UCL Journal of Economics

Article

The Influence of Market Dynamics on Retail Investor Attention

James Symons-Hicks[1]

[1] Student (BSc Economics, 2020-23), Dept. of Economics, UCL, UK; james.symons-hicks.20@ucl.ac.uk
Submission Date: 9th April 2023; Acceptance Date: 29th June 2023; Publication Date: 2nd October 2023

How to cite

Symons-Hicks, J. (2023). The Influence of Market Dynamics on Retail Investor Attention. UCL Journal of Economics, vol. 2, pp. 36-64. DOI: https://doi.org/10.14324/111.444.2755-0877.1611

Peer review

This article has been peer-reviewed through the journal's standard double-blind peer review, where both the reviewers and authors are anonymised during review

---

Abstract

Retail investors have become increasingly active in global markets over the past several years. However, the factors that drive retail investors to focus on particular stocks are unclear. Using a sample of UK FTSE 100 stocks, this paper analyses whether stock volatility, liquidity, returns, and trading volume have the power to attract the attention of UK retail investors, measured using the Google Search Volume Index. Overall, this paper finds positive relationships between three of the dynamics (volatility, returns, and daily trading volume) and increased retail investor attention. Greater stock illiquidity also coincides with an increase in the Google Search Volume Index, although this may be due to liquidity-impacting events. When conditioning on stocks by quartiles of market capitalisation, I find that the effects of returns and trading volume are greater in magnitude for the top 25% of stocks.

Keywords: Retail Investors, Market Dynamics, Financial Markets, FTSE 100, Attention, Investor Behaviour

1. Introduction

Retail investors have caught the attention of global media over the past several years, particularly for their potential to cause extreme market events. There are famous cases of asset prices soaring as a result of retail attention; GameStop, AMC, and Bitcoin are just a few of many examples. This paper seeks to identify which market dynamics capture the attention of retail investors. Attention is a scarce resource, which limits the amount of information that humans are able to process at any one time (Kahneman, 1973). Therefore, it is of interest to consider which market dynamics are sufficiently important to capture retail investors' attention; the dynamics considered in this paper are volatility, liquidity, returns, and daily trading volume.

Much of the literature focuses on the impact that retail investors have on markets: how they impact liquidity and volatility (Barrot et al., 2016; Foucault et al., 2011; Kelley and Tetlock, 2013), how their trading makes returns predictable (Barber et al., 2009), and how they impact price stability (Baig et al., 2022). However, few studies have investigated what attracts them to trading in the first place. Some studies which have looked at this topic include Hsieh et al. (2020) and Seasholes and Wu (2007); these study the impact of upper price limit events on investor attention, finding that such events generate an "attention-grabbing effect". Kaniel et al. (2008) find that volume shocks bring media attention to stocks, which, in turn, captures investors' attention.

Overall, there is a lack of literature that looks more generally, rather than event-specific studies, at the impacts of market dynamics on attention. Furthermore, the number of retail investors in the UK has increased over the past decade as a result of, amongst other factors, increased access to commission-free trading (Statista, 2021). This can be seen by the monthly number of active users on the major trading platforms (Figure 1).[1] As such, we need to consider the modern effects, given that the characteristics of these investors may change over time. There is very little UK-focused literature on this topic, and, since these investors can significantly impact markets, greater effort should be taken to understand their behaviour.

I use a sample of 87 stocks from the UK's FTSE 100 index, and the Google Search Volume Index as a measure of retail investor attention. The empirical results suggest that changes in volatility, returns, and, to some extent, daily trading volume lead to increased retail investor attention. I find a positive relationship between illiquidity and attention, although this is likely due to exogenous, liquidity-impacting shocks that simultaneously capture investors' attention. When conditioning on stocks by quartiles of market capitalisation, I find that the effects of returns and trading volume are greater in magnitude for the top 25% of stocks. This suggests that for larger stocks, which are likely more widely publicised in the media, a shock to market dynamics has a greater impact on attention. For smaller stocks, even with larger shocks to dynamics, the resulting impact on attention is smaller.

The remainder of the paper is set out as follows: Section 2 analyses the current literature; Section 3 discusses the data sets used in the analysis; Section 4 explains the empirical methodology used;

[1] Between January 2017 and May 2021 (the peak), we see a 395% increase in the number of active users each month. Excluding the COVID-19 pandemic, we still observe a 53% increase between January 2017 and February 2020.

https://doi.org/10.14324/111.444.2755-0877.1611

Section 5 presents the findings from the data; Section 6 discusses the implications of these findings; Section 7 concludes.

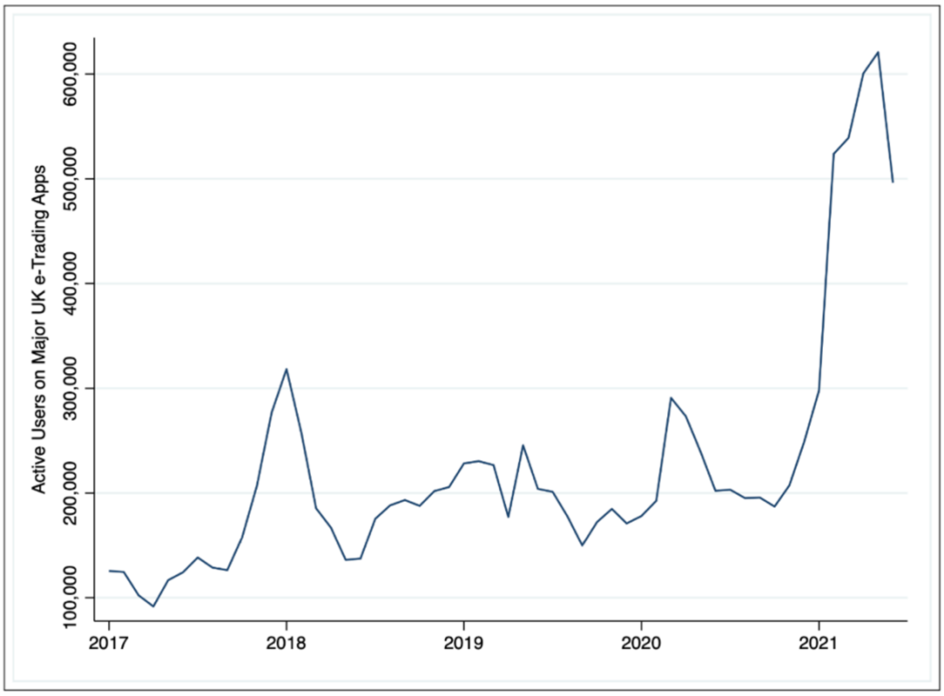

Figure 1: Number of Monthly Active Users on the Major e-Trading Apps (Statista, 2022)

2. Literature Review

A substantial volume of literature has analysed the impacts of retail investors on markets (e.g., Barber et al. (2022), Barber and Odean (2008), Cheng et al. (2021), and Da et al. (2011)); these papers commonly show that retail investors contribute to both market liquidity and volatility. However, very few papers look at this from the opposite direction, that is, considering the effects of preceding dynamics on attention.

Several papers have considered how retail investors' trades respond to attention-grabbing stocks, that is, those that "have experienced extreme returns, inflated trading volume, or been the subject of extensive headlines" (Gavish et al., 2021). Barber and Odean (2008) find that retail investors are net buyers of attention-grabbing stocks; Gavish et al. (2021) find that the magnitude of this effect depends on their level of sophistication. Barber and Odean use three measures to identify stocks that are likely to be attention-grabbing: news, unusual trading volume, and extreme returns. They use these measures as a proxy for whether investors were paying attention to a stock.

Seasholes and Wu (2007) study upper price limit events on the Shanghai Stock Exchange, finding that such events capture individual investors' attention because high returns and high trading volume generate news. Similarly, Hsieh et al. (2020) show that "reaching the price limit generates an attention-grabbing effect" for retail investors. Such events "induce individual investors to buy stocks they have

https://doi.org/10.14324/111.444.2755-0877.1611

not previously owned" (Seasholes and Wu, 2007). This is consistent with the findings of Kaniel et al. (2008): "it is reasonable to assume that individual investors do not follow all the stocks all the time but may be attracted to a certain stock after a volume shock brings media attention to it." Similarly, Engelberg and Parsons (2011) find that individual investors are more likely to trade a stock if it has appeared in the local newspaper. These results suggest that media attention is in response to market dynamics, thus acting as a 'stepping stone' between market dynamics and the resulting retail investor attention. The extent to which the media focuses on these dynamics is a factor that determines the strength of transmission between market dynamics and attention (Engelberg and Parsons, 2011).[2]

This mechanism will not, however, be the primary focus of the paper.

Welch (2022) finds that retail investors are collectively attracted to volatility and to "stocks with high past share volume and dollar-trading volume". Welch hypothesises that they may enjoy the risk associated with holding these stocks. During return- and turnover-based attention-grabbing events, "an increase in retail trading...is associated with an increase in idiosyncratic volatility" (Brandt et al., 2010). Similar to this paper, but without a focus on equities, Urquhart (2018) uses the Google Search Volume Index (SVI) to analyse the effects of shocks on volatility, volume, and returns on Bitcoin attention, finding that volatility and volume significantly impact attention the following day.

A factor not relating directly to market dynamics but still potentially impacting attention is local peer performance.[3] Kaustia and Knüpfer (2012) find a positive effect of local peer performance on participation the following month, but only for positive returns. This supports the idea that individuals extrapolate from peer outcomes and that stories of high returns encourage other investors. Kaustia and Knüpfer suggest that these results "are likely to be stronger in populations with more wide-spread stock market participation", making the UK a good market to consider given the low cost for retail investors to access markets.[4]

Recent literature has identified the Google SVI as an accurate indicator of retail investor attention (Da et al., 2011; Ding and Hou, 2015; Smales, 2021). Whilst many studies analyse the impact of attention-grabbing events on retail investors, they use noisy proxies for what constitutes an attention-grabbing event; Google SVI is a cleaner proxy and a trusted measure for attention. Whereas institutional investors use services such as Bloomberg, the vast majority of retail investors instead use the internet, namely Google, as their source of information. Da et al. (2011) argue that the Google SVI is a direct and accurate measure of retail investor attention for several reasons. Firstly, Google remains the most popular search engine; moreover, "search is a revealed attention measure": if someone is searching for something, they are paying attention to it.

[2] Engelberg and Parsons (2011) find that, when reacting to "the same set of information events (earnings releases of S&P 500 Index firms),... the presence or absence of local media coverage is strongly related to the probability and magnitude of local trading".

[3] Peers are here defined as individuals to which the person of interest is connected, whether this be friends, family, co-workers, or as is common for many online traders, social media connections.

[4] Many platforms have and continue to emerge in the UK over the last decade that offer commission-free trading to retail investors. Previously, high, fixed transaction costs per trade meant that trading was only economically viable for those retail investors who were willing and able to trade larger amounts. This is no longer the case; for example, Trading212 offers trading from as little at £1.

https://doi.org/10.14324/111.444.2755-0877.1611

3. Data

The data used in this analysis is split into two categories: UK financial market data and investor attention data.

