UNIVERSITY OF GOTHENBURG

SCHOOL OF BUSINESS, ECONOMICS AND LAW

THE EFFECT OF ATTENTION ON THE BEHAVIOUR OF INVESTORS USING A SOCIAL TRADING PLATFORM, SHAREVILLE

Maria Olsson and Lisa Reenbom

Centre for Finance – CFF, Gothenburg School of Business, Economics and Law

Bachelor's thesis

Supervised by Jianhua Zhang

Acknowledgements

This thesis has been challenging in many ways but we have learnt a lot and hopefully this is will be of value for someone who is interested in the future research of Behavioural Finance. We have put in a lot of effort and time ourselves, but we are very grateful for all the support and guidelines along the way.

We will start with our supervisor, Jianhua Zhang, your guidance and expertise has taken us through many moments in doubt and your help has been of great importance.

Fabian Grapengiesser, Shareville and Nordnet whose support has made this thesis possible.

Our opponents and course mates have been valuable and supportive in several ways.

Thank you all for keeping us on track towards our goal with this bachelor thesis.

Abstract

When buying a stock, it is impossible to take hundreds or thousands of stocks into consideration. A way for investors to simplify the search problem is to make the choice from stocks that have caught their attention. Motivated by the theories of human cognitive boundaries affecting investor behaviour, this thesis investigates the impact of attention effects on the behaviour of investors using a social trading platform, Shareville. Using a novel dataset from Shareville, we test the causal relation between the order volume and different attention proxies; comments, comments on a Friday and comments' effect on buy orders. In addition, a sub sample with only the thirty largest and the thirty smallest Swedish firms is used. Our results indicate that order volume can be predicted by the number of comments on an asset, but that volume also has a positive and significant effect on the number of comments. Second, there is no evidence for that investors are more likely to show attention driven trading behaviour on a Friday. Third, we find that comments increase buy order volume more, compared to sell order volume. Fourth, the regressions containing firm size and profitability do not show an effect on order volume. We conclude that while there is a significant effect of comments on order volume, it is likely that our equations suffer from endogeneity due to reversed causality.

Keywords: Behavioural Finance, Irrational Investors, Attention, Financial Markets, Social Trading Platform

Contents

1. Introduction

1.1 Background

In the field of behavioural finance one theory is that humans are boundedly rational. There are constraints to how much information the brain can process. In order to make a decision, it is more convenient to make a choice out of a selection of ten alternatives rather than a thousand. A way to downsize the selection is to make a choice from alternatives that have caught our attention (Barber and Odean, 2008). Barber and Odean mean that stocks covered by the media, stocks with abnormal returns and stocks experiencing abnormal trading volume, grab the attention of investors. They confirm this by finding effects of attention driven buying on trading volume and stock prices. Barber and Odean (2008) also find that attention effects seem to have stronger impact on investors' buying behaviour than on their selling behaviour.

Other researchers have some proxies for attention such as a stock's market capitalization, profitability or analyst coverage. Some researchers have also hypothesised that on Fridays, investors are likely to be distracted by the upcoming weekend (DellaVigna and Pollet, 2009). DellaVigna and Pollet test their hypothesis by using a Friday dummy and their findings give support for their argument.

There is an increasing interest for how media might affect investors' trading behaviour amongst researchers and professionals. In the attention literature, researchers use different methods and forms of media as attention proxies, however there seem to be effects of economic significance irrespective of study type. Barber and Odean (2008) test and find that investors buy stocks exposed in economy related magazines. Engelberg *et al.* (2012) investigate effects of media exposure, in the form of a TV-show giving stock recommendations, on investor behaviour. They confirm that investors buy stocks recommended in the show and thus show attention driven buying behaviour. The digital evolution of media, have brought novel media proxies for attention. Mondria and Wu (2013) proxy attention with the number of times a stock ticker is googled per day, and find evidence for attention driven buying. Media, in the shape of social media platforms, has not been investigated to any extent by limited attention literature. The recent emerge of social media has led to an increasing selection of trading podcasts, Facebook investing threads and development of social trading platforms. This further motivates investigation of the role this novel media shape might have.

Compared to earlier research, this study uses exclusive data from the social trading platform Shareville, to test for attention driven buying behaviour amongst retail investors. Shareville gives us a unique data set as well as a new proxy for attention; comments, which is the number of times an asset is commented on the platform. The data also makes it possible to analyse the attention proxy comments in more depth, by investigating the case of reversed causality. The study's main contribution is the novel proxy comments, a reversed causality test of the new proxy and a unique data set. With the impact of social media on today's society, the idea of testing attention driven buying behaviour in the context of a trading forum is timely and interesting both for the investors and the companies themselves.

Shareville, at the time being the only social trading platform in Sweden, aims to facilitate investment choices for investors by enabling them to follow other investors' portfolios and discussions of their investments. The platform is owned and connected to Nordnet, a trading site where investors buy and sell tradeable assets. The idea of Shareville is that the user can chose to be anonymous or not, and Shareville does not register or provide any information about the value of the portfolio in terms of monetary size (Nordnet, 2016b). The portfolio is shown to other followers in percentages of how much the investor owns of each asset. There are 93 024 portfolios registered on the platform (Shareville, 2016a). Several of the most followed portfolios are well known professionals in the finance business (Shareville, 2016b). As a member of the network you can choose to follow successful investors, and get notified by email when they buy or sell a stock. Investors using Shareville can comment on Nordnet's tradeable assets. It is of importance to clarify that comments are not necessarily a buy or a sell recommendation, but it is whenever a stock is mentioned in a discussion thread, at a portfolio wall or as a comment accompanying an order of an asset. When an investor with a Shareville profile comments or mentions a certain asset, other investors are likely to pay attention to that and this might affect their trading behaviour.

1.2 Hypotheses

This study aims to investigate the impact of attention effects on the behaviour of investors using Shareville. By using econometric techniques, we estimate the causal relationship between different attention proxies and order volume. With this aim, we test four hypotheses.

The first hypothesis is that comments on assets should grab investors' attention and by that increase trading volume. We also expect that endogeneity due to reversed causality might be a problem,

hence we do a reversed causality test with comments as the dependent variable and order volume as the explanatory variable.

The second hypothesis is that investors should suffer more from cognitive constraints on Fridays and thus the effect of comments should have stronger attention grabbing effects on a Friday. We hypothesise that a comment on a Friday, therefore should increase order volume more than on other weekdays.

The third hypothesis is that attention effects from comments should have a larger effect, causing the volume to increase more, for buy orders than for sell orders.

The fourth hypothesis is that other proxies for attention; market capitalization and return on equity, also should have a positive effect on the traded volume. We also expect attention effects to be stronger for stocks with a small market capitalization.

The above hypotheses are tested by either running cross sectional or pooled regressions.

1.3 Delimitations and limitations

Due to the characteristics and uniqueness of the data set it is hard to use some econometric techniques, like instrumental variables or additional control variables. An example is that the variable comments is hard to instrument.

