Retail Investor's Active Attention and Stock Liquidity

Rong Ding and Wenxuan Hou*

Abstract

We use the search volume index (SVI) for firm's ticker from Google Trends to capture the retail investor's active attention on the constituents of S&P 500 from January 2004 to December 2009. 95% of its cross-sectional variation cannot be explained by the passive attention measures including the total number of news available on the internet obtained from Google News, and the advertising expenditures. We show that the firms with increased retail investor attention, reflected by the level and the change of SVI, are associated with a larger shareholder base, and with a significant improvement in the stock liquidity. The results are robust to the control of passive attention measures and to the alternative liquidity measures.

***Keywords*:** Investors attention; breadth of ownership; liquidity; bid-ask spread; investor recognition hypothesis.

JEL Classification:

First Draft: 23 January 2011
Preliminary draft
All comments welcome

* Rong Ding is at Middlesex University Business School, and Wenxuan Hou is at Durham Business School, Durham University. Address correspondence to Wenxuan Hou, Durham Business School, Mill Hill Lane, DH1 3LB, United Kingdom. Tel: +44 (0)191 334 5321; Fax: +44 (0)191 334 5201.
E-mail Address: wenxuan.hou@durham.ac.uk

1. Introduction

Prior theoretic studies on asset pricing posit that investor attention is a necessary condition for stock price to fully incorporate public information, as investors have to be aware of the information before they can perceive and react to the information (Hirshleifer and Teoh, 2003; Peng and Xiong, 2006). However, as a scared resource in reality, investor attention is by no means constantly and equally allocated to each listed firms in the global financial markets, but are largely concentrated on the stocks the investor knows, and especially the ones they are interested in or familiar with. Among others, the visibility of a listed firm proxied by advertising expenditures (Grullon et al., 2004) and media coverage (Fang and Peress, 2009) captures the passive attention of the investors by assuming that the investors passively perceive the information available to them.

In this paper, we employ an active attention measure, which does not require such critical assumption, namely the aggregate search volume index (SVI) proposed by Da et al. (2011) and available in Google Trends. It reflects the search frequency of the ticker of stocks. It is intuitive appeal in that one of the most convenient ways for retail investor to access to the financial information of a listed firm, which they are paying attention, is to search its ticker in Google, the dominant search engine. This active attention measure has been found to be distinct from the passive ones: although they are positively related, almost 95% of the cross-sectional variation of the former cannot be explained by the latter.

We then exam how the level of active attention measure and its change affect the breath of ownership of listed firms. With a control of the passive attention measures, both of them have been found to increase the number of investors. This is in line with the argument of Barber

and Odean (2008) that (individual) retail investor tend to search for information about the firm's history, product, environment and strategies when they are selecting stocks, and can be interpreted with the "investor recognition hypothesis" advanced by Merton (1987) that the shareholder base measures the recognition of the firm among investors, so an enlarged shareholder base indicates the firm has been well recognized. In other words, potential investors have to be aware of a firm before they gradually get familiar with it and eventually decide to invest in it, suggesting that investor attention is the necessary condition for a firm to be recognized. So when more people pay attention to a firm, a certain proportion of them would acquire additional information on its history, product, general business environment and competitive strategies. However, as suggested by Da et al. (2011), this is less likely to be the case for institutional investors, who have access to more sophisticated resources of information, and employ explicit criteria to diversify their investments.

The active attention mitigates the information asymmetry. Prior studies (Glosten and Harris, 1988; Stoll, 1989; George *et al*., 1991; Callahan, *et al.* 1997; Brockman and Chung, 1999) show that the bid-ask spread has three components: order-processing cost, inventory-holding cost and adverse selection cost, among which the adverse selection cost is in general more significant[2]. Hence we expect that the active attention enables to lower adverse selection cost and a reduced corresponding component in the bid-ask spread. The empirical results based on constituents of S&P 500 index over the period January 2004 to December 2009 confirm our prediction.

The results are robust to the control of the passive attention measures, firm characteristics and other documented determinants from the literature. In addition to the advertising

[2] In quota-driven markets such as NYSE, the adverse selection costs exist because dealers are facing un-informed traders and informed traders. In such markets, dealers optimize their pay-offs by maximizing the difference between the gains from trades with un-informed traders and losses from trades with informed traders (Coughenour and Shastri, 1999).

expenditures, we incorporate the number of the news available on the internet obtained from Google News and argue that they are more relevant to the retail investors who search information on the internet[4]. The results are also robust to the alternative measures of the liquidity. Apart from relative bid-ask spread (Amihud and Mendelson, 1986; Erwin and Miller, 1998; Grullon *et al*. 2004), we also use effective spread, relative effective spread, and turnover rate (trading volume divided by shares outstanding) as alternative proxies for liquidity.

To the best of our knowledge, our study is the first that documents a cross-sectional relation between investor's active attention, breadth of ownership and stock liquidity. Our paper adds to the burgeoning literature on investor attention and asset pricing dynamics, including Barber and Odean (2008) on investor attention and individual investors' trading behavior, Engelberg and Parsons (2011) on the casual impact of local media coverage on local trading and Da et al. (2011) on investors attention measured by Google search frequency and its effect on IPO returns and price change in the subsequent periods.

Our study also contributes to the stream of literature that examines the "investor recognition hypothesis" (e.g., Grullon *et al.*, 2004; Tetlock et al., 2009; Fang and Peress, 2009).[8] Although investors continuously receive information on assets that are traded on the financial market, they are unlikely to pay sufficient attention to each piece of the information, because there are limits on the central cognitive-processing capacity of human beings. In markets with incomplete information, information asymmetry become more severe for stocks with lower

[4] Yuan (2009), Tetlock (2009) and Fang and Peress (2009) use either LexisNexis database or Dow Jones news archival to search the number of newspaper article related to a stock. Googel news channel includes news from the most popular English-language news sites such The New York Times, Bloomberg, Reuters, Guardian, CBS News, BBC News, Times Online, CNN, and thus offers a broader news coverage.

[8] Our study is related to but different from Grullon *et al*, (2004), in that our paper focus on the relation between investors' active attention (on a stock) and its shareholder base as well as liquidity, while Grullon *et al*, (2004) investigate firm's advertising expenditure as a (passive) approach to reach a broad audience and its impact on breadth of ownership and liquidity.

investor recognition. When individual investors pay more attention to a stock by actively searching it on the Internet, they acquire relevant information on the stock leading to reduced level of asymmetric information problem for the stock. As a result, stocks with higher investor attention become more liquid.

The finding also has implications for companies that wish to enjoy the benefits of better investor recognition. Companies may intentionally promote themselves on the Internet to attract the attention of potential investors; Our results might be of interest for participants of financial market (e.g., liquidity traders) in that they may benefit from sophisticated models that incorporate individual's searching behavior into the prediction of stock liquidity; Finally, our findings may incentivize search engine companies to innovate their business model. For example, if search engine companies can better their data on public's searching behavior in terms of timeliness and accuracy, they might be able to sell these data to interested parties that can analyze and benefit from such information.