3.1 Refinitiv FTSE 100 Data

The FTSE 100 pricing data was obtained from the Refinitiv Data platform, and ranges between 01 January 2017 and 01 July 2022. The information contains opening and closing prices, daily high and low values, daily market capitalisation of each stock, daily trading volume, daily turnover (total trade value for a given market day), and closing bid and ask prices. This data is used to create the measures of liquidity, volatility, and daily returns, which are, in turn, used to form the analysis.

There are five stocks for which the data is truncated due to different listing dates on the London Stock Exchange.

The ticker symbols for these stocks are: *AAF, AVST, EDV, MNG,* and *PSH*. These have been removed from the analysis. For a full list of the stocks in the data set, see Table A1 in the Appendix.

Table 1 shows the summary statistics for the data retrieved from the Refinitiv Data platform. Table 2 shows the summary statistics for the spreads and returns generated using this data.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 1: Summary Statistics for FTSE 100 Constituents

	Mean	**Std. Dev.**	**Min.**	**Max.**
Open Price	50.94	298.64	0.24	4267.65
Close Price	52.94	298.54	0.24	4263.64
High Price	53.61	302.70	0.25	4299.74
Low Price	52.27	294.44	0.24	4195.46
Closing Bid	52.93	298.44	0.24	4259.63
Closing Ask	52.96	298.66	0.24	4263.64
Daily Volume	8.98M	25.70M	11288	1.41B
Market Cap.	21.60B	29.90B	104.00M	233.00B
Turnover	68.10M	257.0M	63882.70	27.70B

All values rounded to two decimal places.

Observations: 120, 835

Note: M refers to million - i.e., 1M = 1×10^6. B refers to billion - i.e., 1B = 1×10^9.

Table 2: Summary Statistics for FTSE 100 Returns & Spreads

	Mean	**Std. Dev.**	**Min.**	**Max.**
Returns				
1-Day Return	0.0003	0.0206	-0.5747	0.5668
5-Day Return	0.0013	0.0461	-0.8134	0.9648
Spreads				
Absolute Spread	0.0373	0.3106	0.00*	24.0657
Relative Spread	0.0006	0.0001	0.00*	0.2026

All values rounded to four decimal places.

*There are only three instances for which the closing spread is 0. When excluding these, the minimum absolute and relative spreads are 0.00005 and 0.0001, respectively. The mean, standard deviation, and maximum values remain unchanged to four decimal places.

https://doi.org/10.14324/111.444.2755-0877.1611

3.2 Google Search Volume Index

As a measure of retail investor attention, I use the Google SVI for the ticker symbol, following Da et al. (2011).[5] The data provides a weekly index for searches of a given term, with values ranging between 0 and 100. Da et al. choose to use the ticker symbol rather than the company name as it is "less ambiguous", stating that if an investor is searching for a particular stock symbol, they are likely doing so because they are interested in the financial information of the company.

In some cases, the search term was combined with the word 'share' to avoid ambiguity with other similar searches; this judgement was made based on the 'related topics' component from Google Trends.[6] To ensure robustness, I conduct dummy variable regressions to test for statistical significance of the effects of noisy searches. These results are reported in Appendix B.1. Overall, I find that the results are not affected significantly when using only the subset of 'clean' ticker symbols.

There are eight stocks for which the Google Trends data is excluded. This is due to the fact that the ticker symbol only is too ambiguous, but there is missing data when using the ticker symbol appended with 'share'. These stocks are: DPH, FLTR, HIK, ICP, LAND, RS1, SDR, and SKG. Thus, when combining the financial pricing data (five stocks removed) and Google SVI data (eight stocks removed), the analysis is conducted with 87 stocks. Although the exclusion of these stocks is not ideal, this is unlikely to significantly affect the empirical results, given that the models are run at the stock-level and the SVI data is an index, not an absolute value, meaning that the analysis considers *changes* in this index.

One potential issue caused by the missing data is for the quantile regression models (discussed in Section 4.4); several of these eight stocks have a lower than average market capitalisation, implying that some of the models conditioning on the bottom 25% of stocks by market capitalisation may be biased to some extent. This is not overly problematic, however, as the number of excluded stocks due to missing data is low.

[5] This data is obtained from Google Trends, which is available at https://trends.google.co.uk/.
[6] For a further explanation, I refer to Appendix A.2.

https://doi.org/10.14324/111.444.2755-0877.1611

4. Methodology

This section will discuss the empirical methods used in the analysis. The results are reported in Section 5 and are discussed in Section 6.

4.1 Measures of Explanatory Variables

Firstly, I define the empirical proxies used for the measures of liquidity and volatility. In the descriptions below, X_d and X_t refer to observations of variable X at the daily and weekly levels, respectively.

4.1.1 Volatility Measures

Financial volatility refers to the fluctuations in the returns of an asset, most commonly, to the standard deviation, $\hat{\sigma}$, or variance, $\hat{\sigma}^2$, over a set of observations (Poon and Granger, 2003). The variance is given by:

$$\hat{\sigma}_{i,t}^2 = \frac{1}{N-1}\sum_{t=1}^{N}(R_{i,d} - \bar{R}_i)^2, \tag{1}$$

where $R_{i,d}$ is the return of stock i on a given day d and $\bar{R}_i$ is the sample mean of stock i. N is the number of observations in a given week. I use the standard deviation measure, $\hat{\sigma}_{i,t}$, in my analysis; in the regressions below, this measure, at the weekly level, will be denoted by $Volatility_t$.

I also use a range-based volatility measure, $LogRange$, following Alizadeh et al. (2002), in which the volatility is given as the difference in the intraday log high and low quoted prices:

$$LogRange_{i,d} = ln(high_{i,d}) - ln(low_{i,d}). \tag{2}$$

The daily value of $LogRange_{i,d}$ is then averaged at the weekly level.

4.1.2 Liquidity Measures

In my analysis, I use two measures of liquidity (or, conversely, illiquidity): the quoted closing bid-ask spread and Amihud (2002)'s illiquidity measure. Firstly, the quoted bid-ask spread is a simple measure of liquidity; a narrower spread refers to greater liquidity and lower trading costs. In absolute terms, the bid-ask spread is:

$$S_{i,d} = a_{i,d} - b_{i,d},$$

where $a_{i,d}$ and $b_{i,d}$ are the daily closing ask and bid quotes, respectively.

https://doi.org/10.14324/111.444.2755-0877.1611

However, we can look at this in relative terms by dividing this by the midpoint of the two prices:

$$s_{i,d} \equiv \frac{S_{i,d}}{m_{i,d}} = \frac{a_{i,d} - b_{i,d}}{m_{i,d}}, \tag{3}$$

where $m_{i,d} = \frac{a_{i,d}+b_{i,d}}{2}$. In the models presented in Section 4.2, I will denote the relative spread by $Spread_{i,t}$.

Unfortunately, given the limited access to microstructure (intra-day) data on bid-ask spreads, this analysis is limited to closing bid-ask spreads. This may be a limitation to some of the quantitative results presented. The second measure of liquidity used in my analysis, Amihud (2002)'s illiquidity ratio, is given by:

$$Illiq_{i,d} = \frac{|R_{i,d}|}{Volume_{i,d}}, \tag{4}$$

where $|R_{i,d}|$ is the absolute daily return of stock i and $Volume_{i,d}$ is the daily monetary trading volume of the corresponding stock. The weekly average is denoted by $Illiq_{i,t}$. In my analysis, I use the logarithmic transformation, $\log(Illiq)_{i,t}$, to make the interpretation of the regression coefficients more intuitive.

4.2 Regression Models

This section will present the panel autoregressive models used in the analysis. The autoregressive distributed lag (ADL) models employed in the analysis account for the potentially persistent nature of the variables of interest, as well as the attention index in prior weeks. In the analysis, I consider both models that separate the variables of interest, as well as those that combine the various specifications; these can be seen in Section 4.3.

Table 3: Recommended Lags of Each Variable by Information Criterion

Variable	AIC	BIC	Figure
Attention	2	2	*Figure 2*
Volatility	3	2	*Figure 3*
LogRange	1	1	
Spread	3	2	*Figure 4*
log(*Illiq*)	4	4	
1*DRet*	0	0	*Figure 5*
5*DRet*	4	1	
log(*Volume*)	3	1	*Figure 6*

https://doi.org/10.14324/111.444.2755-0877.1611

I consider both the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC); when there are differences between the two criteria, I favour the AIC. This is because the AIC has lower penalties than the BIC for additional lags, favouring 9 larger models than the BIC; this allows for the analysis of dynamics of the lags in this paper. The criteria are given by choose p so as to minimise the following expressions (Stock and Watson, 2020):

$$AIC(p) = \log\left(\frac{SSR(p)}{T}\right) + (p+1)\frac{2}{T}, \quad (5)$$

$$BIC(p) = \log\left(\frac{SSR(p)}{T}\right) + (p+1)\frac{log(T)}{T}, \quad (6)$$

where p is the number of lags and T is the number of time periods. The first term is decreasing in p, whilst the second is increasing; hence, the second term is the penalty for higher lags.[7] The recommended lags for each of the variables according to the criteria can be seen in Table 3.

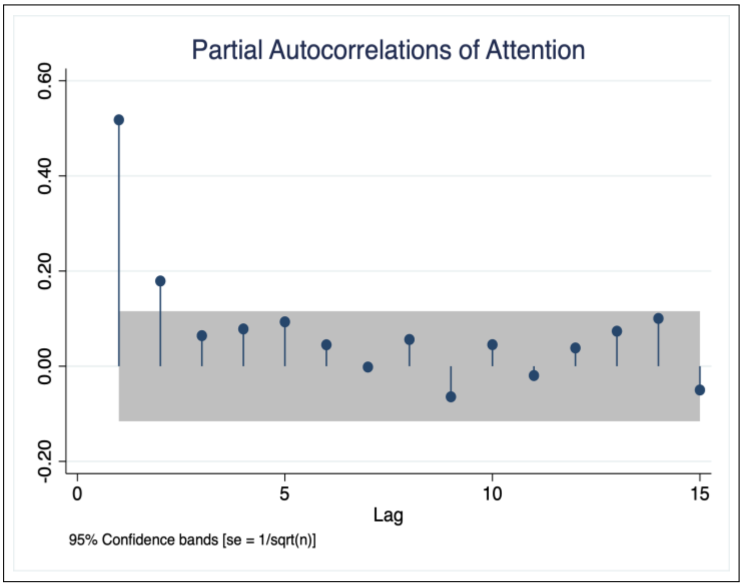

Figure 2: Partial Autocorrelation Function for the Google SVI

Figure 2 shows persistent nature of the Google SVI data, $Attention_t$, through the partial autocorrelation (PAC) function. According to this, after two weeks, the lags of the index produce no additional impact on the current week.

[7] The first term is decreasing, or at least non-increasing, in p as SSR(p) is decreasing, or at least nonincreasing, in p.

https://doi.org/10.14324/111.444.2755-0877.1611

4.2.1 Volatility Models

The volatility models use three lags for $Volatility_t$ and two for $LogRange$. Whilst the AIC recommends one lag for $LogRange$, I increase this to two lags as previous lags may be of interest due to the potential for delayed reactions to the market dynamics, or due to reversals. The PAC functions for the volatility measures can be seen in Figure 3.

