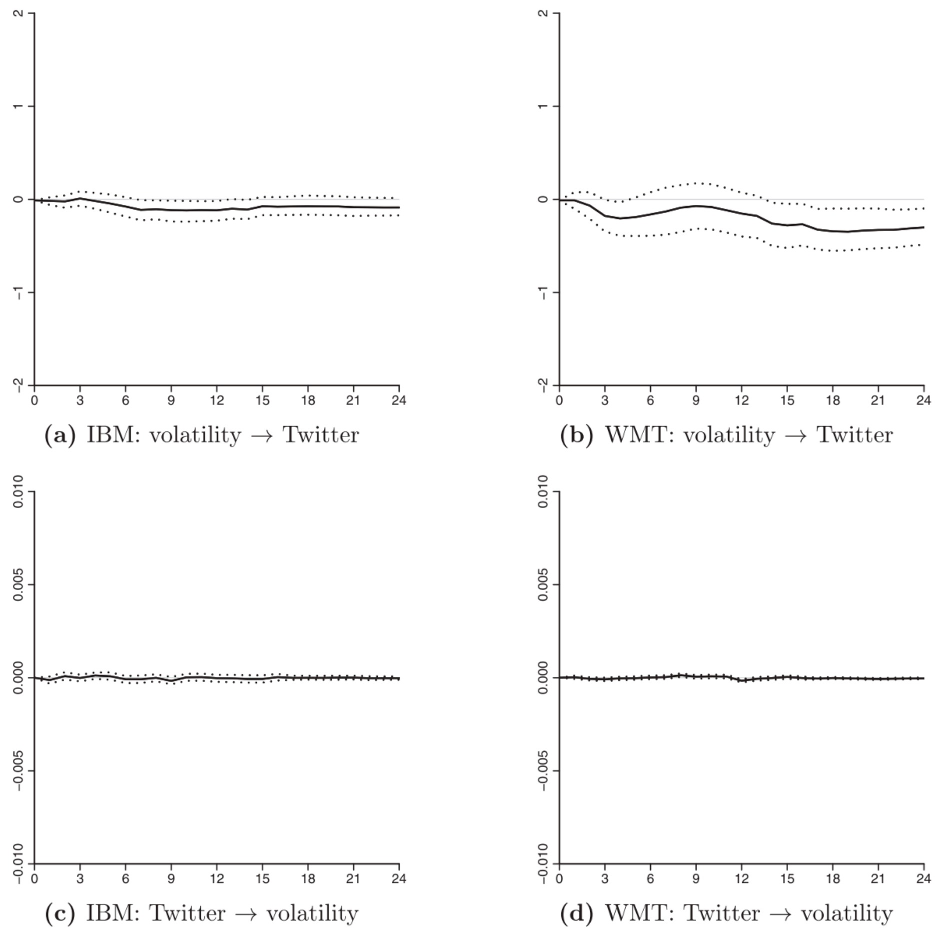
$$\begin{array}{c|c}
\text{IBM} & \text{WMT} \\
\hline
\begin{array}{cc} & \begin{array}{cc} R^* & twit \end{array} \\ \begin{array}{c} R^* \\ twit \end{array} & \begin{pmatrix} 1 & 0.0068 \\ 0.0068 & 1 \end{pmatrix} \end{array} &
\begin{array}{c} \begin{array}{cc} R^* & twit \end{array} \\ \begin{pmatrix} 1 & 0.0191 \\ 0.0191 & 1 \end{pmatrix} \end{array}
\end{array}$$

3. Empirical Result

- Forecast error variance decomposition (FEVD)
  - No relevant contribution of either Twitter variable for all DJIA constituents (less than 2% of the forecast error variance).

3. Empirical Result

- Impulse response analysis (Twitter sentiment)