The rest of the paper is organized as follows. Section 2 describes the data and research design. Section 3 presents the empirical results. Section 4 tests the causality between investors attention and stock liquidity, and section 5 provides robustness check. The last section concludes.

2 Research design and Data

2.1 Active Attention Measures

Google Trends (http://www.google.com/trends) retain the search frequency data over time since the beginning of 2004. It reflects how many searches have been made for a term

relative to the total number of searches done on Google over time. Hence, it is not biased by the fact that there are more people using internet to collect information. The search volume is normalized with fixed scale to ease the comparison of values in different periods. We include the constituents of S&P 500 over a period of 6 years from January 2004 to December 2009. We believe this has little impact on our result: the search volume index (SVI) for small firms is often not available anyway due to their relatively low search volume. Following Da et al. (2011), we obtain the SVI for stock tickers as the identification of stocks. It is less ambiguous and therefore more appropriate than company name because the latter often has multiple meanings (e.g. "Apple" or "Amazon"). Retail investor can get firm's ticker in financial news in which tickers are often reported in parentheses. Same as Da et al. (2011), we also exclude SVI with value of zero.

In addition, we also compute the SVI change with the following approach to capture the change of investors attention:

$$\Delta SVI_t = Ln(SVI_t) - Ln[Med\ (SVI_{t-1}, \ldots\ldots, SVI_{t-8})] \qquad (1)$$

, where SVI is the search volume index during the week *t* from Google Trends and [*Med* (*SVI* $_{t-1}$, ……,*SVI* $_{t-8}$)] is the median value of SVI during the previous eight weeks. SVI and ΔSVI_t are then transformed into monthly frequency by taking the average because other controls variables used in our study are mainly on the monthly or yearly basis.

2.2 Passive Attention Measures

A commonly used passive attention measure is the newspaper coverage. Fang and Peress (2009) focus on four daily news papers with nationwide circulation: *New York Time*, *USA Today*, *Wall Street Journal*, and *Washington Post*. We, however, argue that retail investors are very unlikely to subscribe more than two newspapers. A more convenient and inexpensive way for them to read news is from internet, and every piece of news on the internet is with at least "global circulation" because everyone is able to access to it.

The number of news available related to a term over time is documented in the database of Google News (http://news.google.com), which aggregates news from 4,500 English-language news sources worldwide. News reads become less likely to stick with one publication and then look for interesting headlines, but use search engine which offers a wider variety of perspectives from which to choose. Another advantage is that stories are sorted without regard to political viewpoint or ideology.

The advanced news search section in Google News allows us to get the total number of relevant news for each company in our sample available on the internet from 2004 to 2009 on the annual basis. To obtain the news number, we use the company name instead of the ticker in that tickers are only reported in financial newspapers, but the retail investors do not necessarily get information from financial newspapers only. As we address, the multiple meanings of some companies may add error to our sample, and due to the large number of news, it is unlikely for us to read through each to exclude the irrelevant ones. However, this noise biased against us to find the results.

2.3 Research Design and Data

To exam how the investor attention affect the breath of ownership and stock liquidity, we add the attention measures into the models of Grullon et al. (2004) as follows:

$$\begin{aligned} LnNumS &= \lambda_0 + \lambda_1 SVI + \lambda_2 LnNews + \lambda_3 LnAdv \\ &+ \lambda_4 LnAge + \lambda_5 RET + \lambda_6 ROA + \lambda_7 LnMC + \lambda_8 (1/P_t) \\ &+ \lambda_9 LnTurnover + \lambda_{10} LnVolatility + \lambda_{11} NASDAQ + \varepsilon \end{aligned} \quad (1)$$

$$\begin{aligned} RBAS &= \gamma_0 + \gamma_1 SVI + \gamma_2 Ln.News + \gamma_3 LnAdv \\ &+ \gamma_4 LnAge + \gamma_5 ROA + \gamma_6 LnMC + \lambda_7 (1/P_t) \\ &+ \gamma 8 LnTurnover + \gamma_9 LnVolatility + \gamma_{10} NASDAQ + \varepsilon \end{aligned} \quad (3)$$

, where the number of shareholders (*lnNumS*) and the relative bid-ask spread (*RBAS*) are respectively regressed against the search volume index (*SVI*), the number of news available online (*LnNews*), and the advertising expenses (*lnAdv*). The annual number of shareholders and advertising expenses are both obtained from Compustat. A large proportion of firms do not report their advertising expenses, hence in order to keep the sample size, we assume that all missing observations for advertising expenditures are equal to zero. However, when we perform the study based on the reduced sample, similar results are documented. Likewise, Grullon et al. (2007) also find this assumption does not change much of their results.

The relative bid-ask spread is the monthly average of the ratio of the daily inside spread divided and the midpoint of the daily inside spread from CRSP (Centre for Research in Security Prices). Chung and Zhang (2009) suggest daily CRSP-based spread is a good substitute of the TAQ-based spread in that the former represents at least 91% (78%) of cross-sectional variation of the latter from NASDAQ (NYSE/AMEX) stocks. We also drop the observations of relative spread which is greater than 50% of the midpoint to filter the data for errors.

In the robustness checks, we replicate the analysis by using alternative liquidity measures including the change in relative spread, the effective spread and the relative effective spread, and get similar results. The change of relative spread is the monthly change of relative spread in percentage. The effective spread is constructed as twice the difference of the transaction price and the spread midpoint. The relative effective spread is the effective spared scaled by the midpoint of the spread.

Other control variables include firm age (*LnAge*), average monthly return (*RET*), return of assets (*ROA*), firm market capitalization (*lnMC*), the inverse of closing price (1/*P*), share turnover (*LnTurnover*), return volatility (*LnVolatility*) and an exchange dummy (*NASDQA* equals to 1 for firms listed in NASDQA, and 0 otherwise). Firm age is the number of years the firm has existed in CRSP. Average monthly return is the average of daily stock return from CRSP. Return on assets is constructed from Compustat as the annual operating income before depreciation scaled by total assets. Share turnover constructed Share turnover is constructed from CRSP as the monthly average of share volume divided by shares outstanding. Return volatility is the monthly average of the standard deviation of daily returns from CRSP. Following Grullon et al. (2004), some variables are taken natural logarithm, and average monthly return (*RET*) is only incorporated in the regression model for the number of shareholders (*lnNumS*).

Our sample starts in 2004, the beginning of the Google Trends database. The final sample includes 14,690 firm-month observations over the period of 2004 to 2009. The top and bottom 0.5% of the variables are winsorised to control the impact of the outliers. To confirm

our predictions, we expect a significantly positive λ_1 in Equation (2) and a significantly negative λ_1 in Equation (3).

2.4 Summary statistics

Table 1 presents the descriptive statistics of the variables. It can be inferred that our sample is populated with large and mature firms, as the mean (median) of market capitalization (in logarithm) and firm age are 9.33 (9.17) and 25.53(34.12) respectively. Furthermore, there is significant cross-sectional variance in breadth of ownership, liquidity and profitability. For example, the relative bid-ask spread ranges from 1.81 (25^{th} percentile) to 3.47 (75^{th} percentile), the logarithm of number of individual shareholders ranges from 1.32 (25^{th} percentile) to 3.93 (75^{th} percentile) and ROA ranges from 0.30 (25^{th} percentile) to 0.66 (75^{th} percentile). The mean (median) of SVI and logarithm of number of news coverage are 0.0079 (0.0013), and 7.05 (7.09) respectively.