Shareville has existed for barely three years and the data used covers one and a half years, which gives that it can be hard to capture the true effects. The young age of Shareville is likely a reason why some variables used for the analysis have rather small magnitudes with distributions clustered close to zero, see Figure 1 and 3 in Appendix.

1.4 Structure

The remainder of the thesis starts with a theory section, section two, which is an introduction of the traditional concepts of finance and the efficient market theory. Section three presents reported results from existing literature and explains in more detail the motivation for the hypotheses we test. Section four describes the data and variables used in our equations. In section five, we describe the methodology and the different equations used in order to test our hypotheses. Section six provides the results of the hypotheses tests. Section seven gives a conclusion of the thesis.

2. Theoretical Framework

2.1 Theory

The efficient market hypothesis is a central concept for the traditional theories of finance. A general definition of the efficiency concept is "A market in which prices always "fully reflect" available information is called efficient" (Fama 1970, s. 383). Fama further defines the different forms of market efficiency as the weak, the semi-strong and the strong form. The semi-strong form means that the available information set, is all publicly available information. Strong form means that all information, including private and insider information, is in the information set. Testing the semi-strong form could for example be tested on whether a brokerage firm's recommendation is efficiently incorporated in the stock price or not. Fama (1970) argues that tests of the semi-strong form support the efficient market hypothesis but that the strong form should be viewed as a benchmark since research find that insider trading gives abnormal returns. Considering Fama's review we can assume that the semi-strong form is what best describes the markets today. The semi-strong form implies that second-hand information, like a stock recommendation based on already publicly available information should not enable abnormal returns. It also means that only when adding new information, an increase in trading volume will occur, due to investors' adjusting for the new price to the new intrinsic value. In an efficient market the assumption that security prices fully reflect all available information should invalidate the theory about media affecting abnormal returns, Fama (1970).

Evidence from research suggests that the topic of market efficiency might be more complex than assuming that prices efficiently reflect all available public information. The supporters of behavioural finance, have a growing amount of findings showing situations where the market is not efficient. Robert Shiller (2003) states that we have to accept that life is not as easy as assuming the efficient market theory always holds, and points out that even Fama in 1970 spoke about some anomalies of serial correlation of stock returns, even though he claimed they were too small to be of any significance.

2.2 Existing literature

2.2.1 Bounded rationality

There are many empirical studies of market efficiency which tests the semi-strong form of market efficiency. This concerns whether stock prices efficiently adjust to other information that is obviously publicly available. Bushee *et al.* (2009) investigate how media, as an information intermediary, affects the capital markets. An "information intermediary" is in this context an agent who provides new and useful information about a stock or a company. Bushee *et al.* (2009) test the role of media around earnings announcements, in the sense that media mitigate the information asymmetry around this announcement. The results show that the press as an information intermediary, fulfils multiple roles, including providing investors with new and relevant information about the company. To conclude, Bushee *et al.* (2009) states that the press has potential to influence the degree of information asymmetry across investors, and that greater press coverage during earnings announcements the more reduced will the bid-ask spreads be. They also state that the press provides more depth in the market.

When Barber *et al.* (2011) investigate the theory of market efficiency, they suggest that investors suffer more or less from overconfidence; they tend to be unrealistic about how high their returns are going to be which leads to ignoring information that might be of relevance. The results from their research are supported by the well-known concept of bounded rationality. This term is a contradiction to that humans make rational decisions, because of cognitive limitations and uncertain future predictions (Tseng, 2006). The argument about bounded rationality raises the question whether the financial markets are efficient or not. Traders, investors and other market participants are all exposed to different information and all suffer more or less from time as a scarce resource. If the argument about bounded rationality holds, the efficient market theory is violated.

2.2.2 Attention driven buying behaviour

Odean (1999) finds patterns in excessive trading showing that investors buy stocks experiencing abnormal performance. As explanation for this behaviour, Odean (1999) means there are thousands of stocks to trade and a way to downsize the selection is to make a choice from alternatives that have caught investors' attention. Investors are more likely to pay attention to stocks experiencing extreme returns and thus more likely to trade those. Barber and Odean (2008)

further investigate this theme and find evidence in line with Odean's (1999). Barber and Odean (2008) mean that the rational investor only has a limited number of hours to consider stocks and that attention is a scarce resource. When buying a stock, investors are faced with a search problem in the sense that there are thousands of stocks from which to choose. We are not able to rank all of these, thus we limit our choice set. Barber and Odean (2008) argue that the human cognitive capacity is a scarce resource and thus attention is limited and all available information cannot be processed.

Attention is hard to measure directly as Barber and Odean (2008) conclude with the quote "a direct measure would be to go back in time and, each day, question…investors… as to which stocks they thought about that day". Therefore, researchers use proxies for attention. Odean (1999) proposes that stocks showing abnormal returns and stocks covered by the media are likely to grab an investors attention and thus be proxies for attention. Barber and Odean (2008) investigate and find that stocks covered by media in terms of newspapers and stocks with extreme one day returns grab investors' attention.

The findings of Odean (1999) suggests that individual investors are more likely to buy stocks that are attention grabbing rather than sell. Other researchers have reported similar results (Barber and Odean, 2008), (Engelberg *et al.* 2012). Barber and Odean (2008) reason the stronger buy effect is due to that the search problem is more severe for buying than for selling a stock. Retail investors have limited possibilities to short sell and thus in general only sell stocks they already own, hence the asymmetric behaviour.

Studies investigating investor's trading behaviour have found proof of the disposition effect, that investors sell winners and keep losers. Odean's (1998) research on the topic show support of the disposition effect that Shefrin and Statman (1985) foretold. Barber and Odean (2008) find that there are larger differences between investors buying and selling on days with negative stock returns. They argue that the disposition effect could be a possible explanation for this pattern.

2.2.3 Other proxies for attention

Another approach to proxy for attention is the use of a dummy variable for Friday. DellaVigna and Pollet (2009) argue that if investors suffer from cognitive constraints, the investors' attention should be more scarce on Fridays since investors could be distracted by plans for the upcoming weekend. DellaVigna and Pollet (2009) mean that managers are likely to time bad earnings

announcement on days with low attention, such as a Friday, due to that investors are distracted. They find that Friday earnings announcements have 0.5 percent lower abnormal return and are 45 percent more likely to be a negative earnings surprise. DellaVigna and Pollet (2009) argue this is supportive of the attention hypothesis.

Firm size is another proxy for attention, that is used by Hong *et al.* (2000) when they test for slow information diffusion with momentum returns on stocks. The slower the information diffusion, the more profitable is the momentum trading. Their argument is that investors may have higher costs getting information about small stocks and thus information about those stocks get out slower amongst the investors. Then information diffusion amongst investors shall increase with size. Hong *et al.* (2000) also use analyst coverage as proxy for information flow, where low coverage stocks have slower information diffusion. They find that profitability of momentum strategies declines as market capitalization increases and the same result is found for increasing analyst coverage. Also their findings give evidence for that analyst coverage has largest marginal effect on small stocks.