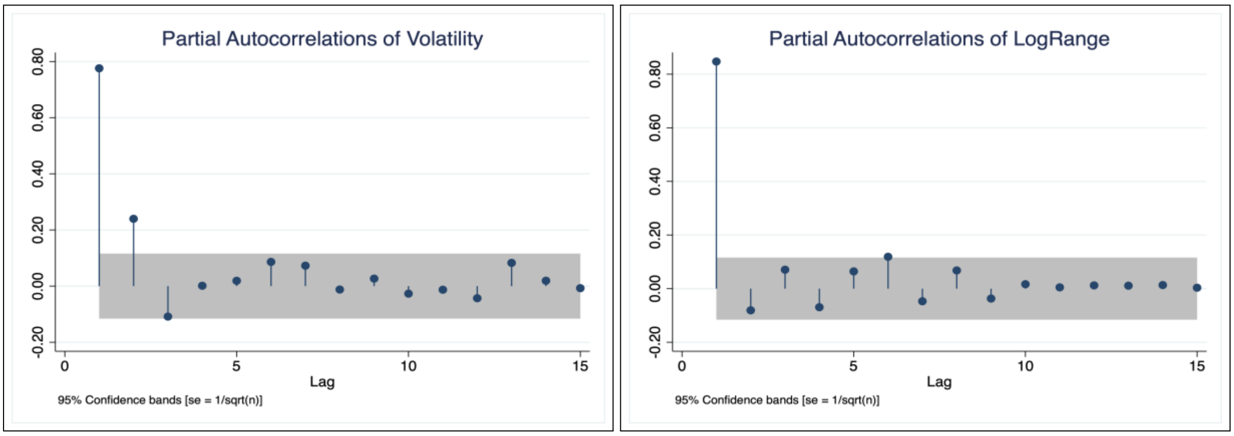

Figure 3: Partial Autocorrelation Functions for *Volatility* and *LogRange.*

Θ_i is a vector of stock-specific characteristics, following Baig et al. (2022).[8] These controls, averaged at the weekly level, are the log of daily turnover $(\log(Turnover)_t)$, closing bid-ask spreads $(ClosingSpread_t)$, and log of the stock's market capitalisation$(\log(MarketCap)_t)$.

Regression Model 1

$$Attention_{i,t} = \alpha + \beta_0 Volatility_{i,t} + \sum_{j=1}^{3} \beta_j Volatility_{i,t-j} + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Theta_i$$

Regression Model 2

$$Attention_{i,t} = \alpha + \beta_0 LogRange_{i,t} + \sum_{j=1}^{2} \beta_j LogRange_{i,t-j} + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Theta_i$$

[8] Although Baig et al. focus on the impact of retail investors on market volatility, the opposite direction of interest to this paper, it is likely that these variables still serve as relevant controls in the models discussed.

https://doi.org/10.14324/111.444.2755-0877.1611

4.2.2 Liquidity Models

Three and four weeks of lags are implemented for $Spread$ and $\log(Illiq)$, respectively. The PACs for the measures are in Figure 3. In the liquidity models, Φ_i is a reduced vector of controls that includes only $\log(MarketCap)_t$ and $\log(Turnover)_t$ due to collinearity when including the absolute spread and relative spread.

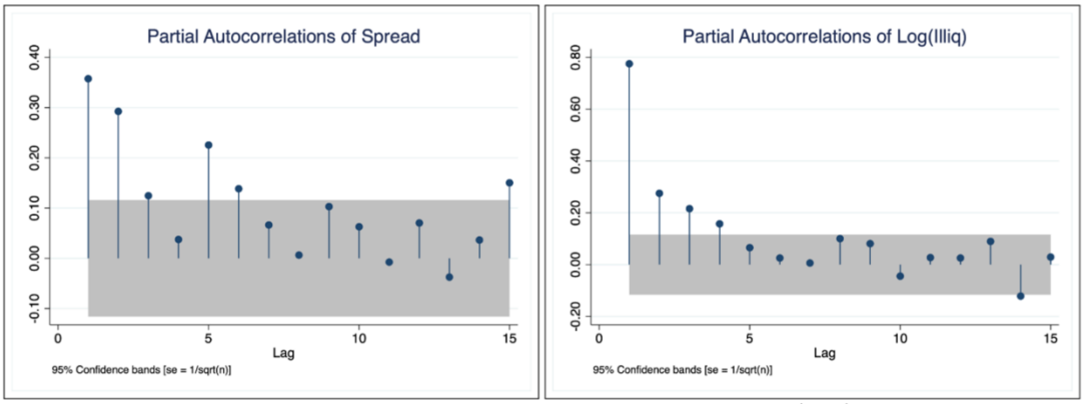

Figure 4: Partial Autocorrelation Functions for $Spread$ and $\log(Illiq)$.

Regression Model 3

$$Attention_{i,t} = \alpha + \beta_0 Spread_{i,t} + \sum_{j=1}^{3} \beta_j Spread_{i,t-j} + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Phi_i$$

Regression Model 4

$$Attention_{i,t} = \alpha + \beta_0 \log(Illiq)_{i,t} + \sum_{j=1}^{4} \beta_j \log(Illiq)_{i,t-j} + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Phi_i$$

Returns Models

In the models below, I use both 1-day ($1DRet_{i,t}$) and 5-day ($5DRet_{i,t}$) returns to analyse how attention responds to changes in the price of a given stock. Both variables are averaged at the weekly level, but the 5-day return allows for the testing of whether returns need to immediate or gradual for them to capture retail attention. $5DRet_{i,t}$ is a 5-day rolling return that is then averaged at the weekly level.

For 1-day returns, the recommendation by the AIC and BIC is to use only the current week, implying no persistence in the time series. However, as mentioned above, it is worth considering the potential for delayed reactions to the change in dynamics; thus, I include two lags of $1DRet$ as a measure. For the 5-day returns, I include four lags. The PACs can be seen in Figure 5.

https://doi.org/10.14324/111.444.2755-0877.1611

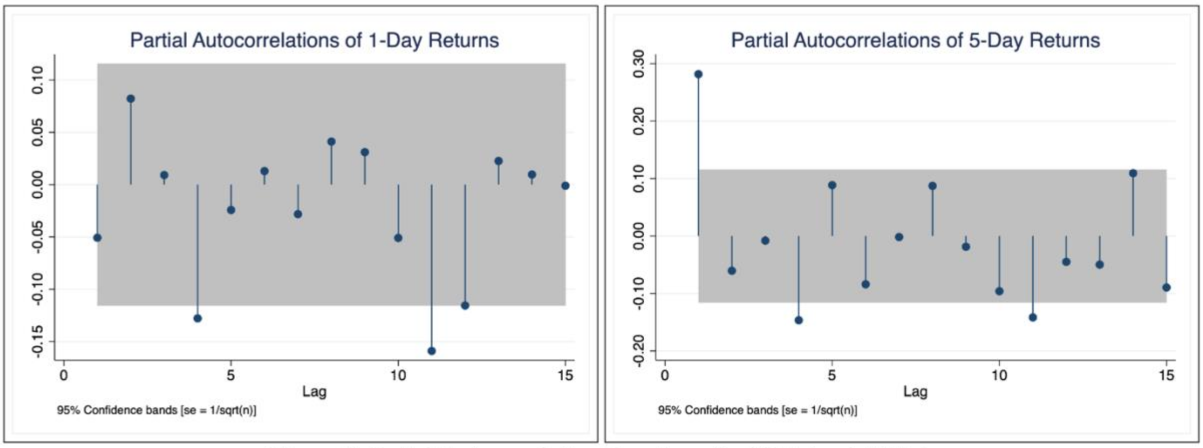

Figure 5: Partial Autocorrelation Functions for $1DRet$ and $5DRet$.

Regression Model 5

$$Attention_{i,t} = \alpha + \beta_0 \cdot 1DRet_{i,t} + \sum_{j=1}^{2} \beta_j \cdot 1DRet_{i,t-j} + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Theta_i$$

Regression Model 6

$$Attention_{i,t} = \alpha + \beta_0 \cdot 5DRet_{i,t} + \sum_{j=1}^{4} \beta_j \cdot 5DRet_{i,t-j} + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Theta_i$$

Trading Volume Model

In the model below, $Volume_{i,t}$ refers to the average daily shares of stock *i* traded in a given week. I use log(*Volume*) to make the interpretation more intuitive. The PAC for log(*Volume*) is shown in Figure 6.

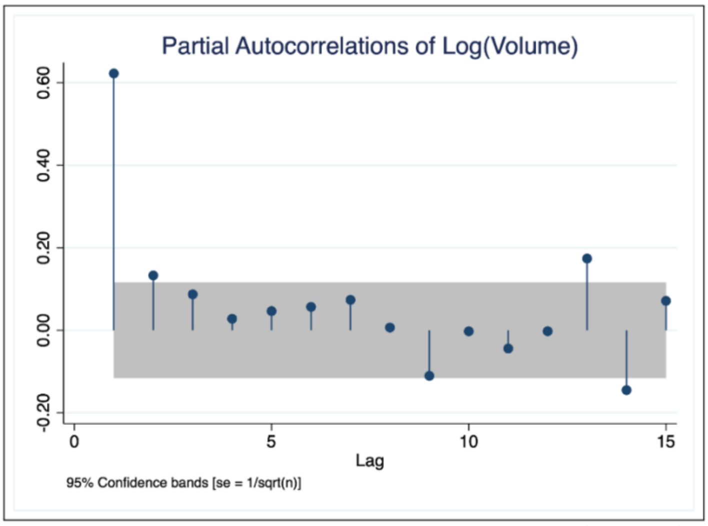

Figure 6: Partial Autocorrelation Function for $\log(Volume)$.

https://doi.org/10.14324/111.444.2755-0877.1611

Regression Model 7

$$Attention_{i,t} = \alpha + \beta_0 \cdot \log(Volume)_{i,t} + \sum_{j=1}^{3} \beta_j \log(Volume)_{i,t-j} + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Theta_i$$

4.3 Combined Regression Specifications

This section combines several of the regression specifications above into a single model to test the effects when considering multiple dynamics. Due to collinearity between the measures, each model will contain only one measure of each variable of interest, except for returns given the differing dynamics of 1-day and 5-day returns. As there are two measures for both volatility and liquidity, the analysis uses four combined models that consider the combinations of these measures. That is, the four will contain the following combination of measures:

$$\{Volatility, Spread\}, \{Volatility, \log(Illiq)\}, \{LogRange, Spread\}, \{LogRange, \log(Illiq)\}.$$

An example for the first specification can be seen in Model 8 below.

Regression Model 8

$$\begin{aligned} Attention_{i,t} \quad = \quad & a + \sum_{j=0}^{3} \beta_{1,j} \cdot Volatility_{i,t-j} + \sum_{j=0}^{3} \beta_{2,j} \cdot Spread_{i,t-j} \\ & + \sum_{j=0}^{2} \beta_{3,j} \cdot 1DRet_{i,t-j} + \sum_{j=0}^{4} \beta_{4,j} \cdot 5DRet_{i,t-j} \\ & + \sum_{j=0}^{3} \beta_{5,j} \cdot \log(Volume)_{i,t-j} + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Phi_i \end{aligned}$$

https://doi.org/10.14324/111.444.2755-0877.1611

4.4 Regressions by Market Capitalisation

It is also of interest to consider how the impacts on attention vary across levels of market capitalisation. I classify a given stock by using the average market capitalisation of that stock over the entire sample; these are split into the bottom 25%, middle 50% (interquartile range), and the top 25% of stocks by market capitalisation. The quartile values for the sample are given in Table 4.