<< Insert Table 1 about here >>

Table 2 provides the correlation between variables. Consistent with our expectation, both SVI and news coverage are positively and significantly correlated with shareholder base,. Advertising expenditure, firm age and firm size are also positively correlated with shareholder base. Both advertising expenditure are firm age are negatively correlated with relative bid-ask spread, suggesting that old firms and firms spending more on advertising have high liquidity in terms of lower spread. SVI is positively correlated with advertising expenditure and firm size, which indicates that large firms and firms with higher advertising expenditure attract more attention from investors. Turnover and volatility are positively

correlated with SVI, suggesting that firms catching more investor attention have high trading volume and return volatility. This is consistent with the findings that investors draw attention to stocks experiencing high turnover and become net buyers of these attention-grabbing stocks (Gervais, 2001; Barber and Odean, 2008). NASDAQ is positively correlated with SVI, which implies that firms traded on NASDAQ attract more attention from investors. Consistent with results that larger and old firms are covered more by major newspaper documented by Falkerestein (1996), news coverage is positively correlated with firm age and firm size. Finally, advertising expenditure is positively related to firma size and NASDAQ, which indicates that larger firms and NASDAQ firms tend to spend more on advertising.

<< Insert Table 2 about here >>

3 Results

3.1 The determinants of SVI/SVI change

Although SVI/SVI change is arguably a direct measure of investors attention (Da *et al*., 2011), it is important to investigate how firm-specific characteristics are related to SVI/ SVI change. In model (1) we regress SVI on Log(number of news), Log (advertising), Log(turnover), firm age, ROA, firma size and return volatility. The results reported in Table 3 suggest that news coverage and advertising expenditure are positively associated with investors attention measured by SVI, which suggests that more investors search for firms that are covered more by news and that have more spending on advertising. Next, the coefficient of turnover, ROA and firm size are positive and significant, which can be interpreted as profitable firms, firms with high trading volume and firm with large market value already gain sufficient investors attention. Furthermore, the coefficient of return volatility is positive and significant, indicating that the investors may increase their search frequency for companies with higher risk, consistent with the findings reported by Seaholes and Wu (2007)

that stocks with higher return or higher risk receive more news coverage and therefore attract more attention among investors. The coefficient of firm age is significantly negative. In model (2) we regress SVI change on the same independent variables. Only the coefficients of turnover, firm size and return volatility are positive and significant, which indicate that large firms and stocks with higher trading volume and higher volatility attract incremental attention from investors.

<< Insert Table 3 about here >>

3.2 The effect of investors attention on breadth of ownership

3.2.1 Univariate analysis

We divide our sample firms according to their level of SVI into high-SVI (above-median) and low-SVI (below median) sub-samples and compare the mean of number of shareholders across the sub-samples. Consistent with our expectation, the number of shareholder (mean) of high-SVI subsample is significantly higher than that of low-SVI sub-sample at 1% level (High-low difference= 47.31, t= 14.38), which suggests that stocks catching more attention tend to have a bigger shareholder base. Next, we partition the sample according to market capitalization into small (below-median) and large (above-median) group, then divide each sub-sample into high-SVI and low-SVI sub-samples before comparing the number of shareholders (mean) for each group. The findings that mean of number of shareholders is higher for high-SVI sub-sample hold for both small and large group, which suggests that higher investors attention is associated with larger shareholder base. The results are provided in Table 4.

<< Insert Table 4 about here >>

3.2.2 Multivariate analysis

In this section we analyze the association between investors attention and the shareholder base. We regress the logarithm of number of shareholder on SVI, news coverage, advertising expenditure and a battery of control variables documented by previous literature to explain cross-sectional differences in the breadth of ownership. The results are reported in Table 5. In model (1), the independent variables are SVI and the control variables. We also introduce the year and industry fixed effect to control for secular trends and other un-modeled industry-specific effects, but for parsimony we don't report these coefficients in the tables. The coefficient of SVI is positive and significant at 5% level (0.02, t= 2.36), suggesting a positive association between investors attention reflected by search frequency and the shareholder base. Among the control variables, the coefficients of firm age, firm size and ROA are significantly positive, consistent with the expectation that profitable firms, large firms and long-standing firms enjoy a larger shareholder base. The coefficients of 1/P is positive and significant, in line with the explanation that individual investors are likely to buy stocks with lower price (higher 1/P). The coefficient of return volatility is negative and significant, which suggests that few investors buy stocks with high volatility. The coefficient of turnover is significantly negative, indicating that few investors tend to hold stocks with high trading volume. Finally, the coefficient of dummy variable NASDAQ is negative and significant. In model (2), we include logarithm of number of news as a second measure of investor attention. The results are consistent with those in model (1), as the coefficients of both SVI and Log (number of news) are positive and significant at 5% level (0.02, t= 2.39; 0.04, t= 5.51). The sign and significance of control variables remain unchanged. In model (3), we introduce logarithm of advertising expenditure as additional independent variable. The coefficients of SVI and Log (number of news) are positive and significant at 5% level (0.02, t= 2.33; 0.04, t= 5.42), while the coefficient of Log (advertising) is positive and insignificant at 10% level (0.006, t= 1.57), which suggests that the positive association between investors

attention measured by search frequency and news coverage and shareholder base is robust after controlling the positive relationship between advertising expenditure and breadth of ownership documented by Grullon *et al*., (2004).

<< Insert Table 5 about here >>

3.3 The effect of investors attention on liquidity

In this section we test the association between investors attention and stock liquidity. Our early findings show that when firms receive more attention reflected by search frequency, they are likely to have an enlarged shareholder base and thus become recognized by more investors. As a result, the increased recognition and familiarity of the firm reduce the information asymmetry between informed and uninformed investors, leading to reduced level of bid-ask spread. Thus, we expect to find support for a negative association between investors attention and bid-ask spread. Furthermore, we use turnover rate as a second measure of liquidity, and expect a positive association between investors attention measured by SVI and turnover rate.

3.3.1 Univariate analysis

We follow the same procedure to divide our sample firms into high-SVI and low-SVI sub-samples and compare the mean of relative bid-ask spread across the sub-samples. The results show that relative bid-ask spread (mean) of high-SVI subsample is significantly lower than that of low-SVI sub-sample at 1% level (High-low difference= -0.20, t= 6.38), which suggests that stocks catching more attention tend to be more liquid. Next, we partition the sample according to market capitalization into small (and large groups, then divide each sub-sample into high-SVI and low-SVI sub-samples before comparing the mean of relative spread for each group. We find that the difference of relative spread is more pronounced for small group, as small firms are on average less recognized by investors, so they are likely to benefit more from increased investors attention. The results are presented in Table 6.