Engelberg and Parsons (2011) investigate investors' home biasedness, meaning that local traders tend to pay more attention to stocks that have been mentioned in local media. Using local media as an information provider, they predict that local media has strong effects on local trading, after controlling for earnings announcements, investor and newspaper characteristics.

2.2.4 Media and novel attention proxies

Media is an information intermediary between firms and investors and might therefore affect their trading behaviour and provide us with new attention proxies. Several studies report different relations between media publications and reactions in the stock market. Seasholes and Wu (2007) show that individual investors tend to be net buyers of attention grabbing events. Their findings are in line with other research on media leading to over reaction of stock prices (Barber and Odean, 2008) (Engelberg *et al*., 2012). Peress (2008) finds that media coverage has a negative effect on the under reaction anomaly PEAD, and thus increases market efficiency. Whether it's insights of under reaction or over reaction, it proposes "a potentially important role for the media in shaping the behaviour of the stock market" Hong and Stein (2007, s. 118).

Today media, covering information about stocks, is not only in the form of newspapers or TV-shows but can be discussion forums for investors or social trading platforms. With media evolving

its new digital form provides us with new ways of investigating attention effects and new proxies for attention. The Google search engine is increasingly used to investigate attention effects. Da *et al.* (2011) find evidence of that google search volume for a ticker predicts stock returns. Mondria and Wu (2013) conduct similar study using Google search volume and find support for attention theories. Social media like Facebook, Twitter and social trading platforms have in contrast been less touched upon by attention researchers. However, other behavioural finance fields have increasingly started exploring this. Karabulut (2013) uses Facebook Gross National Happiness (GNH) to investigate how investor sentiment may affect stock prices. He finds that GNH predicts changes in stock prices as well as trading volume. By using textual analysis on the most used social media platforms in the US, comScore, Chen *et al.* (2014) test and find that written opinions about stocks there can predict stock returns and earnings surprises.

2.2.5 The difficulties in measuring attention

Due to the difficulties of measuring attention, testing attention hypotheses can be hard, and showing a causal relationship between trading volume or stock returns and attention is therefore difficult. Despite researchers using different proxies for investor attention to come around this problem, each proxy still has flaws. For example, the used proxy abnormal stock returns might result in more attention on that stock but more attention might cause extreme stock returns. This makes it hard to interpret the true impact of the attention effects. Clearly, most proxies tend to have the advantage of being simple, intuitive and having roots in causal evidence but they are not results from theoretic work on attention (Michaely, Rubin and Vedrashko 2013).

2.2.6 Implications on an aggregate level

Researchers have reported that attention effects do not only affect investors trading behaviour but do also leave traces on an aggregate level in trading volume or stock returns. When an event about a stock or a company hits the news, trading volume will most likely be greater than normal. Even though this is against the market's ability to incorporate all news, significant events will probably catch investor attention and cause an abnormal trading volume, Barber and Odean (2008). It is reasonable to think that cognitive constraints are most binding for retail investors thus they are more affected by attention effects. The findings by Barber and Odean (2008) and Engelberg *et al*. (2012) support this line of reasoning. Engelberg *et al*. (2012) show that the retail investors' excessive buying of attention grabbing stocks can be seen in abnormal volume effects. The analysis by Engelberg *et al*. (2012) is based on a TV-program about stocks in order to see the

effects on the stock price. They concluded that the more time spent on the stock in the program, the larger the price increase the day after the recommendation of the stock. This reflects the short-run behaviour of the stock, in the long-run the price went back to its original level. This is an effect of how media can provide misleading stock prices at the market level.

2.2.7 Critique of behavioural finance

The empirical findings that media causes abnormal returns on stocks have been criticized for having a vague alternative hypothesis postulating market inefficiency. This is vague because it does not focus on a specific alternative to market efficiency. The alternative should explain the range of results better. It should focus on the expected value of abnormal returns that generates deviations from zero in both directions depending on if the media exposure was negative or positive (Fama, 1998).

In market efficiency, long-term returns, and behavioural finance, there is an argument about whether the selection of events is random or not (Fama, 1998). This can be related to criticism about over reaction, that researchers are more likely to pick events that cause abnormal returns. The researchers are often content with the overreaction or underreaction and are willing to infer that both outcomes reject the hypothesis about market efficiency (Fama, 1998). Fama also argues that if a reasonable change in the method of estimating the abnormal return causes an anomaly to disappear, it may not be evidential enough. The doubts about these anomalies are results of replication and robustness checks that followed publication of the original studies. As a conclusion of Fama's paper about market efficiency, the theory of anomalies is subject to happen by chance and does not provide long term evidence for market inefficiency. Even if the sample is large, it will be interesting to know the average probability of an abnormal return among the market (Fama, 1998).

As a reply to Fama's theory about market efficiency, Robert Shiller publishes in his article from 2003, that market efficiency can be extremely wrong in some senses. He gives the example that efficient market theory may lead to drastically incorrect interpretations of major events such as market bubbles. Schiller means that the arguments about that anomalies tend to disappear as time passes, are weak. In all disciplines, initial claims of important discoveries are often taken down by later research. Also, he means that the fact that markets adjust to normality after an over- or under reaction, is no evidence that markets are fully rational.

3. Data

3.1 Descriptives of the data

Shareville is a social media platform that is connected to Nordnet. The data is from comments and trades through the Shareville platform, on the tradeable assets on Nordnet. Additional information on return of equity and market capitalisation for certain firms was retrieved using the Orbis database (Orbis 2016). The tradeable assets are assets investors can trade on Nasdaq Stockholm, First North, Oslo stock exchange, Copenhagen stock exchange, Helsinki stock exchange, Aktietorget, NGM, Nordic MTF, Nasdaq, NYSE Pink sheet, Bulletin Board, Toronto stock exchange, Xetra, Euronext and London Stock Exchange (Nordnet 2016a). Assets are from the countries Sweden, Norway, Denmark, Finland, Europe, Canada and USA. Assets incorporates stocks, funds, certificates, options and futures. Shareville's investors are primarily from Sweden, Norway, Denmark and Finland (Nordnet 2016c). Most of Nordnet's closed trades 2015, are done by investors located in Sweden, 9.7 million (Nordnet 2016c). Second and third most closed trades are done in Denmark and Finland with 4.0 million and 3.6 million closed trades respectively. Least closed trades in 2015 have Norwegian investors at Nordnet with 2.5 million.

Observations from the period 2014-03-23 until 2014-10-01 are excluded due to that there was no registered order volume from trade through Shareville in that period. This leaves us with a data set with observations from 2014-10-01 till 2016-03-22. Observations where an instrument is traded but not commented on that day has been excluded. This leaves us with 745 381 observations.