Table 4: Average Market Capitalisation Quartiles

25%	Median	75%
6.10B	8.64B	21.30B

Observations: 87. B refers to billion - i.e., 1B = 1×10^9.
Note: These values use the average of each stock over the entire sample period.

I run the models above conditioning on subsets of the data set, split into the three groups. I then test for statistically significant differences across the regression coefficients using:

$$Z = \frac{\hat{\beta} - \tilde{\beta}}{\sqrt{\hat{se}^2 + \tilde{se}^2}}, \tag{7}$$

under the assumption that $Z \sim N(0, 1)$ in large samples. $\hat{\beta}$ and $\widehat{se}$ refer to the regression coefficient and corresponding standard error of one subset of the data and $\hat{\beta}$ and $\widetilde{se}$ refer to the regression of the coefficient and corresponding standard error of another subset.

5. Results

This section presents the empirical results from the models in Section 4.2. Each of these models are split into three columns in the output below: (i) gives the most basic model, with just the variable of interest and controls; (ii) gives the full model using heteroskedasticity- and autocorrelation-consistent (HAC) standard errors but without accounting for panel fixed effects; (iii) is the full model that accounts for fixed effects and uses HAC standard errors.

The models below use the Newey-West estimator for the HAC standard error, with the truncation parameter, m, as given by Stock and Watson (2020):

$$m = 0.75 \times T^{\frac{1}{3}},$$

where T is the number of time periods within the sample. The parameter, m, is rounded up to the nearest integer. The data set used contains $T = 287$ weeks of data; thus, I set $m = 5$.

Tables 5 and 6 present the output from the volatility regression models for *Volatility* and *LogRange*, respectively. Focusing on column (iii), we see that the coefficients are significant at the 1% level for both volatility measures at week t.

https://doi.org/10.14324/111.444.2755-0877.1611

For *Volatility*, only $Volatility_t$ and $Volatility_{t-2}$ show statistical significance, whilst lags one and three are insignificant. The coefficient on $LogRange_t$ is positive and significant at the 1% level, but $LogRange_{t-1}$ is negative at the 5% level. A one standard deviation increase in *Volatility*, equal to 1.41 pence (£0.0141), correlates with a 1.8 unit increase in the attention index. For *LogRange*, a 1% increase in the range between the daily high and low values coincides with a 1.16% increase in the index. Given that the mean value of the Google SVI is 30.93, this corresponds to a 0.36 unit increase.

Tables 7 and 8 present the output from the liquidity regression models for *Spread* and log(*Illiq*), respectively. Overall, the results are mainly statistically insignificant for the spread models, with some significance when using log(*Illiq*). For relative spreads, column (iii) shows that the coefficient on $Spread_{t-2}$ is statistically significant at the 10% level (p-value = 0.093), however, the remainder of the coefficients for *Spread* are statistically insignificant. Table 8, show that $\log(Illiq)_t$ is significant at the 1% level, along with some of the lags of this measure at varying levels of significance. The results suggest that a 1% increase in the illiquidity ratio coincides with a 1.03 point increase in the attention index in that week. The implications of these findings are discussed in Section 6.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 5: *Volatility* Regression Models

	$Attention_t$		
Model 1	(i)	(ii)	(iii)
$Volatility_t$	131.2128***	186.7333***	130.2388***
	(10.0294)	(13.44)	(12.7593)
$Volatility_{t-1}$		7.3881	9.1998
		(10.8825)	(10.396)
$Volatility_{t-2}$		-67.5204***	-39.4893***
		(12.0943)	(10.8827)
$Volatility_{t-3}$		-5.2394	-0.1816
		(10.2674)	(10.2214)
$Attention_{t-1}$		0.3651***	0.2150***
		(0.0066)	(0.0075)
$Attention_{t-2}$		0.3289***	0.1801***
		(0.0069)	(0.0077)
$\log(Turnover)_t$	5.0519***	0.7449***	3.9762***
	(0.3188)	(0.1955)	(0.2102)
$ClosingSpread_t$	1.2759	-2.7068***	1.6476**
	(1.0336)	(0.6707)	(0.6575)
$\log(Marketcap)_t$	-3.9664***	0.7893***	-2.8887***
	(0.4624)	(0.1849)	(0.2038)
Constant	33.6098***	-23.8176***	15.4382***
	(10.0166)	(2.7796)	(3.2358)
Fixed Effects	-	-	Yes
HAC Standard Errors	-	Yes	Yes
N	24,969	24,708	24,708

Note: $\log(Turnover)$, $ClosingSpread$, and $\log(MarketCap)$ are weekly averages by stock.
( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 6: *LogRange* Regression Models

	Attention$_t$		
Model 2	(i)	(ii)	(iii)
$LogRange_t$	107.2477***	186.0002***	115.6583***
	(10.0294)	(23.4716)	(19.6155)
$LogRange_{t-1}$		-43.6401***	-28.7237**
		(14.4246)	(13.117)
$LogRange_{t-2}$		-44.5276***	-16.3331
		(14.9338)	(12.8642)
$Attention_{t-1}$		0.3666***	0.2159***
		(0.0065)	(0.0075)
$Attention_{t-2}$		0.3269***	0.1779***
		(0.0069)	(0.0077)
$\log(Turnover)_t$	5.3436***	0.9004***	4.3165***
	(0.3186)	(0.2032)	(0.2203)
$ClosingSpread_t$	1.1882	-2.9907***	1.5945**
	(1.0351)	(0.6764)	(0.6583)
$\log(Marketcap)_t$	-4.0494***	0.7395***	-3.1460***
	(0.4707)	(0.1917)	(0.2114)
Constant	30.3459***	-25.4194***	15.7448***
	(10.1405)	(2.8528)	(3.2881)
Fixed Effects	-	-	Yes
HAC Standard Errors	-	Yes	Yes
N	24,969	24,795	24,795

Note: $\log(Turnover)$, $ClosingSpread$, and $\log(MarketCap)$ are weekly averages by stock.
( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 7: *Spread* Regression Models

	$Attention_t$		
Model 3	(i)	(ii)	(iii)
$Spread_t$	-569.3428**	170.5256	-291.5251
	(258.9486)	(207.4634)	(238.1030)
$Spread_{t-1}$		139.0290	-308.0519
		(201.3943)	(206.6847)
$Spread_{t-2}$		189.3880	-329.7626*
		(212.6291)	(196.4395)
$Spread_{t-3}$		529.0155**	16.0428
		(229.5922)	(261.7005)
$Attention_{t-1}$		0.3708***	0.2148***
		(0.0065)	(0.0075)
$Attention_{t-2}$		0.3298***	0.1779***
		(0.0068)	(0.0076)
$\log(Turnover)_t$	6.5969***	0.9053***	5.4481***
	(0.2983)	(0.1666)	(0.1948)
$\log(MarketCap)_t$	-5.9796***	0.6745***	-4.7588***
	(0.4504)	(0.1712)	(0.2137)
Constant	56.5495***	-22.5957***	36.0385***
	(10.1526)	(3.019)	(3.823)
Fixed Effects	-	-	Yes
HAC Standard Errors	-	Yes	Yes
N	24,969	24,708	24,708

Note: $\log(Turnover)$ and $\log(MarketCap)$ are weekly averages by stock. ( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 8: *Illiq* Regression Models

	$Attention_t$		
Model 4	(i)	(ii)	(iii)
$\log(Illiq)_t$	1.9668***	1.2668***	1.0306***
	(0.2103)	(0.2346)	(0.2217)
$\log(Illiq)_{t-1}$		0.6092***	0.2115
		(0.2369)	(0.2196)
$\log(Illiq)_{t-2}$		0.7538***	0.4338*
		(0.244)	(0.2233)
$\log(Illiq)_{t-3}$		0.8838***	0.6483***
		(0.2352)	(0.2194)
$\log(Illiq)_{t-4}$		0.6963***	0.5353**
		(0.2241)	(0.2097)
$Attention_{t-1}$		0.3637***	0.2135***
		(0.0065)	(0.0074)
$Attention_{t-2}$		0.3257***	0.1775***
		(0.0068)	(0.0076)
$\log(Turnover)_t$	7.7065***	4.4560***	6.7731***
	(0.3215)	(0.2638)	(0.2857)
$\log(MarketCap)_t$	-4.8174***	1.2887***	-3.0056***
	(0.4505)	(0.1510)	(0.1896)
Constant	53.4931***	-4.2015	35.0750***
	(9.8888)	(2.7090)	(3.355)
Fixed Effects	-	-	Yes
HAC Standard Errors	-	Yes	Yes
N	24,969	24,621	24,621

Note: $\log(Turnover)$ and $\log(MarketCap)$ are weekly averages by stock.
( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

Tables 9 and 10 present the 1-day and 5-day return models, respectively. For 1*DRet*, only the current week, $1DRet_t$, is statistically significant at the 1% level, with the first and second lags having *p*-values of 0.095 and 0.069, respectively. The coefficient on $1DRet_t$ suggests that a 10% increase in the daily return is associated with a 5.4 point increase in the attention index. For 5*DRet*, a 10% increase in the returns over a 5-day period correlates with an increase in the Google SVI of 1.76 points. I find that, when using a one-sided hypothesis test, the coefficient on $1DRet_t$ is statistically greater than the coefficient for $5DRet_t$ (*p*-value = 0.005).[9] The results suggest that a one standard deviation increase in $1DRet_t$ coincides with a 1.11 unit increase in the SVI, whereas a one standard deviation increase in $5DRet_t$ coincides with a 0.81 unit increase, despite 5*DRet* having a standard deviation that is twice as large as that for 1*DRet*.[10]

Table 11 presents the empirical results from the trading volume regressions. Column (iii) shows that $\log(Volume)_t$ and $\log(Volume)_{t-2}$ are both statistically significant at the 1% level, whereas $\log(Volume)_{t-1}$ and $\log(Volume)_{t-3}$ are statistically insignificant. The coefficient on $\log(Volume)_t$ in column (iii) suggests that a 10% increase in the trading volume correlates with a 0.79% increase in the attention index; using the mean value of the SVI, this corresponds to a 0.24 point increase. The coefficient on $\log(Volume_{t-2})$ is also statistically significant at the 1% level. This coefficient suggests that a 10% increase in trading volume two weeks prior is associated with a decrease in the attention index of 0.15 points. This will be discussed in Section 6.

[9] This hypothesis was tested using the formula for the critical values given in Equation (7).