<< Insert Table 6 about here >>

3.3.2 Multivariate analysis

Table 7 provides results relating relative bid-ask spread to investors attention measured by SVI and new coverage. In model (1), SVI is included as the independent variable to reflect investors attention. We also control for year and industry fixed effects. The coefficient of SVI is negative and significant at 5% level (-0.01, t= -2.18), suggesting that higher level of investors attention measured by search frequency leads to reduced bid-ask spread. Among the control variables, the coefficients of reverse of share price, turnover and return volatility are positive and significant. In model (2), we include Log (number of news) as an additional measure for investors attention and find the coefficients of SVI is significantly negative (-0.01, t= -2.17), while the coefficient of Log (number of news) is positive and marginally significant (0.007, t= 1.71).[13] In model (3), we include Log (advertising) as independent variable. The coefficient of SVI remains negative and significant (-0.01, t= -2.11), whereas the coefficient of Log (number of news) is positive and significant (0.007, t= 1.80). The coefficient of Log (advertising) is negative but insignificant (-0.003, t= -1.36). The results confirm the negative association between investors attention reflected by search frequency and relative spread after controlling for advertising expenditure, which is documented by Grullon *et al*. (2004) to have negative effect on relative spread.

<< Insert Table 7 about here >>

Table 8 provides the results on analysis relating turnover rate to investor attention reflected by search frequency. In model (1), the coefficient of SVI is positive and significant at 1%

[13] We attribute the positive relationship between Log(number of news) and relative spread to the fact that news coverage of a stock can be a noisy measure of investor attention. For instance, users of Windows of Microsoft may search the company on Googel news to have updated information about the latest version of Windows. In contrast, people who use company ticker to search on Google are more likely to be investors of Microsoft.

level (0.01, t= 3.10), suggesting that higher level of investors attention leads to higher turnover rate. Among the control variables, the coefficients of firm age, ROA, firm size and reverse of share price are negative and significant, whereas the coefficients of return volatility and NASDAQ are significantly positive. In model (2), we include Log (number of news) as an additional measure for investors attention and find the coefficients of SVI and Log (number of news) are positive and significant (0.01, t= 3.14; 0.01, t= 3.58). In model (3), we add Log (advertising) as independent variable. The coefficients of SVI and Log (number of news) remains positive and significant (0.01, t= 2.77; 0.01, t= 2.99), and the coefficient of Log (advertising) is strongly positive (0.01, t= 10.35), consistent with findings documented by Grullon *et al*., (2004). The findings show that the positive relationship between turnover and investor attention is robust after the effect of advertising on turnover being controlled for. The findings are generally consistent with the investor recognition hypothesis, in that when firms are searched more by Internet users or have more news coverage, their stocks become more liquid, as these firms obtain higher recognition among investors.

<< Insert Table 8 about here >>

So far all we have identified are time-series association between investors attention and liquidity. We cannot yet say anything about the direction of causality. It's easier to argue that more investors attention leads to higher liquidity. But the reverse can also be true. For example, investors pay more attention to liquid stocks. Sorting out causality requires rigid test, which we turn to in the next section.

4 Test on the causality between investor attention and stock liquidity

In this section we rely on the Granger test to examine the causality between stock liquidity and investors attention measured by SVI. That is, we explore whether changes in search frequency lead to changes in liquidity reflected by relative spread, or *vice versa*. Apart from

including the same battery of control variables, we control for time fixed effects in the regression.

First, in order to check whether SVI changes leads to change in liquidity, we use standard F to test the joint hypothesis that $\delta_1 = \delta_2 = ...\delta_{10} = 0$ for the following regression:

$$RelativeSpread_{i,t} = \alpha + \beta_1 RelativeSpread_{i,t-1} + \beta_2 RelativeSpread_{i,t-2} + ...\beta_{10} RelativeSpread_{i,t-10} + \delta_1 SVIChange_{i,t-1} + \delta_2 SVIChange_{i,t-2} + ..\delta_{10} SVIChange_{i,t-10} + Controls + TimeFixedeffect + \varepsilon$$

Next, to examine whether change in liquidity causes SVI change, we test the joint hypothesis that $\kappa_1 = \kappa_2 = ...\kappa_{10} = 0$ for the following regression:

$$SVIChange_{i,t} = \sigma + \kappa_1 RelativeSpread_{i,t-1} + \kappa_2 RelativeSpread_{i,t-2} + ...\kappa_{10} RelativeSpread_{i,t-10} + \gamma_1 SVIChange_{i,t-1} + \gamma_2 SVIChange_{i,t-2} + ..\gamma_{10} SVIChange_{i,t-10} + Controls + TimeFixedeffect + \varepsilon$$

The results show that the hypothesis that SVI change does not cause change in liquidity is rejected at 10% confidence level. In contrast, we fail to reject the hypothesis that change in liquidity does not cause SVI change at conventional level. The findings stay robust after excluding firm size as control variable, as investors may use firms size as a proxy for liquidity. This confirms our conjecture that there is a casual relationship from SVI change to change in liquidity.

5 Robustness check

We perform the following tests to check the robustness of our findings.

First, to ensure that our results are not driven by outliers we truncate our sample at 1 and 99 percentile, then re-run the regressions. The results (unreported) are qualitatively unaffected.

Second, we use SVI change (the level of SVI in current week relative to the average of SVI in the previous eight weeks) as alternative measure for investor attention and re-run the

regressions. We start with the investigation of the association between shareholder base and investor attention reflected by SVI change. The results are displayed in Table 9. In model (1), the independent variable of interest is SVI change. We consistently introduce the year and industry fixed effect. The coefficient of SVI change is positive and significant at 10% level (0.21, t= 1.93), suggesting a positive association between change of investors attention and shareholder base. In model (2), we include logarithm of number of news as an additional measure of investor attention. Consistent with results in model (1), the coefficients of both SVI change and Log (number of news) are positive and significant at 5% level (0.21, t= 1.96; 0.04, t= 5.52). In model (3), we include logarithm of advertising expenditure as independent variable. The coefficients of SVI change and Log (number of news) are positive and significant at 5% level (0.21, t= 1.96; 0.04, t= 5.42), while the coefficient of Log (advertising) is positive and significant at 10% level (0.006, t= 1.64), which suggests that the results stay robust after controlling the positive relationship between advertising expenditure and breadth of ownership documented by previous studies.

<< Insert Table 9 about here >>

Next, we examine the association between relative spread and investor attention measured by SVI change. Table 10 presents results. In model (1), the coefficient of SVI change is negative and significant at 1% level (-0.24, t= -4.30), suggesting that higher level of investors attention reflected by incremental search frequency leads to reduced bid-ask spread. In model (2), after we include Log (number of news) as a second measure for investors attention, the coefficient of SVI change remain significantly negative (-0.24, t= -4.30) while the coefficient of Log (number of news) is positive and marginally significant (0.007, t= 1.70). In model (3), we include Log (advertising) as additional independent variable. The coefficient of SVI change stays negative and significant (-0.24, t= -4.30), whereas the coefficient of Log (number of

news) is positive and significant (0.007, t= 1.79). The coefficient of Log (advertising) is negative but insignificant.