The data has been adjusted for outliers using the Grubbs test with a confidence level of 99. The Grubbs test excluded 25,514 number of observations which is 3.42% of the total observations. The dataset used for the analysis then contains 719,867 observations[1].

---

[1] Which is 539 trading days and 21 854 traded assets. The total number of comments in the data set is 1.981 million.

3.2 Variable presentation

Section four continuous with an introduction and further description of the variables used in the regressions. This is followed by descriptive statistics. A correlation table as well as distribution tables are provided in Appendix.

Table 1: Description of variables
The table presents the variables used in our regressions.

VARIABLE NAME	DESCRIPTION
Volume	The daily, closing, buy/sell order volume.
Comments	The daily, number of comments on Shareville.
Comments*DummyFriday	Interaction variable modelling the effect of comments on a Friday.
DummyFriday	Models difference between Fridays and all other weekdays.
Comments*DummyBuy	Interaction variable modelling the effect of a comment on buy order volume.
DummyBuy	Models difference between buy and sell orders.
MC	The average market capitalization for a firm in the year of 2014.
ROE	Return on equity, for 2014, calculated using the net income.
Comments*DummySmallcap	Interaction variable modelling the effect of a comment on a small cap firm.
DummySmallcap	Models the difference between small cap and large cap firms.

In order to validate our attention proxies and test whether there are attention effects present amongst the investors using Shareville, the variable volume is used as dependent variable. Since this variable is unknown in terms of monetary size, it is a nominal variable, which means it can be used only as a measure of interest. The daily order volume is divided in buy order volume and sell order volume whereas the comments variable is the total number of comments for a stock per day. The distribution of volume is positively skewed, see Appendix Figure 1 and Figure 2. When we exclude outliers from the data set the skewness is reduced.

We use the variable comments as a proxy for the interest of one particular asset. The variable comments is our main proxy for attention and it is the daily number of times a certain asset, that is tradeable through Nordnet, is mentioned on Shareville. It is important to highlight that comments are not necessarily buy or sell recommendations, but are whenever an asset is mentioned by an investor on the Shareville platform. When a tradeable asset is commented, an investor is likely to pay attention to this asset and therefore it should be a measure of attention. The distribution of the variable comments is positively skewed. When we exclude outliers from the data set the skewness is reduced, see Appendix Figure 3 and Figure 4.

With inspiration from DellaVigna and Pollet (2009) an interaction term is used to measure the effect of comments made on a Friday. This captures time varying effects of attention.

To test the attention proxy firm size, as used by Hong *et al.* (2000), the variable market capitalisation is used as a proxy. The variable return on equity is tested as attention proxy, since investors are likely to pay attention to a firm showing high return on equity. With inspiration from the findings of Hong *et al.* (2000), that analyst coverage has a larger marginal effect on small stocks, an interaction term of comments and the small cap dummy is used to capture this effect. Due to that our data set contains not only firms, but also funds and derivatives we test the Market cap variable, ROE and attention effect on small cap by using data from a sub sample of sixty Swedish firms listed on the Swedish markets. This sixty firm sub sample consists of the thirty firms with the largest market capitalization and the thirty firms with the smallest market capitalization, on average during the last three years. The choice of subset is motivated by previous research on home bias and by that most of Nordnet's customers are located in Sweden (Nordnet, 2016c). Correlation amongst the regressors comments, ROE and MC is low, see Table 6 in Appendix.

Table 2: Descriptives of the variables

VARIABLES	N	mean	sd	min	max
Volume	719,867	2.721	3.244	0	22
Comments	719,867	5.075	6.614	1	45
DummyBuy	707,419	0.525	0.499	0	1
Comments*DummyBuy	707,419	2.576	5.170	0	45
DummyFriday	719,867	0.190	0.392	0	1
Comments*DummyFriday	719,867	0.963	3.502	0	45
DummySmallcap	719,867	0.005	0.069	0	1
Comments*DummySmallcap	719,867	0.017	0.430	0	43
MC	719,867	2,683.000	22,206.000	5	466,539
ROE	719,867	0.688	24.980	-327	941

4. Methodology

The method is set in order to test our hypotheses and this is primarily to assess the causal relation between our proxies for attention and the closing order volume. Specifically, we investigate how attention effects can alter investor behaviour and thus be visible as an effect on the closing order volume.

In order to test the causality, we run two type of regressions; cross sectional and pooled. The cross sectional regressions are run by using a data set that only varies by asset, not over time. The pooled regressions are run by using a data set that consists of an asset's average, over time, of each variable.

For the pooled regressions we use panel data, data that varies by asset and by time. Due to the nature of our data the panel data is unbalanced, since some assets are traded on more dates than others.

To test our first hypothesis, we run regressions on equation 1a and 1b:

$$Volume_{it} = \alpha + \beta_1 Comments_{it} + \varepsilon_{it} \qquad \text{(Eq. 1a)}$$

$$Comments_{it} = \theta + \gamma_1 Volume_{it} + \varepsilon_{it} \qquad \text{(Eq. 1b)}$$

where i is instrument and t is time. Equation 1a is used to tests that comments on assets should grab investors' attention and should increase volume by estimating how comments impact closing order volume. To investigate the expected case of reversed causality being present in estimations of equation 1a, we use equation 1b. Equation 1b estimates how comments might be affected by volume. We test both equation 1a and 1b by running both cross sectional and pooled regressions.

To test the second and third hypothesis we run a pooled regression on equation 2:

$$Volume_{it} = \alpha + \beta_1 Comments_{it} + \beta_2 Comments_{it} * DummyFriday +$$

$$+\beta_3 Comments_{it} * DummyBuy + \beta_4 DummyFriday + \beta_5 DummyBuy + \varepsilon_{it} \quad \text{(Eq. 2)}$$

where t is day.

When testing the fourth hypothesis we use data from the sub sample consisting of the 30 largest and the 30 smallest Swedish firms, with respect to three-year average market capitalization, that is

tradeable on Shareville. To test the fourth hypothesis, we use the sub sample to run a pooled regression on equation 3:

$$Volume_{it} = \alpha + \beta_1 Comments_{it} + \beta_2 MC + \beta_3 ROE +$$

$$+\beta_4 Comments_{it} * DummySmallcap + \beta_5 DummySmallcap + \varepsilon_{it} \qquad \text{(Eq. 3)}$$

Another widely used method in order to investigate the effect of an event on asset prices or volume is to do an event study. This is used when measuring the effect of how company policy changes affect the stock performance. The main limitation is that precise estimation periods are not easy to determine (Brown and Warner, 1984). In the case of this study with Shareville, the trade-off between the number of days used in an event study and the potential risk of another parameter affecting the closing order volume is another reason to why we chose not to use this type of study in order to analyse our data.

5. Results and Analysis

In the following, we will present the results in detail, in order of the four hypothesis. The results are analysed and commented on in relation to the theoretical framework.