[10] The standard deviations for $1DRet_t$ and $5DRet_t$ are 0.0206 and 0.0461, respectively (Table 2).

https://doi.org/10.14324/111.444.2755-0877.1611

Table 9: $1DRet$ Regression Models

	$Attention_t$		
Model 5	(i)	(ii)	(iii)
$1DRet_t$	67.7632***	33.2121**	54.037***
	(12.9101)	(14.5796)	(13.6429)
$1DRet_{t-1}$		-7.8089	20.8679*
		(13.0000)	(12.5103)
$1DRet_{t-2}$		-9.9311	21.6406*
		(12.3462)	(11.9163)
$Attention_{t-1}$		0.3702***	0.2151***
		(0.0065)	(0.0074)
$Attention_{t-2}$		0.3282***	0.1764***
		(0.0068)	(0.0076)
$\log(Turnover)_t$	6.6259***	1.6152***	5.4945***
	(0.2980)	(0.2229)	(0.2313)
$ClosingSpread_t$	1.5431	-4.1144***	1.8796***
	(1.0363)	(0.7410)	(0.6716)
$\log(MarketCap)_t$	-5.8404***	-0.0800	-4.6252***
	(0.4417)	(0.1934)	(0.2190)
Constant	52.4018***	-16.4502***	31.5435***
	(9.9391)	(2.6322)	(3.2985)
Fixed Effects	-	-	Yes
HAC Standard Errors	-	Yes	Yes
N	24,969	24,795	24,795

Note: $\log(Turnover)$, $ClosingSpread$, and $\log(MarketCap)$ are weekly averages by stock.
( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 10: *5DRet* Regression Models

	$Attention_t$		
Model 6	(i)	(ii)	(iii)
$5DRet_t$	20.1509***	11.7312***	17.5698***
	(3.2646)	(3.9274)	(3.6867)
$5DRet_{t-1}$		-6.1761*	-0.3866
		(3.3258)	(3.0896)
$5DRet_{t-2}$		0.0744	6.4317**
		(3.2074)	(3.1214)
$5DRet_{t-3}$		-1.1756	1.8341
		(3.3134)	(3.1726)
$5DRet_{t-4}$		-2.4439	2.5111
		(3.0756)	(3.0644)
$Attention_{t-1}$		0.3717***	0.2166***
		(0.0066)	(0.0075)
$Attention_{t-2}$		0.3281***	0.1769***
		(0.0068)	(0.0077)
$Log(Turnover)_t$	6.6722***	1.6226***	5.5806***
	(0.2992)	(0.2242)	(0.2336)
$ClosingSpread_t$	1.4897	-4.1214***	1.9116***
	(1.0369)	(0.7426)	(0.6696)
$Log(MarketCap)_t$	-5.9373***	-0.1064	-4.7911***
	(0.4435)	(0.1942)	(0.2221)
Constant	53.8583***	-16.0200***	33.8236***
	(9.9764)	(2.6274)	(3.3236)
Fixed Effects	-	-	Yes
HAC Standard Errors	-	Yes	Yes
N	24,882	24,534	24,534

Note: $\log(Turnover)$, $ClosingSpread$, and $\log(MarketCap)$ are weekly averages by stock.
( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 11: log($Volume$) Regression Models

	$Attention_t$		
Model 7	(i)	(ii)	(iii)
$\log(Volume)_t$	9.1303***	6.8632***	7.918***
	(0.5112)	(0.4420)	(0.4233)
$\log(Volume)_{t-1}$		-1.0823**	0.4377
		(0.4595)	(0.4243)
$\log(Volume)_{t-2}$		-3.2797***	-1.5254***
		(0.4690)	(0.4317)
$\log(Volume)_{t-3}$		-1.1997***	-0.1730
		(0.3908)	(0.3678)
$Attention_{t-1}$		0.3608***	0.2099***
		(0.0066)	(0.0075)
$Attention_{t-2}$		0.3278***	0.1772***
		(0.0070)	(0.0078)
$\log(Turnover)_t$	-0.8507*	-0.2313	-0.7473***
	(0.5103)	(0.1836)	(0.213)
$ClosingSpread_t$	1.9265*	0.0562	1.7604**
	(1.0307)	(0.6275)	(0.7098)
$\log(MarketCap)_t$	-0.6034	0.5620***	-0.0668
	(0.5276)	(0.1659)	(0.1909)
Constant	-75.778***	-18.7270***	-65.3375***
	(12.2371)	(2.7067)	(3.2005)
Fixed Effects	-	-	Yes
HAC Standard Errors	-	Yes	Yes
N	24,969	24,708	24,708

Note: $\log(Turnover)$, $ClosingSpread$, and $\log(MarketCap)$ are weekly averages by stock.
( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

5.1 Results for Combined Models

This section presents the results from the four combined models in Section 4.3. In the interest of space, Table 12 shows the coefficient of each variable for the current week only, however, the regressions are calculated using the full set of lags. For the extended regression table, see Table B2 in Appendix B.2.

The combined regression models again show that an increase in volatility, using either measure, coincides with an increase in the level of attention towards a particular stock. A one standard deviation increase in $Volatility_t$ approximately correlates with a 1.74 point increase in the SVI. For *LogRange*, a 1% increase correlates with a 1.1 unit increase. Some of the lags of volatility have statistically significant, negative coefficients in the combined models; $Volatility_{t-2}$ is negative in combined models (1) and (2), and $LogRange_{t-1}$ is negative in models (3) and (4). This is discussed in Section 6.

In the combined models, there is little evidence of statistical significance for the liquidity measures, with only model 4 showing some evidence that $\log(Illiq)_t$ is significant at the 5%. This coefficient suggests that a 10% increase the illiquidity ratio corresponds to a 0.05 point increase in the SVI.

Whilst a 10% increase in the 5-day return coincides with a 2.4 point increase, on average across models, in the SVI, 1-day returns seem to have no statistically significant impact. In terms of log(*Volume*), we see that the current week is significant at the 1% level in all four of the combined models. On average, a 10% increase in the daily trading volume correlates with a 0.63 point increase in the index. The first lag of log(*Volume*) is also statistically significant and positive at either the 5% or 10% in all four of the models. Considering only coefficients above the 5% significance level, we see that the coefficients on $\log(Volume)_{t-2}$ are negative in models (3) and (4); a 10% increase in daily trading volume leads to a 1.3 point decrease in the index after 2 weeks.[11]

[11] For the full list of regression coefficients, I refer the reader to Table B2.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 12: Combined Regression Models

	$Attention_t$			
	(1)	(2)	(3)	(4)
$Volatility_t$	121.7839***	125.7509***		
	(12.5112)	(16.7997)		
$LogRange_t$			118.3787***	101.1842***
			(19.032)	(21.5185)
$Spread_t$	2.1243		2.1046	
	(1.4439)		(1.4435)	
$\log(Illiq)_t$		-0.2252		0.5527**
		(0.2982)		(0.2689)
$1DRet_t$	-3.8751	-2.6466	-4.4913	-7.4808
	(25.6738)	(25.7264)	(25.9556)	(25.8730)
$5DRet_t$	22.1644***	20.7105***	27.3421***	27.4927***
	(7.4752)	(7.5374)	(7.5596)	(7.5471)
$\log(Volume)_t$	6.6466***	5.4792***	6.9942***	6.1680***
	(0.4536)	(0.5926)	(0.4679)	(0.5473)
$Attention_{t-1}$	0.2109***	0.2106***	0.2112***	0.2108***
	(0.0076)	(0.0076)	(0.0076)	(0.0076)
$Attention_{t-2}$	0.1761***	0.1760***	0.1754***	0.1751***
	(0.0078)	(0.0078)	(0.0078)	(0.0078)
$\log(Turnover)_t$	-1.1337***	-0.0742	-1.0966***	0.2281
	(0.2061)	(0.5616)	(0.2064)	(0.5013)
$\log(MarketCap)_t$	0.4902***	0.4694***	0.4671**	0.4383***
	(0.1872)	(0.1641)	(0.1888)	(0.1660)
Constant	-68.4804***	-62.3073***	-69.3914***	-60.2172***
	(3.2257)	(4.1247)	(3.2705)	(4.1402)
N	24,534	24,534	24,534	24,534

Note: $\log(Turnover)$ and $\log(MarketCap)$ are weekly averages by stock. ( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

5.2 Results by Market Capitalisation

The models in this section use Models 1 through 7 with fixed effects and HAC standard errors, but condition on subsets of the data, namely, quartiles of market capitalization. The tables below report only the coefficients of the variables of interest, however, the regression is run using the full model, including the control variables and lags of *Attention*.

We can see from Table 13 that the coefficients on $Volatility_t$ are statistically significant at the 1% level for all quartiles of market capitalisation. The coefficient on $Volatility_t$ is greater for the top 25% than the bottom 25% at the 5% level (one-sided p-value = 0.044). The coefficients for $Volatility_{t-2}$ are significant at the 5% and 1% levels for the middle 50% and top 25%, respectively, but not for the bottom 25%. The coefficients on $LogRange_t$ are statistically significant for all three groups, and for $LogRange_{t-2}$ in the model conditional on the top 25%. Despite the coefficient being large for the top 25%, the difference between the three groups is statistically insignificant at the 10% level; this suggests that the magnitude of the effect of $LogRange_t$ on $Attention_t$ does not vary with market capitalisation.

Whilst there are no significant results for *Spread* in Table 14, the results show that the coefficients for $\log(Illiq)_t$ are statistically significant across all levels of market capitalisation. For the top 25% of stocks, a 10% increase in illiquidity correlates with a 0.11 increase in the SVI. However, the coefficients are statistically insignificant across levels of market capitalisation, with p-values greater than 10%.

When considering 1-day and 5-day returns using the conditional models (Table 15), we see that the coefficients on $1DRet_t$ are positive and significant for the middle 50% and top 25%, but not for the bottom 25%. As with the models using the complete data set, the coefficients on lags of 1*DRet* are insignificant above the 5% level. The coefficient for the top 25% is statistically greater than for the bottom 25% (one-sided p-value = 0.063). However, the magnitude of the effect for the top 25% is not statistically greater than the middle 50% (one-sided p-value = 0.275).

The coefficients on $5DRet_t$ are positive and significant for the middle 50% and top 25%, but not for the bottom 25%. For the middle 50%, $5DRet_{t-2}$ is significant at the 5% level. The impact for the top 25% is greater than that for the bottom 25% of stocks (p-value = 0.029). Furthermore, I find that the coefficients on $5DRet_t$ for the top 25% and middle 50% are statistically indifferent at the 10% level, however, the magnitude of $5DRet_t$ for the top 25% is greater than that for the overall model in Table 10 (one-sided p-value = 0.049). As with the models using the complete data set, I find that the coefficients for $1DRet_t$ are statistically greater than those for $5DRet_t$ when considering the middle 50% and the top 25% of stocks; the one-sided p-values are 0.014 and 0.027, respectively.