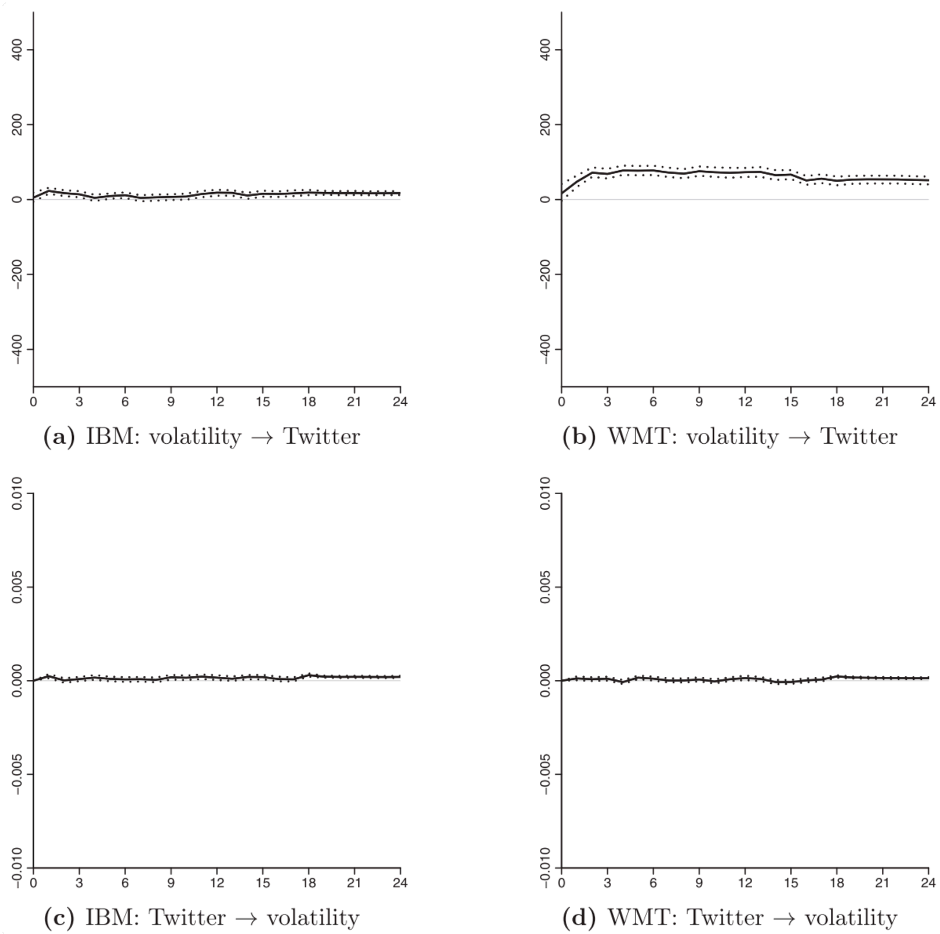


(a) IBM: volatility → Twitter (b) WMT: volatility → Twitter

(c) IBM: Twitter → volatility (d) WMT: Twitter → volatility

3. Empirical Result

- Impulse response analysis (Twitter count)



(a) IBM: volatility → Twitter
(b) WMT: volatility → Twitter
(c) IBM: Twitter → volatility
(d) WMT: Twitter → volatility

3. Empirical Result

- The panel HAR model
  - P-values are given in parentheses.

	Panel A: Twitter sentiment		Panel B: Twitter count	
	IBM	WMT	IBM	WMT
$R^{*}_{t,n-1}$	0.0441**	0.0540***	0.0440**	0.0536***
	(0.0227)	(0.0000)	(0.0226)	(0.0000)
$R^{*12}_{t,n-1}$	0.1496***	0.0971**	0.1496***	0.1013**
	(0.0027)	(0.0282)	(0.0029)	(0.0203)
$R^{*24}_{t,n-1}$	-0.0805**	-0.1228***	-0.0805**	-0.1230***
	(0.0348)	(0.0000)	(0.0371)	(0.0000)
$sgn(\bar{R}_{t,n-1})$	0.0000	0.0000	0.0000	0.0000
	(0.8952)	(0.3343)	(0.8674)	(0.3169)
$twit_{t,n-1}$	0.0000	0.0000	0.0000***	0.0000*
	(0.7516)	(0.7270)	(0.0065)	(0.0714)
$twit^{12}_{t,n-1}$	-0.0001	0.0001**	0.0000	0.0000
	(0.5852)	(0.0487)	(0.4354)	(0.3725)
$twit^{24}_{t,n-1}$	0.0001**	0.0000	0.0000	0.0000
	(0.0288)	(0.9384)	(0.9543)	(0.1507)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3. Empirical Result

- The panel HAR model
  - Forecast evaluation by RMSE ($\times 10^4$).