<< Insert Table 10 about here >>

Furthermore, we explore the relationship between turnover and investors attention measured by SVI change. Table 11 displays the results. In model (1), the coefficient of SVI change is positive and significant at 1% level (0.12, t= 3.80), suggesting that higher level of investors attention reflected by incremental search frequency leads to higher turnover rate. In model (2), after including Log (number of news) as an additional measure for investors attention, we find the coefficients of SVI change and Log (number of news) stay positive and significant (0.12, t= 3.80; 0.009, t= 3.59). In model (3), after controlling the effect of Log (advertising) on turnover, the coefficients of SVI change and Log (number of news) remain positive and significant (0.12, t= 3.80; 0.008, t= 2.99). The coefficient of Log (advertising) is also positive and significant (0.011, t= 10.44), consistent with findings of Grullon *et al*., (2004). Overall the results show that the relation between investors attention and liquidity remain invulnerable for alternative measures of investors attention.

<< Insert Table 11 about here >>

As a final robustness check we employ effective spread and relative effective spread as alternative liquidity measures and re-run the regressions. Following previous literature (Grullon *et al*., 2004), we calculate the effective spread as twice the difference of the transaction price and the spread mid-point, and relative effective spread as the effective spread scaled by the mid-point of the spread. The results are presented in Table 12.

Consistent with our main findings, the coefficients of SVI change and Log (number of news) are negative and significant at 5% level when both effective spread and relative effective spread are used as the dependent variable, which confirms that firms with more investors attention tend to have higher liquidity in terms of alternative measures of stock liquidity.

<< Insert Table 12 about here >>

6 Conclusion

In this paper we use search frequency provided by Google as a direct measure of investors attention and address the following research questions: 1). What is the association between investors attention measured by search frequency and the breadth of ownership? 2). What is the association between the investors attention and stock liquidity?

Based on the analysis of constituents of S&P 500 index over the period January 2004 to December 2009, we find that increased investors attention in terms of higher search volume index (SVI) and number of news covered by Google contributes to a broader shareholder base. Furthermore, increased investors attention is associated with reduced relative bid-ask spread and higher turnover rate. Our findings are robust after controlling for well-documented determinants of both bid-ask spread and turnover rate.

Our study contributes to the growing literature on the role of investor attention in the dynamics of asset pricing. Among studies in this strand of literature are Barber and Odean (2008) on investor attention and individual investors' trading behavior, Yuan (2008) on recording-breaking events of the Dow index and front page coverage in newspaper as proxy of investor attention and its impact on trading behavior and market returns and Da *et al*., (2011) on investor attention measured by Google search frequency and its effect on IPO returns and price pressure hypothesis proposed by Barber and Odean (2008).

Our study also extends to the stream of literature that examines the "investor recognition hypothesis" (e.g., Grullon *et al.*, 2004; Fang and Peress, 2009). In markets with incomplete information, investors are less likely to possess the information of all securities. Consequently, securities with lower investor recognition become less liquid and have to offer higher return to compensate for their "illiquidity". The fact that a security is being attended

by investors can partially alleviate the problem of information asymmetry. For example, when people pay more attention to a stock by actively searching it on the Internet, they acquire relevant information on the stock and may eventually become investor of it, thus enabling the stock to be better recognized. As a result, stocks with increased investors attention become more liquid. Our results generally provide support to the "investor recognition" hypothesis.

Reference

Agarwal, P., 2009. Institutional ownership and stock liquidity. Working paper, Cornell University.

Amihud, Y., Mendelson., H. 1986. Asset Pricing and The Bid-Ask Spread. Journal of Financial Economics 17, 223-249.

Barber, B.M., Odean, T., 2008. All that Glitters: The Effect of Attention and News on the Buying Behaviour of Individual and Institutional Investors. Review of Financial Studies 21, 785-818.

Biais, B., Bossaerts, P., 1998. Asset Prices and Trading Volume in a Beauty Contest. Review of Economic Studies 65, 307-340.

Brav, A., Heaton, J.B., 2002. Competing Theories of Financial Anomalies. Review of Financial Studies 15, 575-606.

Brennan, M. J., Subrahmanyama, A. 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics 41, 441-464.

Brockman, P., Chung, D.Y., 1999. Bid-ask spread components in an order driven environment. The Journal of Financial Research 12, 227-246.

Callahan, C.M., Lee, C.M.C., Yohn, T.L., 1997. Accounting information and bid-ask spread. Accounting Horizons 11, 50-60.

Chan, L.K. C., Lakonishok, J. 2004. Value and Growth Investing: Review and Update. Financial Analysts Journal 60, 71-86.

Chemmanur, T., Yan, A., 2009. Advertising, attention and stock returns. Working paper, Boston College and Fordam University.

Chung, K.H., Zhang, H. (2009). A simple approximation of intraday spreads using daily data. Working paper, State University of New York at Buffalo. Available at http://ssrn.com/abstract=1346363

Copeland, T. E., Galai, D. 1983. Information Effects on the Bid-Ask Spread. Journal of Finance 38, 1457-1469.

Da, Z., Engelberg. J., Gao, P. 2011. In Search of Attention. Journal of Finance, forthcoming

Erwin, R., Miller, J.M. 1998 The liquidity effects associated with addition of a stock to the S&P 500 index: Evidence from bid/ask spreads. Financial Review 33, 131-146.

Fang, L., Peress, J. 2009. Media Coverage and the Cross-section of Stock Returns. Journal of Finance 64, 2023–2052.

Engelberg. J., Parsons, C. 2011. The casual impact of media in financial markets. Journal of Finance 66, 67-97.

Frino, A., Jarnecic, E., Lepone, A. 2007. The determinants of the price impact of block trades: further evidence. Abacus 43, 94-106

George, T.J., Kaul, G., Nimalendran, M., 1991. Estimation of the bid-ask spread and its components: a new approach. Review of Financial Studies 4, 623-656.

Gervais, S., 2001. The High-Volume Return Premium. Journal of Finance 56, 877-919.

Glosten, L., Harris, L., 1988. Estimating the components of the bid-ask spread, Journal of Financial Economics 21, 123-142.

Grullon, G., Kanatas, G., Weston, J., 2004. Advertising, Breadth of Ownership, and Liquidity. Review of Financial Studies 17, 439-461.

Hou, K., Peng, L., Xiong, W., 2008. A tale of two anomalies: the implications of investor attention for price and earnings momentum. Working paper, Ohio State University and Princeton University.

Merton, R., 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information. Journal of Finance, 42, 483-510.

Morris, S., 1996. Speculative Investor Behavior and Learning. Quarterly Journal of Economics 110, 1111-1133.

Pashler, H., Johnston, J., 1998. Attention limitations in dual-task performance. In: Pashler, H (Eds). Attention, Psychology Press.

Scheinkman, J., Xiong, W., 2003. Overconfidence and Speculative Bubbles. Journal of Political Economy 111, 1183-1219.

Seasholes, M. S., Wu, G., 2007. Predictable behavior, profits, and attention. Journal of Empirical Finance 14, 590-610.

Stoll, H., 1976. Dealer Inventory Behaviour: An Empirical Investigation of NASDAQ Stocks. Journal of Financial and Quantitative Analysis 11, 359-380.