Table 16 shows that the coefficients for $\log(Volume_t)$ are significant across all three groups at the 1% level. The coefficient on $\log(Volume_t)$ is greater in magnitude for the top 25% than for the bottom 25% and middle 50% (one-sided p-values are <0.0001 in both cases). These results suggest that a 10% increase in daily trading volume is associated with a 2.9 point increase in the Google SVI for largest 25% of stocks, but only 1.4 (0.4) points if the market capitalisation of the stock in question is in the middle 50% (bottom 25%).

https://doi.org/10.14324/111.444.2755-0877.1611

Table 13: Conditional Volatility Models - Market Capitalisation

	$Attention_t$		
	$(< 25\%)$	$(25\% - 75\%)$	$(> 75\%)$
Model 1			
$Volatility_t$	107.8026***	125.4182***	169.4088***
	(23.5497)	(15.7617)	(27.2237)
$Volatility_{t-1}$	15.7332	15.6350	-21.7287
	(17.7681)	(15.2210)	(22.5849)
$Volatility_{t-2}$	-19.1930	-37.8247**	-78.5462***
	(18.0220)	(16.1672)	(22.8793)
$Volatility_{t-3}$	8.7832	13.5062	-47.1725**
	(18.282)	(14.0192)	(22.2586)
N	5,964	12,496	6,248
Model 2			
$LogRange_t$	90.1998**	125.1812***	126.4181***
	(35.2115)	(19.167)	(36.6571)
$LogRange_{t-1}$	-20.0249	-28.0993	-38.8627
	(20.9006)	(18.4965)	(32.9047)
$LogRange_{t-2}$	-2.7910	-2.9353	-82.2820***
	(20.3983)	(17.4676)	(28.1223)
N	5,985	12,540	6,270

The models above use fixed effects and HAC (Newey-West) standard errors.
*: $p - value < 0.1$, **: $p - value < 0.05$, ***: $p - value < 0.01$.
( ) = standard errors.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 14: Conditional Liquidity Models - Market Capitalisation

	$Attention_t$		
	$(< 25\%)$	$(25\% - 75\%)$	$(> 75\%)$
Model 3			
$Spread_t$	-27.7814	-459.8807	32.6191
	(392.2514)	(484.5686)	(240.2769)
$Spread_{t-1}$	-353.6316	254.6537	32.8835
	(400.0749)	(450.0512)	(224.0972)
$Spread_{t-2}$	579.7220	-1103.0110*	-188.5224
	(424.1875)	(562.1837)	(144.4238)
$Spread_{t-3}$	569.3762	460.7781	243.3696
	(485.1770)	(480.3467)	(320.8273)
N	5,964	12,496	6,248
Model 4			
$Log(Illiq)_t$	1.4592***	0.7172**	1.0622***
	(0.4923)	(0.3102)	(0.3982)
$Log(Illiq)_{t-1}$	0.0578	0.2786	-0.0627
	(0.4762)	(0.3054)	(0.4073)
$Log(Illiq)_{t-2}$	0.6360	0.3734	0.0480
	(0.4522)	(0.3266)	(0.4064)
$Log(Illiq)_{t-3}$	1.4804***	0.5093	-0.1282
	(0.4840)	(0.3105)	(0.387)
$Log(Illiq)_{t-4}$	0.4325	0.4834	0.5769
	(0.4389)	(0.2980)	(0.3952)
N	5,943	12,452	6,226

The models above use fixed effects and HAC (Newey-West) standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.
( ) = standard errors.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 15: Conditional Returns Models - Market Capitalisation

	$Attention_t$		
	$(< 25\%)$	$(25\% - 75\%)$	$(> 75\%)$
Model 5			
$1DRet_t$	24.7490	63.6720**	83.8709***
	(26.5784)	(18.8736)	(27.9293)
$1DRet_{t-1}$	26.4432	26.3124	14.4953
	(22.3552)	(17.9796)	(26.1269)
$1DRet_{t-2}$	25.6377	31.8878*	10.7152
	(22.3700)	(16.3020)	(24.9886)
N	5,985	12,540	6,270
Model 6			
$5DRet_t$	8.9805	20.4122***	27.9787***
	(6.6562)	(5.3337)	(7.5093)
$5DRet_{t-1}$	6.9665	0.6296	-10.4196
	(5.8489)	(4.2501)	(6.2620)
$5DRet_{t-2}$	1.7721	11.1796**	7.1928
	(5.5259)	(4.4008)	(6.5668)
$5DRet_{t-3}$	0.3212	-2.2792	17.4227**
	(5.4183)	(4.7442)	(6.2387)
$5DRet_{t-4}$	1.8299	6.4657	-4.6227
	(6.1567)	(4.0597)	(6.5831)
N	5,922	12,408	6,204

The models above use fixed effects and HAC (Newey-West) standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.
( ) = standard errors.

https://doi.org/10.14324/111.444.2755-0877.1611

Table 16: Conditional Volume Models - Market Capitalisation

	$Attention_t$		
	$(< 25\%)$	$(25\% - 75\%)$	$(> 75\%)$
Model 7			
$Log(Volume)_t$	4.5007***	14.2636***	29.4081***
	(0.8494)	(0.6747)	(1.4165)
$Log(Volume)_{t-1}$	1.2581	-0.1867	0.3602
	(0.8789)	(0.5749)	(0.8483)
$Log(Volume)_{t-2}$	-0.9755	-1.5833***	-2.4893***
	(0.8544)	(0.5925)	(0.9050)
$Log(Volume)_{t-3}$	-0.4076	0.0330	-0.8914
	(0.7542)	(0.4942)	(0.8048)
N	5,964	12,496	6,248

The models above use fixed effects and HAC (Newey-West) standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.
( ) = standard errors.

6. Discussion

Firstly, we must consider the extent to which we can identify causality in these results, that is, the extent to which the market dynamics discussed cause increased retail investor attention. There may be a bi-directional relationship between retail investors and the market dynamics in question; Welch (2022) suggests that, whilst retail investors may be a cause of increased volatility, they may also be attracted by it, adding to the complexity of identifying the direction of the effect. One may hypothesise that an increase in the number of traders may increase volatility, which could in turn, attract more retail investors. Similarly, those investors that are attracted by higher price returns may, in turn, contribute to increased stock prices through increased purchasing. A similar case could be made for higher trading volume attracting investors' attention, leading to more trading in that stock. However, by accounting for the lags of both *Attention* and the variables of interest, that is, if the model is correctly specified, one should be able to identify, to a certain degree, the effect of such a loop. Thus, if the feedback effects are accounted for, the remaining effects, unless affected by other factors, should be the causal impact of the variable of interest on attention.

The positive coefficients for both $Volatility_t$ and $LogRange_t$ suggest that an increase in volatility coincides with increased investor attention. One economic interpretation for a causal relationship is that more volatile markets offer the potential for higher returns and short-term profits, incentivizing retail investors to trade. Furthermore, more volatile stocks are likely to attract media attention; the more volatile the stock, the greater the price fluctuations, which, in turn, increases the likelihood that this stock is discussed in the media. Increased investor attention will likely follow from this increased media attention. An additional explanation for a causal effect is the idea of risk-seeking behaviour among retail

https://doi.org/10.14324/111.444.2755-0877.1611

investors; Grinblatt and Keloharju (2009) identifies ‘sensation-seeking’ behaviour as an motivation for retail trading. Hence, more volatile stocks offer higher risks and rewards, appealing to some investors.

From the empirical results, the relative spread models are largely statistically insignificant. However, this is more likely a limit of the data than the true correlation, especially when the Amihud illiquidity measure shows statistical significance. With access to intra-day bid-ask spread data, the statistical power would be improved. There is evidence that periods of lower liquidity correlate with increase attention, as evidence by the positive coefficients on $\log(Illiq)_t$. This is the case in the individual models, as well as some evidence being seen for this in the combined models. The market capitalisation models also show significant results for this. However, it is more likely that these are correlated with a market event rather than the effect of liquidity itself. For example, we saw a drastic decrease in market liquidity at the start of the COVID-19 pandemic, which also brought significant attention to markets. Retail investors replaced much of this liquidity when institutional investors were constrained (Ozik et al., 2021; Pagano et al., 2021).

It would be difficult to say, however, that liquidity itself was a driver of this retail investors trading and attention.

I find strong evidence that increased returns lead to increased investor attention, as suggested by the positive results for $1DRet_t$ and $5DRet_t$. De Bondt (1993) documents that retail investors “expect the continuation of apparent past trends in prices”; this extrapolation likely leads retail investors to become attracted to a stock, believing that it will be an opportunity to earn the same returns that they have just observed. Furthermore, extreme returns are often widely covered by the media, which would translate into increased attention. The impact of returns on attention may also be causal due to the peer effect (Kaustia and Knu¨pfer, 2012); peers earning high returns from a stock would likely incentivise other retail investors to focus their attention on this stock, leading to them trading it.

Table 15 shows that the effect for the top 25% of stocks is greater than for the bottom 25%, suggesting that, for an increase in returns of equal size, the larger stocks attract more investor attention; I put forward two potential reasons for this. Firstly, stocks with larger market capitalisation naturally attract more media attention, given that they are more ‘important’ in the composition of the FTSE 100. Hence, an increase in the returns for larger stocks is more likely to be discussed. Secondly, stocks with higher market capitalisation are less volatile,[12] and, hence, a significant jump in daily return for a stock in the top 25% is likely to capture more attention.

The results suggest that daily trading volume may also cause an increase in retail investor attention, as evidenced by the positive coefficients of $\log(Volume)_t$. One hypothesis for this may be that retail investors like to trade similar stocks to both their peers and to other retail investors in general; there have been extreme cases of this, such as the GameStop episode and the trading of ‘meme stocks’. Retail investors may suffer from confirmation bias; they could already be aware of a stock but a large increase in trading volume may confirm that their stock is popular, given that others are trading it.

[12] Using the standard deviation measure of volatility (Equation (1)), I find that the volatility for the top 25% (0.0171)

https://doi.org/10.14324/111.444.2755-0877.1611

However, it may not be possible to identify an entirely casual relationship between volume and attention. Increased trading volume often coincides with stock events, such as earnings announcements, which makes it difficult to disentangle the impacts of trading volume from the stock event on attention. As such, the extent to which higher trading volume *causes* increased retail investor attention remains unclear; to disentangle these effects, an analysis around stock events using more accurate, intra-day data would be a good starting point.

An interesting observation from the results is the short-run nature of the effects of market dynamics, as evidenced by the either negative or insignificant lags of volatility, returns, and trading volume. $Volatility_{t-2}$ and $LogRange_{t-1}$ both have negative coefficients, suggesting that attention has already fallen approximately two weeks after a shock to volatility, potentially reverting towards the mean. For both 1-day and 5-day returns, shocks at time $(t-1)$ are insignificant (beyond the 10% level) for $Attention_t$; only a shock to returns in week t appears to impact attention. For log(*Volume*), we see that the coefficient is insignificant for the first lag and negative for the second. These results collectively suggest that the attention of retail investors may be primarily short term, with prior shocks to dynamics having little effect.

https://doi.org/10.14324/111.444.2755-0877.1611

7. Conclusion

In this paper, I analyse the impact of changes in markets dynamics (volatility, liquidity, returns, and trading volume) on UK retail investor attention, as measured by the Google SVI. Using a sample of 87 stocks from the FTSE 100 index, I run regression models on both the entire data set, and on subsets of the data, filtered by market capitalisation, to identify potentially heterogeneous effects.

Overall, I find that changes in volatility, returns, and, to some extent, daily trading volume lead to increased retail investor attention in UK markets. Whilst liquidity is negatively correlated with attention, that is, more illiquid stocks attract greater attention, this is likely due to the liquidity-reducing events, such as the COVID-19 market shock. The effects of increased returns and trading volume are more pronounced for the top 25% of stocks by market capitalisation than for the bottom 25%..[13]

I hypothesise that the positive effect of volatility could be due to increased media attention, retail investors seeking short-term gains, or sensation-seeking behaviour. For returns, amongst other reasons, this effect may be due to extrapolation of past returns, with investors expecting that they can earn similar returns. One reason for the impact of trading volume is due to confirmation bias; seeing other investors trade stocks that they have considered may encourage them to follow them to follow the stock more closely. However, it is difficult to identify a purely causal relationship as there exist issues in disentangling the impact of trading volume from a stock event.