Table 3: Regressions of eq. 1a and eq. 1b
This table shows the results from testing the first hypothesis by running both a cross sectional regression and a pooled regression on eq. 1a and eq. 1b.

	(Eq. 1a) Cross sectional	(Eq. 1b) Cross sectional	(Eq. 1a) Pooled	(Eq. 1b) Pooled
	Volume	Comments	Volume	Comments
Comments	0.454***		0.429***	
	(0.002)		(0.001)	
Volume		2.141***		1.782***
		(0.005)		(0.002)
Constant	0.471***	-0.938***	0.546***	0.228***
	(0.003)	(0.008)	(0.003)	(0.005)
Observations	21,854	21,854	719,867	719,867
R-squared	0.972	0.972	0.764	0.764
Adj. R-squared	0.972	0.972	0.764	0.764

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

The results of the cross sectional and pooled regression of equation 1a suggests that comments on assets grab investors' attention and increase trading volume. The cross sectional regression of equation 1a has a positive and statistically significant coefficient implying that one more comment will increase the closing order volume with 0.454 units on average. Also the pooled regression on equation 1a shows a statistically significant, positive effect on the daily closing order volume. Due to characteristics of the data set, that Shareville is a novelty, it's hard to tell if the results are economically significant. Economic significance is in terms of whether the estimated effects are large enough to have an economic impact that actually matters on the dependent variable or not. When putting our estimations in relation to other research, it's not obvious if our results are economically significant due to different methodology. However, our method and results are similar to the study of Engelberg *et al*. (2012) and they argue their findings should be considered economically significant. Thus, the result indicate that volume can be predicted by the social media coverage comments.

Our result is in line with the findings of Barber and Odean (2008) showing that investors purchase stocks that have a first impact on their attention. They suggest that if an unusual high number of investors trade a stock after an event, it is related to the investors paying attention to that stock. This implies that the event causes the investors' attention and also affects their trading behaviour. Engelberg and Parsons (2011) have related research outcome, they conclude that local media coverage on specific firm events can predict the interest of local trading. The coefficient of comments, in the estimation of equation 1a, is lower in the pooled regression compared to the cross sectional which has no obvious explanation. A possible explanation could be that the positive skew of the variables is more prominent in the pooled regression, which would lower the magnitude of the pooled regression. In the pooled cases the R-squared is lower, which can be due to increased total variability.

The dataset used makes it possible to address the likely endogeneity problem in terms of reversed causality; do the number of comments on an asset cause an increase in volume or does the volume cause more comments? It's rational to hypothesise that trading volume on an asset may affect comments. This reversed causality problem is discussed in the similar study of Engelberg and Parsons (2011), where local media might reflect the behaviour of the investors more than they are affected by media coverage. To investigate the case of reversed causality, we run the regressions with closing order volume as the explanatory variable in equation 1b (Bell and Bryman 2011). The results from cross sectional and pooled regressions on equation 1b tests and show that there is a reversed causality since volume has a positive effect on comments, that is statistically significant and possibly of economic significance. The results suggest that the estimation of equation 1a suffers from endogeneity in shape of reversed causality. The high correlation between volume and comments is in line with the reversed causality, see Table 6 in Appendix. Similar to the estimation of equation 1a, the estimation of 1b is of smaller magnitude when running the pooled regression. It seems the effects are somewhat different when letting the variables also vary over time. However, it's hard to argue in favour for a particular explanation.

From the results of the cross sectional and pooled regression of equation 1a and 1b we cannot reject our first hypothesis. Moreover, the results from the regressions of 1a and 1b showing dual causality further motivates investigation of comments as attention proxy. Further investigation with additional variables is one approach. Also results from conducting Ramsey RESET tests on the pooled equation 1a motivates adding more variables and continuing to assume a linear relation between comments and volume.

The results presented in Table 3 are robust for tests with data including outliers, see Appendix Table 3.1. Worth noting is also that although the variables volume and comments are positively skewed, the residuals of estimating equation 1a can be seen as normally distributed, see Appendix Figure 7. The residuals can be seen as a proxy for the error term, which is assumed to be normally distributed in order for OLS to hold.

Table 4: Regression of eq. 2
This table shows the results from testing the second and the third hypothesis by running a pooled regression of equation 2.

	(Eq. 2) Volume
Comments	0.364***
	(0.001)
Comments*dummyFriday	-0.003
	(0.002)
Comments*dummyBuy	0.139***
	(0.001)
DummyFriday	0.014**
	(0.006)
DummyBuy	-0.155***
	(0.005)
Constant	0.621***
	(0.004)
Observations	707,419
R-squared	0.791
F-statistic	80 816.500
Prob. > F	0.000
Adj. R-squared	0.791

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results from the pooled regression of equation 2 reject the second hypothesis, that is due to attention being more scarce on Fridays, investors shall show stronger attention driven trading on Fridays than on other weekdays. The coefficient of comments*dummyFriday estimates that a comment on a Friday decreases the trading volume. One possible explanation is DellaVigna's and Pollet's (2009) argument that managers are likely to time bad earnings announcement on days with

low attention, such as a Friday, due to that investors are distracted. Then one might think that comments on a Friday reduce volume because the comments may be on negative news and the disposition effect makes investors to keep losers. However, due to the insignificance of the coefficient for comments*dummyFriday, the negative effect cannot be said to differ from zero. When DellaVigna and Pollet (2009) discuss their findings they argue that another possible explanation is that the firms releasing news on Fridays may have different characteristics compared to other firms. In our case it's possible that the investors commenting on a Friday does not have as many followers, compared to more popular investors, and therefore the comments on a Friday has no significant effect on volume. According to the findings of DellaVigna and Pollet (2009) it is highly uniquely that the different firm characteristics should affect both earnings announcements and the following drift at the same time. However, in our case, it's reasonable to think there may be a difference in characteristics between investors commenting on a Friday and other investors, that can cluster on a Friday. The argument would be that investors, in contrast to firms, have a more pronounced difference in behaviour on certain weekdays. Amongst the most popular portfolios, several of them claim to be well known professionals (Shareville 2016b). Then one might think the well-known investors end their trading week when they end their work week and thus they might not be the ones commenting assets on Fridays.

The dummy Friday shows that trading volume is significantly higher on Fridays compared to other weekdays. Barber and Odean (2008) comment that more sell limit orders execute on a day when the market is rising. According to the common debate, on Fridays, returns tend to be higher and thus the volume would be higher. Our data does not distinguish between what is market orders and limit orders, which makes it hard to analyse further.

The third hypothesis, that attention effects shall increase buy order volume more than sell order volume, cannot be rejected based on the results from regression of equation 2. The results also suggest that the buy order volume on average is significantly lower than sell order volume. Again the reasoning of Barber and Odean (2008) about the disposition effect may be a possible explanation for the difference between buying and selling behaviour. During Shareville's lifetime one could argue there has been longer upmarket periods than downmarket. This may give that investors sell their winners which gives that the sell order volume is estimated to be relatively higher.