Stoll, H., 1989. Inferring the components of the bid-ask spread: Theory and empirical tests. Journal of Finance 44, 115-134.

Tetlock, P.C., 2009. Does Public Financial News Resolve Asymmetric Information. Working Paper, Columbia University.

Yuan, Y., 2009. Attention and trading. Working paper, University of Iowa.

Table 1: Descriptive statistics

	Mean	Std.Dev	25th percentile	Median	75th percentile	N
Firm Characteristics						
SVI change	0.0079	0.12	-0.04	0.0013	0.047	14690
Log (number of news)	7.05	1.85	5.75	7.09	8.11	14690
Log (advertising)	-0.19	4.97	-4.61	-4.61	5.10	14690

Log(firm age)	3.24	0.93	2.71	3.53	3.85	14690
Stock return	0.0098	0.10	-0.04	0.01	0.06	14690
ROA	0.57	0.50	0.30	0.44	0.66	14690
Firm size	9.33	1.15	8.51	9.17	10.08	14690
1/Share price	0.04	0.04	0.02	0.03	0.04	14690
Log (return volatility)	-4.02	0.56	-4.42	-4.08	-3.68	14690
NASDAQ	0.15	0.36	0	0	0	14690
Breadth of Ownership						
Log (number of Shareholders)	2.66	1.93	1.32	2.68	3.93	14690
Liquidity Measures						
Relative Spread	2.96	1.78	1.81	2.41	3.47	14690
Log (turnover)	0.60	0.74	0.13	0.57	1.06	14690

SVI and Number of news are downloaded from Google trends and Google news respectively. SVI change is calculated as $\Delta SVI_t = SVI_t - Med (SVI_{t-1}, ...SVI_{t-8})$. The advertising expenditure, monthly share price and bid-ask spreads are collected from CRSP. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. Stock return is calculated as: Return= $(P_t-P_{t-1})/ P_{t-1}$. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured by logarithm of market value (share price multiplied by number of shares outstanding). Return volatility is calculated as the standard deviation of monthly returns over the year. NASDAQ is a dummy variable taking 1 for stocks traded on NASDAQ, 0 other wise. The number of common shareholders is downloaded from Compustat. Relative spread is calculated as the quoted spread divided by the mid-point of the bid and ask prices. Turnover is calculated as a monthly trading volume divided by number of shares outstanding.

T-statistics are reported in parentheses below the coefficient estimates.
*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 2: Correlation

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.Log(shareholder)	1												
2.Relative Spread	-0.1524*	1											
3.SVI change	0.0243*	0.0137	1										
4.Log(news)	0.2716*	0.0052	0.0051	1									
5.Log(advertising)	0.1528*	-0.0282*	0.0175*	0.1258*	1								
6.Log(firm age)	0.2729*	-0.0930*	-0.0057	0.0770*	-0.0118	1							
7.Return	-0.0229*	-0.1412*	-0.0019	0.0018	-0.0228*	-0.0305*	1						
8.ROA	0.1299*	-0.0513*	0.0028	0.0671*	-0.0253*	-0.0590*	-0.0071	1					
9.Firm size	0.4868*	-0.2676*	0.0315*	0.3728*	0.2134*	0.1268*	0.0298*	0.1322*	1				
10.1/P	0.003	0.3334*	0.0007	-0.0102	0.0453*	-0.0015	-0.0910*	-0.0807*	-0.3231*	1			
11.Log(turnover)	-0.2466*	0.6092*	0.0267*	-0.0362*	-0.0159*	-0.1842*	-0.0194*	-0.0921*	-0.2837*	0.2049*	1		
12.Log(return volatility)	-0.1955*	0.8831*	0.0348*	-0.0071	-0.0384*	-0.1337*	-0.0866*	-0.0732*	-0.2801*	0.2939*	0.6657*	1	
13.NASDAQ	-0.1577*	0.0548*	0.0278*	-0.1279*	0.0921*	-0.2726*	0.0031	-0.1022*	-0.0392*	0.0844*	0.2232*	0.0821*	1

* Coefficient is significant at 5% level (2-tailed);

Table 3: The effect of firm-specific characteristics on SVI /ASVI

	Dependent Variables	
	Model (1) : SVI	Model (2) : SVI change
Log (number of news)	0.013***	-0.0004
	(6.29)	(0.42)
Log (advertising)	0.008***	0.0002
	(7.64)	(1.16)
Log (turnover)	0.08***	0.0004**
	(9.13)	(2.2)
Log(firm age)	-0.03***	-0.0004
	(-4.24)	(-0.35)
ROA	0.07***	-0.0001
	(6.71)	(-0.06)
Firm size	0.08***	0.005***
	(14.5)	(5.02)
Log(return volatility)	0.03***	0.007***
	(2.65)	(2.75)
N	14,690	14,690
Adjusted-R^2	0.05	0.003

The table presents the estimates of regressions relating SVI/SVI change to a series of firm-specific characteristics. SVI is downloaded from Google trends. SVI change is calculated as $\Delta SVI_t = SVI_t$ – Med (SVI_{t-1}, ...SVI_{t-8}). Number of news are downloaded from Google news. The advertising expenditure is collected from Compustat. Share turnover rate is calculated as a monthly trading volume divided by number of shares outstanding. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured as the logarithm of market value of equity(stock price multiplied by number of shares outstanding). Return volatility is computed as the standard deviation of monthly returns over the year.

*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 4: Univariate analysis on the effect of investors attention on the breadth of ownership

	Low-SVI	High-SVI	High-low difference	t-statistics
All stock	46.96	94.27	47.31***	14.38
Small stock	22.45	24.72	2.27**	2.06
Large stock	78.86	153.24	74.38***	11.54

The table presents a comparison of mean for number of shareholders (‘000) between low and high SVI sub samples. The sub-samples are constructed according to the sample mean of SVI level. Further, we partition the sample into small stock and larger stock groups according to market capitalization, then divide each group into low and high SVI subsample and compare the mean of number of shareholders across sub samples. Reported average number of shareholders is based on equally weighted cross-sectional means. SVI is collected from Google trends and the number of shareholders is collected from Compustat. The market value of equity is constructed as stock price multiplied by number of shares outstanding.