A point of interest is the short-term attention of retail investors evidenced in the data. The lags of many of the market dynamics variables are either negative or insignificant, suggesting that, after the current week, these dynamics have little effect on the attention of these investors. These investors are quick to react to changes but their attention not be sustained over periods of more than one week.

Acknowledgements

I am grateful to the UCL Journal of Economics' peer reviewers for their comments that aided in this revision. I would like to thank Rasmus Duret and Suraj Sridhar for their advice and feedback throughout the research process.

[13] is statistically smaller in magnitude than for the bottom 25% (0.0193) (one-sided p-value < 0.001), as calculated using Welch's t-test.

https://doi.org/10.14324/111.444.2755-0877.1611

References

Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-Based Estimation of Stochastic Volatility Models. The Journal of Finance, 57 (3), 1047–1091. https://doi.org/ 10.1111/1540-6261.00454

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets, 5 (1), 31–56. https://doi.org/10.1016/S1386- 4181(01) 00024-6

Baig, A. S., Blau, B. M., Butt, H. A., & Yasin, A. (2022). Do retail traders destabilize financial markets? An investigation surrounding the COVID-19 pandemic. Journal of Banking & Finance, 144, 106627. https://doi.org/10.1016/J.JBANKFIN.2022. 106627

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2022). Attention-Induced Trading and Returns: Evidence from Robinhood Users. The Journal of Finance, 77 (6), 3141– 3190. https://doi.org/10.1111/JOFI.13183

Barber, B. M., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. Review of Financial Studies, 21 (2), 785–818. https://doi.org/10.1093/rfs/hhm079

Barber, B. M., Odean, T., & Zhu, N. (2009). Do Retail Trades Move Markets? The Review of Financial Studies, 22 (1), 151–186. https://doi.org/10.1093/RFS/HHN035

Barrot, J.-N., Kaniel, R., & Sraer, D. (2016). Are retail traders compensated for providing liquidity? Journal of Financial Economics, 120 (1), 146–168. https://doi.org/10. 1016/j.jfineco.2016.01.005

Brandt, M. W., Brav, A., Graham, J. R., & Kumar, A. (2010). The Idiosyncratic Volatility Puzzle: Time Trend or Speculative Episodes? Review of Financial Studies, 23 (2), 863–899. https://doi.org/10.1093/rfs/hhp087

Cheng, F., Chiao, C., Wang, C., Fang, Z., & Yao, S. (2021). Does retail investor attention improve stock liquidity? A dynamic perspective. Economic Modelling, 94, 170–183. https://doi.org/10.1016/j.econmod.2020.10.001

Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. Journal of Finance, 66 (5), 1461–1499. https://doi.org/10.1111/J.1540-6261.2011.01679.X 35

De Bondt, W. P. (1993). Betting on trends: Intuitive forecasts of financial risk and return. International Journal of Forecasting, 9 (3), 355–371. https://doi.org/10.1016/0169- 2070(93)90030

Q Ding, R., & Hou, W. (2015). Retail investor attention and stock liquidity. Journal of International Financial Markets, Institutions and Money, 37, 12–26. https://doi.org/ 10.1016/j.intfin.2015.04.001

Engelberg, J. E., & Parsons, C. A. (2011). The Causal Impact of Media in Financial Markets. The Journal of Finance, 66 (1), 67–97. https: / / doi. org / 10. 1111 / j. 1540 - 6261.2010.01626.x

Foucault, T., Sraer, D., & Thesmar, D. J. (2011). Individual Investors and Volatility. The Journal of Finance, 66 (4), 1369–1406. https://doi.org/10.1111/j.1540-6261.2011. 01668.x

Gavish, L. R., Qadan, M., & Yagil, J. (2021). Net Buyers of Attention-Grabbing Stocks? Who Exactly Are They? Journal of Behavioral Finance, 22 (1), 26–45. https://doi. org/10.1080/15427560.2020.1716360

Grinblatt, M., & Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. Journal of Finance, 64 (2), 549–578. https: / /doi. org / 10. 1111 /J. 1540 - 6261.2009.01443.X

Hsieh, S. F., Chan, C. Y., & Wang, M. C. (2020). Retail investor attention and herding behavior. Journal of Empirical Finance, 59, 109–132. https://doi.org/10.1016/J. JEMPFIN.2020.09.005

Kahneman, D. (1973). Attention and Effort. Prentice-Hall.

Kaniel, R., Saar, G., & Titman, S. (2008). Individual investor trading and stock returns. Journal of Finance, 63 (1), 273–310. https://doi.org/10.1111/J.1540- 6261.2008. 01316.X

Kaustia, M., & Kn¨upfer, S. (2012). Peer performance and stock market entry. Journal of Financial Economics, 104 (2), 321–338. https://doi.org/10.1016/j.jfineco.2011.01. 010

Kelley, E. K., & Tetlock, P. C. (2013). How Wise Are Crowds? Insights from Retail Orders and Stock Returns. The Journal of Finance, 68 (3), 1229–1265. https://doi.org/10. 1111/jofi.12028 36

https://doi.org/10.14324/111.444.2755-0877.1611

Ozik, G., Sadka, R., & Shen, S. (2021). Flattening the Illiquidity Curve: Retail Trading During the COVID-19 Lockdown. Journal of Financial and Quantitative Analysis, 56 (7), 2356–2388. https://doi.org/10.1017/S0022109021000387

Pagano, M. S., Sedunov, J., & Velthuis, R. (2021). How did retail investors respond to the COVID-19 pandemic? The effect of Robinhood brokerage customers on market quality. Finance Research Letters, 43. https://doi.org/10.1016/j.frl.2021.101946

Poon, S.-H., & Granger, C. W. J. (2003). Forecasting Volatility in Financial Markets: A Review. Journal of Economic Literature, 41 (2), 478–539. https://doi.org/10.1257/ 002205103765762743

Seasholes, M. S., & Wu, G. (2007). Predictable behavior, profits, and attention. Journal of Empirical Finance, 14 (5), 590–610. https://doi.org/10.1016/j.jempfin.2007.03.002

Smales, L. (2021). Investor attention and global market returns during the COVID-19 crisis. International Review of Financial Analysis, 73, 101616. https://doi.org/10.1016/j. irfa.2020.101616

Stock, J. H., & Watson, M. W. (2020). Introduction to Econometrics (4th ed.). Pearson Education.

Urquhart, A. (2018). What causes the attention of Bitcoin? Economics Letters, 166, 40–44. https://doi.org/10.1016/j.econlet.2018.02.017

Welch, I. (2022). The Wisdom of the Robinhood Crowd. Journal of Finance, 77 (3), 1489– 1527. https://doi.org/10.1111/JOFI.13128

https://doi.org/10.14324/111.444.2755-0877.1611

Appendix

A. Data

A.1 Refinitiv FTSE Data

This section contains the list of the companies that formed the FTSE 100 as of the start of data collection, 13th June 2022.

Table A1: Stock Symbols of FTSE 100 Constituents

AAF*	BP	HLMA	MRO	SGE
AAL	BRBY	HSBA	NG	SGRO
ABDN	BT-A	HWDN	NWG	SHEL
ABF	CCH	IAG	NXT	SKG
ADM	CPG	ICP	OCDO	SMDS
AHT	CRDA	IHG	PHNX	SMIN
ANTO	CRH	III	PRU	SMT
AUTO	DCC	IMB	**PSH***	SN
AV	DGE	INF	PSN	SPX
AVST*	DPH	ITRK	PSON	SSE
AVV	**EDV***	ITV	REL	STAN
AZN	ENT	JD	RIO	STJ
BA	EXPN	KGF	RKT	SVT
BARC	FLTR	LAND	RMG	TSCO
BATS	FRES	LGEN	RMV	TW
BDEV	GLEN	LLOY	RR	ULVR
BKG	GSK	LSEG	RS1	UU
BLND	HBR	MGGT	RTO	VOD
BME	HIK	MNDI	SBRY	WPP
BNZL	HL	**MNG***	SDR	WTB

*The available data for these stocks does not start from 01 January 2017. As such, observations on these stocks have been dropped and the analysis is conducted excluding these.

A.2 Google SVI Data

In some cases, the Google Trend search term was appended with the word ‘share’ to avoid ambiguity from other related terms that were unrelated to the share in question. According to Google Trends, the phrase ‘stock price’ is search more frequently than ‘share price’ in the U.S.; the opposite is true for the U.K. The average value of the search index for both phrases between 2004 and 2022 can be

https://doi.org/10.14324/111.444.2755-0877.1611

seen in Figure A1 below. Therefore, given that my analysis focuses on UK investors, I choose to use the term ‘share’ over ‘stock’ when making these changes to the search term. For example, to avoid confusion with the search term ‘ITV’ for ITV plc., this was combined with ‘share’ to identify those search about the share price. Therefore, the search term for this stock is “ITV share”.

To determine which ticker symbols were not sufficiently precise to ensure that the search was related to investment, I compared the related searches for the ticker symbol. Google Trends shows ‘related queries’ and ‘related topics’; if these sections were not related to similar searches for the stock ‘XYZ’ or to financial topics (i.e., ‘XYZ share price’ or searches for other stocks), then the search term was changed to include ‘share’.

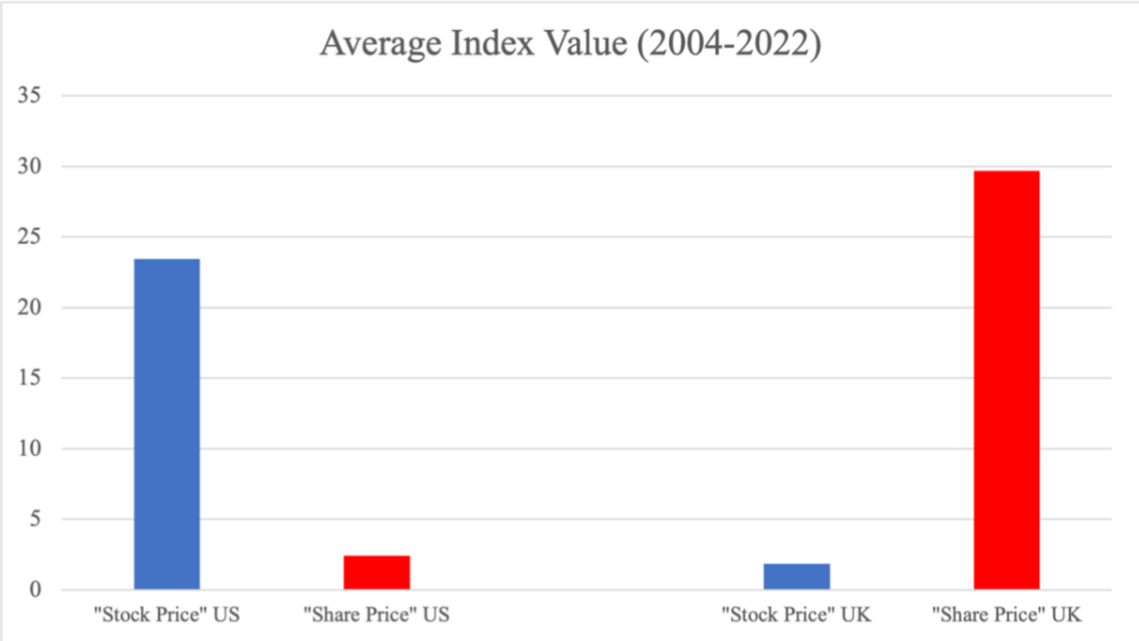

Figure A1: Google SVI: “Stock Price” vs “Share Price”

https://doi.org/10.14324/111.444.2755-0877.1611

A.2.1 List of Noisy Stock Symbols

Below is a table containing the list of stock symbols that are potentially ambiguous. In the main data set, these search terms are appended by 'share'. For the robustness checks in the regressions that use this data, see Appendix B.1.