In the pooled regression of equation 2, the R-squared has increased compared to the pooled regression of 1a, which implies that the additional explanatory variables in equation 2 better explain the variation in the dependent variable volume. R-squared can increase due to additional variables rather than the additional variables explanatory power. Therefore, adjusted R-squared is interesting, since it only increases if the additional variables explain more than what can be predicted by chance. For equation 2 the adjusted R-squared has increased, compared to the pooled regression of equation 1b, which indicates that the additional variables are beneficial.

The F-statistic shows that our dependent variables are jointly significant and that at least one of our estimated coefficients are different from zero. The F-statistic then indicates that equation 2 can be used to predict volume in some sense.

Table 5: Regression of eq. 3
This table shows the results from testing the fourth hypothesis with a pooled regression equation 3.

	(Eq. 3) Volume
Comments	0.440***
	(0.004)
MC	-0.000
	(0.000)
ROE	-0.000*
	(0.000)
Comments*dummySmallcap	0.008
	(0.011)
DummySmallcap	-0.114***
	(0.044)
Constant	0.603***
	(0.033)
Observations	20,901
R-squared	0.747
F-statistic	6482.640
Prob. > F	0.000
Adj. R-squared	0.747

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

The fourth hypothesis, that other proxies for attention shall increase order volume, is rejected when tested by estimating equation 3. The effect of market capitalization has the opposite sign compared to expected and it's insignificant. The estimated effect of return on equity is the opposite and it's statistically significant at a 10 percent level. However, it should be considered as economically insignificant. The volume effect of a comment on a small cap stock is positive as hypothesized but it is both statistically and economically insignificant. The closing order volume is significantly lower for small cap stocks. Although our primary attention proxy using a stock size variable is insignificant, the significant coefficient for small cap stock suggests there might be some support for that size is a proxy for attention. The argument would then be that the stocks considered to be in the category of largest stocks get relatively more attention and thus they are traded more.

In the pooled regression of equation 3, the R-squared has decreased compared to the pooled regression of 1a, which implies that the additional explanatory variables in equation 3 does not improve the explanation of the variation in the dependent variable volume. The adjusted R-squared for equation 3 is lower compared to the pooled regression of equation 1a. The F-statistic for equation 3 has also decreased compared to the F-statistic of equation 1a. However, the F-statistic still shows that our dependent variables are jointly significant and that at least one of our estimated coefficients are different from zero. This can be seen as an indication that we can predict volume in some sense with equation 3.

6. Conclusion

We have tested the attention hypothesis, that investors trade assets that grabs their attention. The test was done by running cross sectional and pooled regressions on equations with different proxies for attention. A unique data set from the social trading platform Shareville was used. The attention hypothesis has been tested by primarily using the novel attention proxy; comments on assets on Shareville. The attention hypothesis could not be rejected, in the sense that comments on tradeable assets significantly increase closing order volume, which is in line with findings from related research. In contrast to most of the other research we also tested the expected case of reversed causality. The outcome show that volume significantly causes comments to increase. This suggest reversed causality is likely present when estimating the effect of comments on volume. The attention hypothesis was also tested by adding more variables to investigate it from more perspectives. We find that investors are not more prone to attention driven trading on Fridays and we also find that attention driven trading is stronger for buying than for selling as Barber and Odean (2008) argue. Our findings show no support for that attention effects are stronger for small capitalization firms, but it suggests that there is a significantly lower trading volume for the 30 smallest stocks compared to the 30 largest stocks traded in Sweden.

The results are supportive for theories of attention driven buying behaviour amongst retail investors. In other research, the attention driven buying seem to be costly for the retail investors. For institutional investors the attention driven buying behaviour amongst retail investors has been shown to predict stock returns. For stock brokers this implicates that attention driven behaviour should generate more brokerage on buy orders. However, due to the test of the reversed causality showing this is a problem, our results primarily motivates further research on presence of attention effects in the context of social trading platforms.

Previous research has reported attention effects leaving traces on an aggregate level in terms of stock returns or turnover volume. The limitations of the data set used in this study, makes it impossible to investigate any effects on returns due to that Shareville is very small relative to the stock market as a whole. But the increasing popularity of social media and social trading platforms, like Shareville, makes it an interesting research area. Thus an idea for further research is to use similar data but use other econometrics techniques such as lagged variables, VAR analysis or use other proxies for attention. Then it might be possible to mitigate the problem of endogeneity and reversed causality. Also, combining several different methods testing the same proxy should be

interesting since other researchers use event studies on exogenous shocks and treatment groups to minimize the endogeneity problem[2]. Furthermore, an interesting aspect would be to see if the investors that do buy attention grabbing assets, benefit from picking them or not.

---

[2] See Engelberg and Parsons (2011) and Shive (2012) who compare a treatment group's behaviour to a control group that has not been exposed to the treatment.

References

Barber, Brad M., and Terrance Odean, 2011, Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly Journal of Economics* 116, 261-292

Barber, and Terrance Odean, 2008, All That Glitters: The Effect of Attention and News on the Buying Behaviour of Individual and Institutional Investors, *The Review of Financial Studies* 21(2), 785-818.

Bell E., and Bryman A., 2011, Företagsekonomiska forskningsmetoder, upplaga 2:1, 334-336

Brown J. S., and Warner B. J., 1984, Using Daily Stock Returns, the Case of Event Studies, *Journal of Financial Economics,* 14(1985), 3-31

Bushee, Core J., Wayne Guay, and Sophia Hamm, 2009, The Role of the Press as an Information Intermediary, *Journal of Accounting Research* 48(1), 1-19

Campbell CJ. and C.E Wasley, 1996, Measuring Abnormal Daily Trading Volume for samples of NYSE/ASE and NASDAQ Securities Using Parametric and Nonparametric Test Statistics., *Review of Quantitative Finance and Accounting*, 6 (1996), 309-326

Chen, H., P. De, Y. Hu, and B.-H. Hwang, 2014, Wisdom of crowds: The value of stock opinions transmitted through social media, *Review of Financial Studies* 27, 1367–1403.

Da, Z., J. Engelberg, and P. Gao, 2011, In search of attention, *Journal of Finance*, 66, 1461–1499.

DellaVigna, J. M. Pollet, 2009, Investor inattention and Friday earnings announcements, *Journal of Finance*, 64, 709–749.

Engelberg E. J. and C.A. Parsons, 2011, The causal impact of media in financial markets, *Journal of Finance* 66, 67-97.

Engelberg J., C. Sasseville and J. Williams, 2012, Market Madness? The Case of Mad Money, *Management Science* 58(2), 351-364.

Fama, E., 1970, Efficient capital markets: A review of theory and empirical evidence, *Journal of Finance* 25(2), 383-517.

Fama, E., 1998, Market efficiency, long-term returns, and behavioural finance, *Journal of Financial Economics* 49(1998), 283-306.