*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 5: Regression analysis: the effect of investors attention on the breadth of ownership:

	Dependent Variable: Log (Number of Shareholders)		
	(1)	(2)	(3)
SVI	0.02**	0.02**	0.02**
	(2.36)	(2.39)	(2.33)
Log (number of News)		0.04***	0.04**
		(5.51)	(5.42)
Log (advertising)			0.01
			(1.57)
Log(firm age)	0.27***	0.27***	0.27***
	(12.20)	(12.25)	(12.25)
Return	-0.20	-0.20	-0.20
	(-1.59)	(-1.58)	(-1.56)
ROA	0.31***	0.31***	0.31***
	(8.66)	(8.77)	(8.80)
Firm size	0.90***	0.88***	0.87***
	(73.06)	(66.07)	(64.40)
1/Share price	8.28***	8.18***	8.16***
	(13.09)	(12.97)	(12.96)
Log(turnover)	-0.07**	-0.08**	-0.08**
	(-2.47)	(-2.56)	(-2.62)
Log(return volatility)	-0.13***	-0.14***	-0.13***
	(-3.74)	(-3.80)	(-3.72)
NASDAQ	-0.67***	-0.65***	-0.65***
	(-13.93)	(-13.55)	(-13.53)
Year	Y	Y	Y
Industry	Y	Y	Y
Observation	14,690	14,690	14,690
Adjusted-R^2	0.397	0.398	0.398

The table presents the estimates of regressions relating Logarithm of number of shareholders to measures of investors attention. The number of common shareholders is downloaded from Compustat. SVI and Number of news are downloaded from Google trends and Google news respectively. The advertising expenditure, monthly share price and bid-ask spreads are collected from CRSP. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. Stock return is calculated as: Return= $(P_t-P_{t-1})/ P_{t-1}$. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured by logarithm of market value (share price multiplied by number of shares outstanding). Turnover is constructed as a monthly trading volume divided by number of shares outstanding. Return volatility is calculated as the standard deviation of monthly returns over the year. NASDAQ is a dummy variable taking 1 for stocks traded on NASDAQ, 0 other wise.

T-statistics are reported in parentheses below the coefficient estimates.
*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 6 Univariate analysis on the effect of investors attention on liquidity:

	Low-SVI	High-SVI	High-low difference	t-statistics
All stock	3.08	2.89	-0.19***	6.38
Small stock	3.44	3.27	-0.17***	3.43
Large stock	2.62	2.56	-0.06*	1.72

The table presents a comparison of means for relative spread between low and high SVI sub-samples. The sub-samples are constructed according to the sample mean of SVI level. Further, we partition the sample into small stock and larger stock groups according to market capitalization, then divide each group into low and high SVI subsample and compare the mean of relative spread across sub samples. Reported relative spread is based on equally weighted cross-sectional means. SVI is collected from Google trends, stock price and bid-ask spread from CRSP. The market value of equity is constructed as stock price multiplied by number of shares outstanding.

*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 7: Regression analysis: the effect of investors attention on bid-ask spread:

	Dependent Variable: Relative spread		
	(1)	(2)	(3)
SVI	-0.01**	-0.01**	-0.01***
	(-2.18)	(-2.17)	(-2.11)
Log (number of News)		0.01*	0.01*
		(1.71)	(1.80)
Log (advertising)			-0.003
			(-1.42)
Log(firm age)	-0.01	-0.01	-0.01
	(-0.73)	(-0.69)	(-0.70)
ROA	-0.01	-0.01	-0.01
	(-0.46)	(-0.41)	(-0.44)
Firm size	0.003	-0.001	0.001
	(0.39)	(-0.17)	(0.16)
1/Share price	2.95***	2.93***	2.94***
	(8.04)	(7.99)	(7.99)
Log(turnover)	0.12***	0.12***	0.12***
	(7.68)	(7.64)	(7.69)
Log(return volatility)	2.47***	2.47***	2.47***
	(84.99)	(84.95)	(84.86)
NASDAQ	-0.03	-0.03	-0.03
	(-1.28)	(-1.14)	(-1.14)
Year	Y	Y	Y
Industry	Y	Y	Y
Observation	14,690	14,690	14,690
Adjusted-R^2	0.793	0.793	0.793

The table presents the estimates of regressions relating relative spread to measures of investors attention. Relative spread is calculated as the quoted spread divided by the mid-point of the bid and ask prices. SVI and Number of news are downloaded from Google trends and Google news respectively. The advertising expenditure, monthly share price and bid-ask spreads are collected from CRSP. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured by logarithm of market value (share price multiplied by number of shares outstanding). Turnover is constructed as a monthly trading volume divided by number of shares outstanding. Return volatility is calculated as the standard deviation of monthly returns over the year. NASDAQ is a dummy variable taking 1 for stocks traded on NASDAQ, 0 other wise.

T-statistics are reported in parentheses below the coefficient estimates.
*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 8: Regression analysis: the effect of investors attention on turnover

	Dependent Variable: Log (turnover)		
	(1)	(2)	(3)
SVI	0.01***	0.01***	0.01***
	(3.10)	(3.14)	(2.77)
Log (number of News)		0.01***	0.01***
		(3.58)	(2.99)
Log (advertising)			0.01***
			(10.35)
Log (firm age)	-0.05***	-0.05***	-0.05***
	(-9.41)	(-9.35)	(-9.32)
ROA	-0.06***	-0.06***	-0.06***
	(-6.61)	(-6.49)	(-6.23)
Firm size	-0.08***	-0.08***	-0.09***
	(-17.03)	(-15.93)	(-17.45)
1/ Share price	-0.63***	-0.66***	-0.70***
	(-5.71)	(-5.85)	(-6.07)
Log (return Volatility)	0.66***	0.66***	0.66***
	(62.48)	(62.44)	(62.68)
NASDAQ	0.27***	0.28***	0.27***
	(20.61)	(21.12)	(21.03)
Year	Y	Y	Y
Industry	Y	Y	Y
Observation	14,690	14,690	14,690
Adjusted-R^2	0.483	0.484	0.487

The table presents the estimates of regressions relating Log (turnover) to measures of investors attention. Turnover is constructed as a monthly trading volume divided by number of shares outstanding. SVI and Number of news are downloaded from Google trends and Google news respectively. The advertising expenditure, monthly share price and bid-ask spreads are collected from CRSP. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured by logarithm of market value (share price multiplied by number of shares outstanding). Return volatility is calculated as the standard deviation of monthly returns over the year. NASDAQ is a dummy variable taking 1 for stocks traded on NASDAQ, 0 other wise.

T-statistics are reported in parentheses below the coefficient estimates.
*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 9: Regression analysis with alternative investors attention measure: the effect of investors attention on the breadth of ownership

	Dependent Variable: Log (Number of Shareholders)		
	(1)	(2)	(3)
SVI change	0.21**	0.21**	0.21**
	(1.93)	(1.96)	(1.96)
Log (number of News)		0.04***	0.04**
		(5.52)	(5.42)
Log (advertising)			0.01
			(1.64)
Log(firm age)	0.27***	0.27***	0.27***
	(12.15)	(12.20)	(12.20)
Return	-0.21	-0.21	-0.20
	(-1.64)	(-1.63)	(-1.60)
ROA	0.31***	0.31***	0.31***
	(8.59)	(8.70)	(8.73)
Firm size	0.90***	0.88***	0.88***
	(72.68)	(66.70)	(64.10)
1/Share price	8.24***	8.14***	8.12***
	(13.13)	(13.00)	(13.00)
Log(turnover)	-0.07**	-0.08**	-0.08**
	(-2.47)	(-2.57)	(-2.64)
Log(return volatility)	-0.14***	-0.14***	-0.14***
	(-3.83)	(-3.90)	(-3.81)
NASDAQ	-0.67***	-0.65***	-0.65***
	(-13.95)	(-13.56)	(-13.55)
Year	Y	Y	Y
Industry	Y	Y	Y
Observation	14,690	14,690	14,690
Adjusted-R^2	0.397	0.398	0.398

The table presents the estimates of regressions relating Logarithm of number of shareholders to measures of investors attention. The number of common shareholders is downloaded from Compustat. SVI and Number of news are downloaded from Google trends and Google news respectively. SVI change is calculated as $\Delta SVI_t = SVI_t - Med (SVI_{t-1}, ...SVI_{t-8})$. The advertising expenditure, monthly share price and bid-ask spreads are collected from CRSP. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. Stock return is calculated as: Return= $(P_t-P_{t-1})/ P_{t-1}$. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured by logarithm of market value (share price multiplied by number of shares outstanding). Turnover is constructed as a monthly trading volume divided by number of shares outstanding. Return volatility is calculated as the standard deviation of monthly returns over the year. NASDAQ is a dummy variable taking 1 for stocks traded on NASDAQ, 0 other wise.