The list of noisy stock symbols are given in Table A2

Table A2: List of Noisy Google SVI Ticker Symbols

AUTO	DCC	KGF	PSN	STAN
BA	ENT	MNG	REL	SVT
BATS	GLEN	MRO	RIO	TW
BME	IHG	NG	RR	UU
BP	III	NWG	RTO	VOD
BT-A	INF	NXT	SHEL	WTB
CCH	ITV	PHNX	SN	
CPG	JD	PRU	SSE	

These search symbols are appended with the word 'share' to avoid ambiguity from using the ticker symbol alone.

There are eight stocks for which the Google SVI data is excluded. This is due to the fact that the ticker symbol only was too ambiguous, but there was missing data on the stock symbol plus 'share'. These stocks are in Table A3.

Table A3: List of Excluded Google SVI Searches

DPH	FLTR	HIK	ICP	LAND	RS1	SDR	SKG

https://doi.org/10.14324/111.444.2755-0877.1611

B. Additional Results

B.1 Robustness Checks for Google Trends Data

This section contains the additional regression outputs when considering potentially noisy ticker searches in the Google SVI data. The list of noisy ticker symbols is given in Appendix A.2.1.

In this section, I test the null hypothesis that the regression coefficients do not change across subsets of the Google SVI data, namely, whether a stock is ‘noisy’ or not. I use a dummy variable, Noisy, which is equal to 1 if a stock is appended with the term ‘share’, and 0 otherwise. I then interact the variable of interest with the dummy variable, and use this to create a dummy regression model. For example, the model for $Volatility$ (Model 1) is given by:

$$Attention_{i,t} = \alpha + \beta_0 Volatility_{i,t} + \sum_{j=1}^{3} \beta_j Volatility_{i,t-j} + \sum_{x=0}^{3} \phi_x \left(Noisy \times Volatility_{i,t-x}\right) + \sum_{k=1}^{2} \gamma_k Attention_{i,t-k} + \delta \cdot \Theta_i$$

Table B1 below shows the regression output for the interaction terms and the corresponding F-statistic, with the null hypothesis $\phi_0 = \phi_1 = \cdots = \phi_N$ (N is the number of lags of the variable of interest that have been used).[14] I find that only the coefficient for $1DRet_t$ is affected by the noisy measures beyond the 10% significance level.

Several of the F-statistics suggest that, when considering all lags, there is some effect that is due to the noisy variables. This is the case for $Volatility$, $LogRange$, log($Illiq$) and log($Volume$). Each of these variables are largely insignificant individually, but there is evidence of some joint effect. As the analysis primarily considers the current week and the first two weeks of lags, I also run an F-test for only these interaction terms. Here, apart from $LogRange$ which does not change, the interaction terms are insignificant. This could be due to over-specification of the models, or indeed due to some underlying effect. Overall, these results suggest that there are minor effects, although not completely absent, caused by the noisy attention data.

[14]Whilst the regressions were run with the entire model (including fixed effects and HAC standard errors), only the interaction terms are reported.

https://doi.org/10.14324/111.444.2755-0877.1611

Table B1: Robustness Checks: $Noisy \times$ Variable of Interest

	$Volatility$	$LogRange$	$Spread$	$\log(Illiq)$	$1DRet$	$5DRet$	$\log(Volume)$
$Noisy \times \text{Variable}_t$	9.1602	15.7951	47.2368	0.5175	57.2989**	13.6594*	-0.368
	(22.4826)	(31.4773)	(698.0509)	(0.4214)	(27.4672)	(7.3838)	(0.7993)
$Noisy \times \text{Variable}_{t-1}$	-0.8330	-3.5435	608.315	0.0713	-7.7165	-6.9165	-0.2085
	(20.218)	(26.2698)	(758.621)	(0.4427)	(24.9727)	(6.2734)	(0.8574)
$Noisy \times \text{Variable}_{t-2}$	25.8289	45.4765*	409.8175	-0.0712	-14.6581	0.2206	0.9501
	(22.1219)	(26.7919)	(749.5927)	(0.4523)	(23.9824)	(6.3317)	(0.8741)
$Noisy \times \text{Variable}_{t-3}$	11.4486		390.6142	-0.0126		12.0889*	1.0548
	(19.8276)		(814.7634)	(0.4344)		(6.4736)	(0.7394)
$Noisy \times \text{Variable}_{t-4}$				0.2268		3.2744	
				(0.4171)		(6.1963)	
N	24,708	24,795	24,708	24,621	24,795	24,534	24,708
F-test	0.0089	0.0000	0.8610	0.0000	0.1567	0.1460	0.0000
F-test (2 lags)	0.4302	0.0000	0.8652	0.5576	0.1567	0.2389	0.7456

The models above use fixed effects and HAC standard errors. *: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.
( ) = standard errors.

https://doi.org/10.14324/111.444.2755-0877.1611

B.2 Additional Tables

Table B2 presents the full regression table for Table 12 in Section 5.1.

Table B2: Combined Regression Models - Extended

	$Attention_t$			
	(1)	(2)	(3)	(4)
$Volatility_t$	121.7839***	125.7509***		
	(12.5112)	(16.7997)		
$Volatility_{t-1}$	-4.6062	-8.4201		
	(11.3317)	(14.3951)		
$Volatility_{t-2}$	-35.8628***	-53.0886***		
	(11.8206)	(13.9118)		
$Volatility_{t-3}$	-2.0331	-19.6089		
	(10.886)	(13.4118)		
$LogRange_t$			118.3787***	101.1842***
			(19.032)	(21.5185)
$LogRange_{t-1}$			-52.0147***	-59.8384***
			(14.0494)	(14.8943)
$LogRange_{t-2}$			-6.0147	-16.6895
			(13.9807)	(14.5471)
$Spread_t$	2.1243		2.1046	
	(1.4439)		(1.4435)	
$Spread_{t-1}$	-0.9472		-1.0172	
	(0.9367)		(0.9245)	
$Spread_{t-2}$	-0.6717		-0.673	
	(0.979)		(0.9692)	
$Spread_{t-3}$	-0.1664		-0.1405	
	(1.0154)		(1.0175)	
$\log(Illiq)_t$		-0.2252		0.5527**
		(0.2982)		(0.2689)
$\log(Illiq)_{t-1}$		0.0934		0.2331
		(0.2827)		(0.2509)
$\log(Illiq)_{t-2}$		0.5358*		0.1079
		(0.2748)		(0.2494)
$\log(Illiq)_{t-3}$		0.5161*		0.3087
		(0.2736)		(0.2329)
$\log(Illiq)_{t-4}$		0.1227		0.1278
		(0.2141)		(0.2163)
$1DRet_t$	-3.8751	-2.6466	-4.4913	-7.4808
	(25.6738)	(25.7264)	(25.9556)	(25.8730)
$1DRet_{t-1}$	-14.1609	-13.3464	-32.7843	-37.3857
	(36.3447)	(36.2664)	(36.4902)	(36.3747)
$1DRet_{t-2}$	16.4058	16.3343	-2.1579	-4.7574
	(28.664)	(28.6354)	(28.4476)	(28.421)
Continued	⋮	⋮	⋮	⋮

https://doi.org/10.14324/111.444.2755-0877.1611

	(1)	(2)	(3)	(4)
Continued	⋮	⋮	⋮	⋮
$5DRet_t$	22.1644***	20.7105***	27.3421***	27.4927***
	(7.4752)	(7.5374)	(7.5596)	(7.5471)
$5DRet_{t-1}$	7.3548	6.3061	11.9927	12.0936
	(7.9815)	(7.9932)	(8.0665)	(8.0501)
$5DRet_{t-2}$	8.1689	7.692	8.481	7.7389
	(5.4487)	(5.476)	(5.4486)	(5.4570)
$5DRet_{t-3}$	5.0854	5.2188	4.0047	3.5688
	(3.2704)	(3.2768)	(3.2652)	(3.2897)
$5DRet_{t-4}$	5.083*	4.8735	4.6834	4.7491
	(3.0654)	(3.0572)	(3.0783)	(3.0658)
$\log(Volume)_t$	6.6466***	5.4792***	6.9942***	6.168***
	(0.4536)	(0.5926)	(0.4679)	(0.5473)
$\log(Volume)_{t-1}$	0.986**	1.1093**	1.2334***	1.5396***
	(0.4609)	(0.5178)	(0.4641)	(0.4908)
$\log(Volume)_{t-2}$	-0.8628*	-0.3587	-1.4254***	-1.1673**
	(0.4707)	(0.5188)	(0.4688)	(0.4947)
$\log(Volume)_{t-3}$	-0.424	0.098	-0.4008	-0.212
	(0.4022)	(0.4676)	(0.3688)	(0.3886)
$Attention_{t-1}$	0.2109***	0.2106***	0.2112***	0.2108***
	(0.0076)	(0.0076)	(0.0076)	(0.0076)
$Attention_{t-2}$	0.1761***	0.176***	0.1754***	0.1751***
	(0.0078)	(0.0078)	(0.0078)	(0.0078)
$\log(Turnover)_t$	-1.1337***	-0.0742	-1.0966***	0.2281
	(0.2061)	(0.5616)	(0.2064)	(0.5013)
$\log(MarketCap)_t$	0.4902***	0.4694***	0.4671**	0.4383***
	(0.1872)	(0.1641)	(0.1888)	(0.1660)
Constant	-68.4804***	-62.3073***	-69.3914***	-60.2172***
	(3.2257)	(4.1247)	(3.2705)	(4.1402)
N	24,534	24,534	24,534	24,534

Note: $\log(Turnover)$ and $\log(MarketCap)$ are weekly averages by stock.
( ) = standard errors.
*: $p-value < 0.1$, **: $p-value < 0.05$, ***: $p-value < 0.01$.

https://doi.org/10.14324/111.444.2755-0877.1611

C. Additional Materials

The data sets and stata .do files are available at:

https://www.dropbox.com/sh/mvv7aoioljh0rl0/AAAOACkFkQQj-ARyVMD5ec ba?dl=0

- FTSE PricingData Full Clean Py.dta is the (cleaned) FTSE 100 pricing data set obtained from the Refinitiv Data platform.
- Google Trends-Stock Symbols Clean.dta is the Google SVI (attention) data including the week indicators and the *Noisy* dummy variable.
- RI Dynamics Attention Code Revision.do is the Stata commands for the formatting of the variables, merg- ing of the pricing and attention data, as well as the regression commands.

https://doi.org/10.14324/111.444.2755-0877.1611