Fang, L., and J. Peress, 2009, Media coverage and the cross-section of stock returns, *Journal of Finance*, 64, 2023–2052.

Hong, H. G., T. Lim, and J. C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance*, 55, 265–295.

Hong H., J. C. Stein, 2007, Disagreement and the Stock Market, *The Journal of Economic Perspectives* 21(2), 109-128.

Karabulut, Y., 2013, Can Facebook predict stock market activity?, Unpublished working paper, Goethe University Frankfurt.

Michaely, R., A. Rubin, and A. Vedrashko, 2013, Firm heterogeneity and investor inattention to Friday earnings announcements, Unpublished working paper, Cornell University, Simon Fraser University.

Mondria J., and T. Wu, 2013, Asymmetric attention and stock returns, Unpublished working paper, University of Toronto, University of California.

Odean T., 1998, Are investors reluctant to realize losses? *Journal of Finance*, 53:1775-79

Odean T., 1999, Do investors trade too much? *American Economic Review*, 1279–98.

Peress J., 2008, Media coverage and investors' attention to earnings announcements, *Unpublished working paper, Insead.*

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

Shiller R., 2003, From Efficient Markets Theory to Behavioral Finance, *Journal of Economic perspectives* 17(1), 83-104

Shive S. A., 2012, Local investors, price discovery and market efficiency, *Journal of Financial Economics,* 104, 145-161

Shefrin H. and M. Statman, 1985, The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance*, 40:777-90

Tseng K. C., 2006, Behavioural Finance, Bounded Rationality, Neuro-Finance, and Traditional Finance, *Investment Management and Financial Innovations,* Volume 3, Issue 4

Internet references

Orbis, 2016, *available at* https://orbis-bvdinfo-com.ezproxy.ub.gu.se/version-2016614/home.serv?product=orbisneo (last visited April 19 2016)

Nordnet, 2016a, Alla våra priser och räntor, *available at* https://www.nordnet.se/mux/web/nordnet/pricelist.html#/categories/retail (last visited May 20 2016)

Nordnet, 2016b, Shareville, *available at* https://www.nordnet.se/tjanster/investeringshjalp/shareville.html (last visited May 20)

Nordnet, 2016c, Årsredovisning 2016, *available at* http://org.nordnet.se/shared/stories/reports/2015/annual_report_nordnet_2015.pdf (last visited April 25 2016)
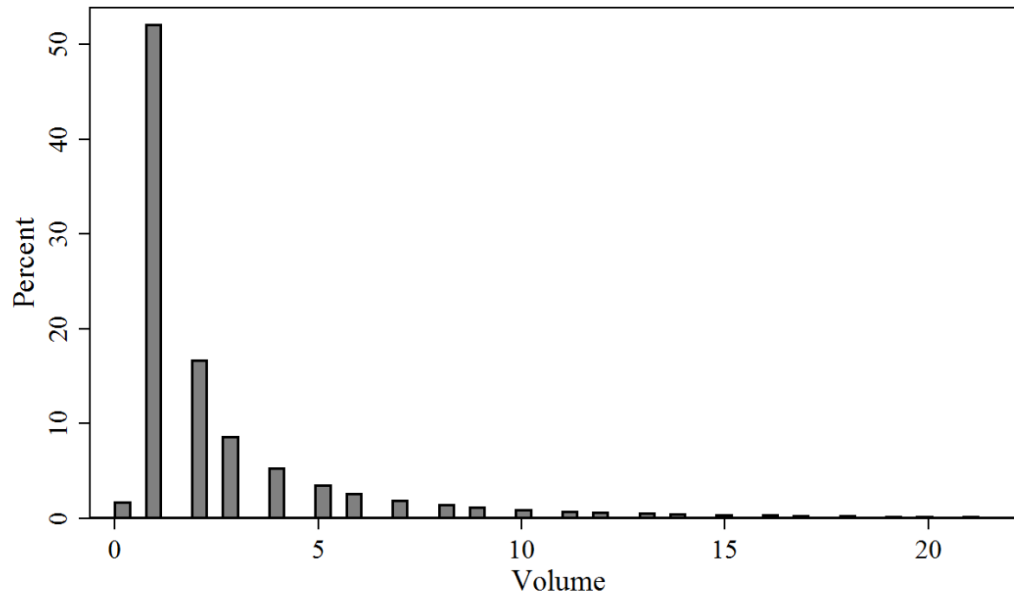
Shareville, 2016a, Shareville, *available at* https://www.shareville.se/ (last visited May 23 2016)

Shareville, 2016b, Mest populära, *available at* https://www.shareville.se/medlemmar/popular (last visited May 23 2016)
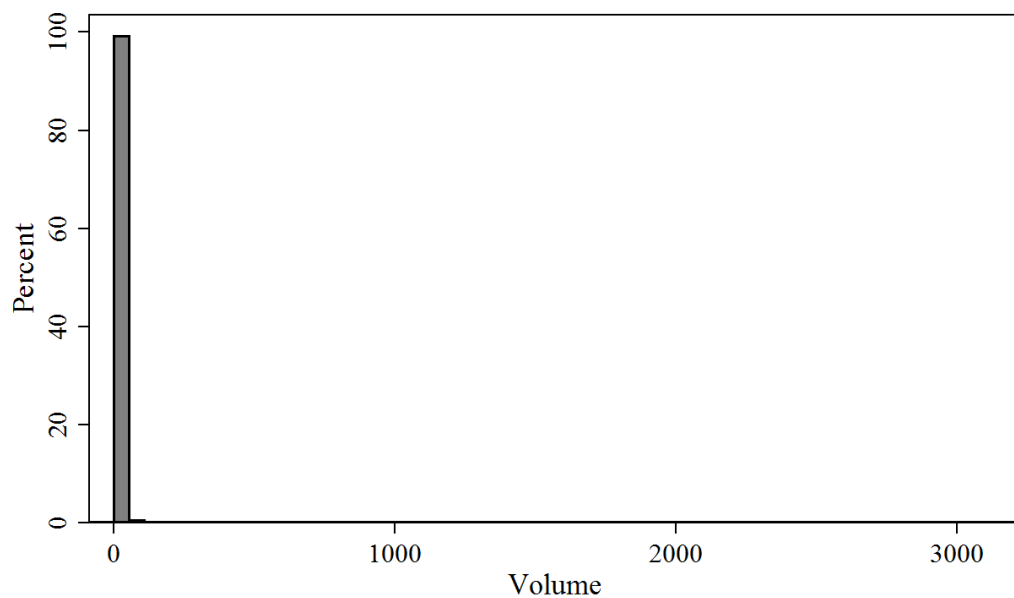
Appendix

Figure 1

This figure shows a histogram of Volume.