	IBM	WMT
HAR	6.8934	9.0461
HAR + *TS*	6.8798	9.0309
HAR + *TC*	6.8706	8.8692

3. Empirical Result

- Result for all 30 stocks
  - Lower trading volume more often show a statistically significant influence of Twitter sentiment and count than those higher.
  - HAR and VAR model are not entirely robust across all.

					Sentiment		Count	
Ticker	Volume	VAR	HAR1	HAR2	$R^* \nrightarrow T$	$T \nrightarrow R^*$	$R^* \nrightarrow T$	$T \nrightarrow R^*$
GE	453.81	×			**2.1684**	0.6976	**5.9543**	**2.8337**
AAPL	367.75			×	**1.4721**	0.6154	**5.6040**	**6.0759**
MSFT	285.29				0.9605	**1.4860**	1.2739	0.8573
PFE	264.22	×			1.0670	0.1066	**3.8680**	1.1818
INTC	247.81				1.3820	0.3625	**2.9899**	**1.7498**
CSCO	233.83			×	1.0993	**1.5003**	0.7855	**2.8876**
JPM	154.75				**1.9473**	0.9993	**1.8870**	1.0505
VZ	148.85				0.7375	0.6828	0.4858	1.1721
KO	128.42				1.2546	1.3345	0.9386	**1.7326**
XOM	123.74	×	×		1.2731	0.7470	1.1769	**1.7594**
MRK	102.41				0.6854	1.1696	1.2390	0.4046
PG	96.07				0.6206	0.9298	**33.2960**	**1.5941**
NKE	94.80	×			0.5325	1.2436	**3.4281**	**2.7298**
WMT	94.38	×	×		0.8019	**1.7018**	**40.4163**	**2.8456**
V	85.88		×	×	0.8170	0.5741	1.3083	1.0021
DIS	82.61	×			**1.7140**	0.6003	**25.7273**	**9.8504**
DD	76.55	×			1.2967	1.0419	**9.6421**	**4.7287**
CVX	76.36	×			1.1816	1.1762	**3.0121**	**2.4587**
JNJ	69.96	×			**2.6715**	**2.1258**	**3.4540**	**2.6726**
CAT	52.74				**2.9053**	1.0739	**10.3373**	**4.5908**
HD	48.94	×			0.9799	**2.1205**	**9.3463**	**3.3585**
AXP	48.54	×			0.7538	**2.5288**	**15.9147**	**6.1008**
MCD	48.41	×		×	1.0901	1.0417	**1.9112**	**1.5883**
IBM	42.22	×	×		1.2607	1.0957	**5.9968**	**3.2663**
UTX	40.30	×			0.6168	**1.8660**	**20.1767**	**18.2685**
BA	37.95	×			0.8354	0.6984	**8.1744**	**2.1790**
UNH	35.05	×			0.8481	1.3064	**4.9244**	**2.4981**
GS	33.80	×			1.0125	**1.6159**	**3.5307**	**2.1091**
MMM	20.98				**1.5369**	0.9916	0.9079	**1.6085**
TRV	16.64				0.5003	1.1115	0.8267	0.5104

4. Conclusion

- There are indeed statistically significant feedback effects in a bivariate VAR framework, but coefficients are of small absolute magnitude and do not have a significant economic impact.
- Incorporating exogenous information from Twitter into intraday prediction for return volatility doesn’t have a significant impact.
- The performance of liquid blue-chip stocks such as the DJIA constituents should be linked to information related to fundamentals.

4. Conclusion

- The intraday frequencies (5-minute) are too high for investors to react appropriately, lower frequencies can be investigated in the future.
- Future research should base on less liquid stocks, in order to test the validity and robustness of the findings presented in this paper.

Comparison

	Deng's	Behrendt & Schmidt's
Data	StockTwits & Traditional news media	Twitter
Variable	Sentiment (Both negative and positive)	Sentiment & Count
Sample	44 most frequently mentioned stocks	30 most liquid DJIA constituents
Level	Individual & Index level	Individual level
Frequency	1 hour & 1 day	5 mins
Method	Both use VAR, Granger causality test, FEVD and Impulse response analysis.	
Conclusion	1. Sentiment on stock return is both statistically and economically significant at the hour level. 2. Microblog sentiment is also largely driven by stock market. 3. Stock return has a stronger influence on negative sentiment.	High-frequency Twitter information is not particularly useful for intraday volatility assessment and forecasting when considering individual-level stocks.