T-statistics are reported in parentheses below the coefficient estimates.
*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 10: Regression analysis with alternative investors attention measure: the effect of investors attention on bid-ask spread

	Dependent Variable: Relative spread		
	(1)	(2)	(3)
SVI change	-0.24***	-0.24***	-0.24***
	(-4.30)	(-4.30)	(-4.30)
Log (number of News)		0.01*	0.01*
		(1.70)	(1.79)
Log (advertising)			-0.003
			(-1.42)
Log(firm age)	-0.01	-0.01	-0.01
	(-0.67)	(-0.63)	(-0.64)
ROA	-0.01	-0.01	-0.01
	(-0.39)	(-0.35)	(-0.39)
Firm size	0.003	-0.001	0.001
	(0.38)	(-0.19)	(0.16)
1/Share price	2.97***	2.95***	2.96***
	(8.11)	(8.05)	(8.06)
Log(turnover)	0.12***	0.12***	0.13***
	(7.76)	(7.71)	(7.77)
Log(return volatility)	2.48***	2.47***	2.47***
	(85.28)	(85.24)	(85.16)
NASDAQ	-0.03	-0.02	-0.02
	(-1.24)	(-1.11)	(-1.11)
Year	Y	Y	Y
Industry	Y	Y	Y
Observation	14,690	14,690	14,690
Adjusted-R^2	0.794	0.794	0.794

The table presents the estimates of regressions relating relative spread to measures of investors attention. Relative spread is calculated as the quoted spread divided by the mid-point of the bid and ask prices. SVI and Number of news are downloaded from Google trends and Google news respectively. SVI change is calculated as $\Delta SVI_t = SVI_t - Med\ (SVI_{t-1}, ...SVI_{t-8})$. The advertising expenditure, monthly share price and bid-ask spreads are collected from CRSP. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured by logarithm of market value (share price multiplied by number of shares outstanding). Turnover is constructed as a monthly trading volume divided by number of shares outstanding. Return volatility is calculated as the standard deviation of monthly returns over the year. NASDAQ is a dummy variable taking 1 for stocks traded on NASDAQ, 0 other wise.

T-statistics are reported in parentheses below the coefficient estimates.
*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 11: Regression analysis with alternative investors attention measure: the effect of investors attention on turnover

	Dependent variable: Log (turnover)		
	(1)	(2)	(3)
SVI change	0.12***	0.12***	0.12***
	(3.80)	(3.80)	(3.80)
Log (number of News)		0.01***	0.01***
		(3.59)	(2.99)
Log (advertising)			0.01***
			(10.44)
Log (firm age)	-0.06***	-0.05***	-0.05***
	(-9.47)	(-9.41)	(-9.37)
ROA	-0.06***	-0.06***	-0.06***
	(-6.71)	(-6.60)	(-6.31)
Firm size	-0.08***	-0.08***	-0.09***
	(-16.87)	(-15.79)	(-17.37)
1/ Share price	-0.65***	-0.68***	-0.71***
	(-5.87)	(-6.00)	(-6.21)
Log (return Volatility)	0.65***	0.65***	0.65***
	(62.04)	(62.01)	(62.25)
NASDAQ	0.27***	0.28***	0.27***
	(20.62)	(21.13)	(21.03)
Year	Y	Y	Y
Industry	Y	Y	Y
Observation	14,690	14,690	14,690
Adjusted-R^2	0.483	0.484	0.487

The table presents the estimates of regressions relating Log (turnover) to measures of investors attention. Turnover is constructed as a monthly trading volume divided by number of shares outstanding. SVI and Number of news are downloaded from Google trends and Google news respectively. SVI change is calculated as $\Delta SVI_t = SVI_t - Med\ (SVI_{t-1}, ...SVI_{t-8})$. The advertising expenditure, monthly share price and bid-ask spreads are collected from CRSP. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured by logarithm of market value (share price multiplied by number of shares outstanding). Return volatility is calculated as the standard deviation of monthly returns over the year. NASDAQ is a dummy variable taking 1 for stocks traded on NASDAQ, 0 other wise.

T-statistics are reported in parentheses below the coefficient estimates.
*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).

Table 12: Regression analysis with alternative liquidity measures

	Dependent Variables	
	Model (1) Effective spread	Model (2) Relative effective spread
SVI change	-0.11**	-0.01***
	(-2.02)	(-2.47)
Log (number of news)	-0.01***	-0.001***
	(-3.07)	(-3.87)
Log (advertising)	-0.02***	-0.001***
	(-4.97)	(-5.46)
Log (firm age)	0.05***	0.001***
	(3.87)	(4.74)
ROA	0.11***	0.01***
	(4.75)	(5.22)
Firm size	0.13***	0.01***
	(5.28)	(5.67)
1/Share price	0.62***	0.03***
	(3.19)	(3.91)
Log(turnover)	1.00***	0.05***
	(5.32)	(5.80)
Log(return volatility)	-0.68***	-0.03***
	(-5.39)	(-5.79)
NASDAQ	-0.25***	-0.01***
	(-5.21)	(-5.66)
Year	Y	Y
Industry	Y	Y
N	14,690	14,690
Adjusted-R^2	0.79	0.79

The table presents the estimates of regressions relating SVI/SVI change to alternative measures of liquidity. Effective spread is constructed as twice the difference of the transaction price and the mid-point of bid-ask spread. Relative effective spread is constructed as the effective spread scaled by the mid-point of the bid--ask spread. SVI and Number of news are downloaded from Google trends and Google news respectively. SVI change is calculated as $\Delta SVI_t = SVI_t - Med\ (SVI_{t-1}, ...SVI_{t-8})$. The advertising expenditure, monthly share price and bid-ask spreads are collected from CRSP. Firm age is calculated as the difference in years between the current date and the first date of stock data appeared in CRSP. ROA is calculated as operating income before depreciation scaled by total assets. Firm size is measured by logarithm of market value (share price multiplied by number of shares outstanding). Return volatility is calculated as the standard deviation of monthly returns over the year. NASDAQ is a dummy variable taking 1 for stocks traded on NASDAQ, 0 other wise.
T-statistics are reported in parentheses below the coefficient estimates.
*** Coefficient is significant at 1% level (2-tailed);
** Coefficient is significant at 5% level (2-tailed);
* Coefficient is significant at 10 % level (2-tailed).