The value of skewness is 2.86 and the value of kurtosis is 12.30.
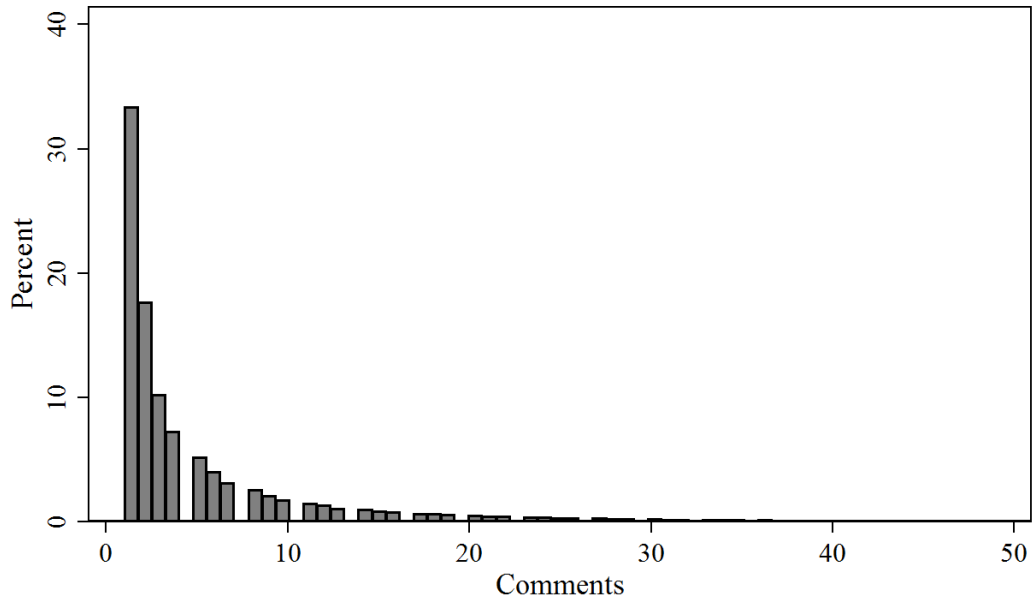
Figure 2

This figure shows a histogram of Volume using data including outliers.

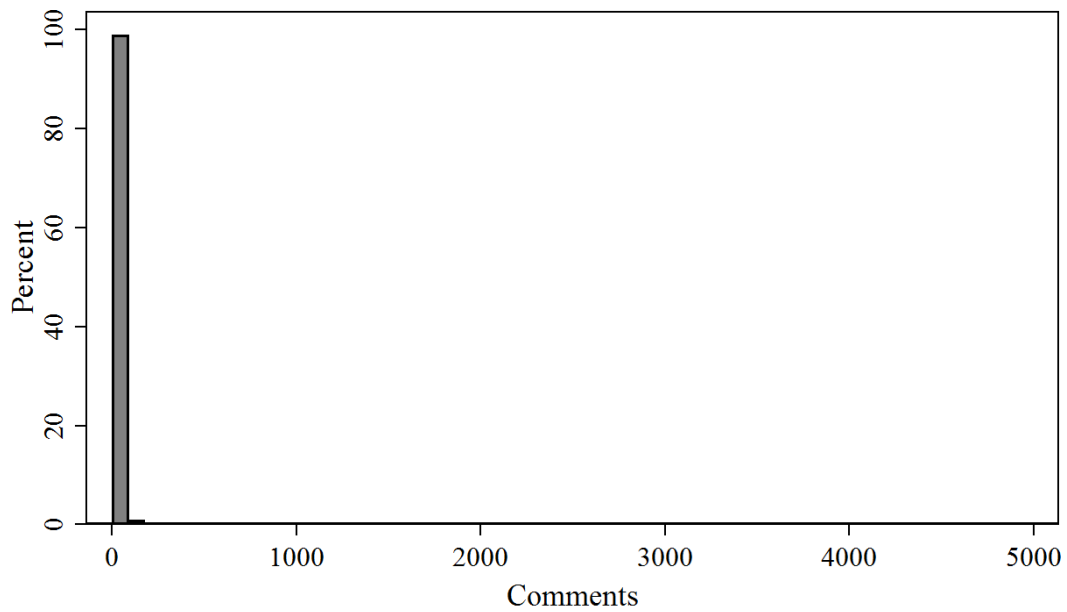The value of skewness is 34.57 and the value of kurtosis is 3347.59.

Figure 3

This figure shows a histogram of Comments.

The value of skewness is 2.73 and the value of kurtosis is 11.58.

Figure 4

This figure shows a histogram of Comments using data including outliers.

The value of skewness is 33.65 and the value of kurtosis is 2897.94.

Figure 5

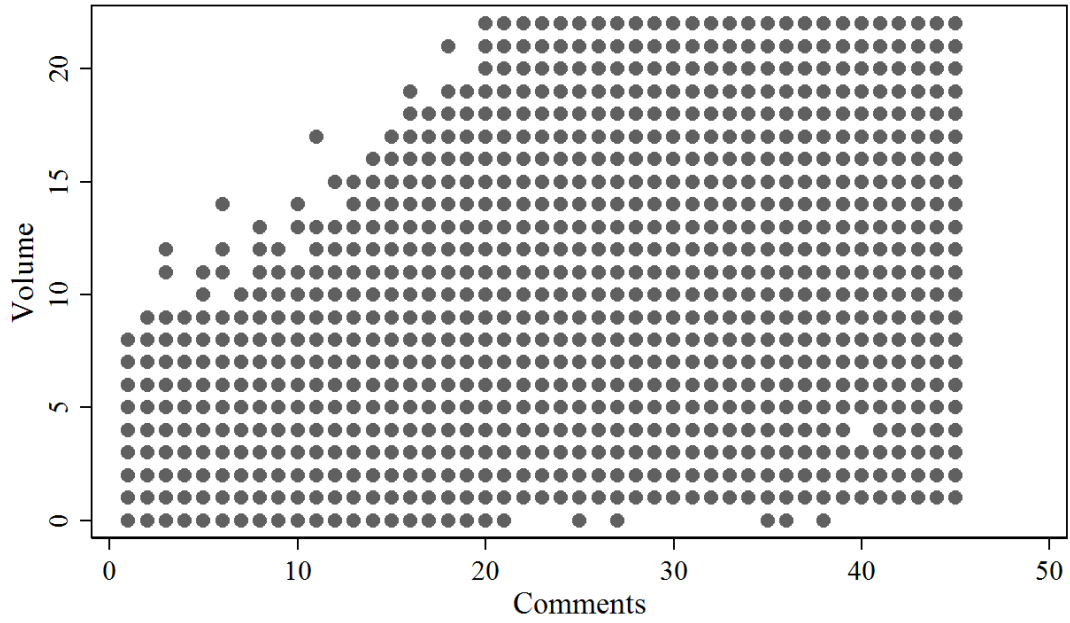This figure shows a scatter plot of Volume and Comments.

Figure 6

This figure shows a scatter plot of Volume and Comments using data with outliers.
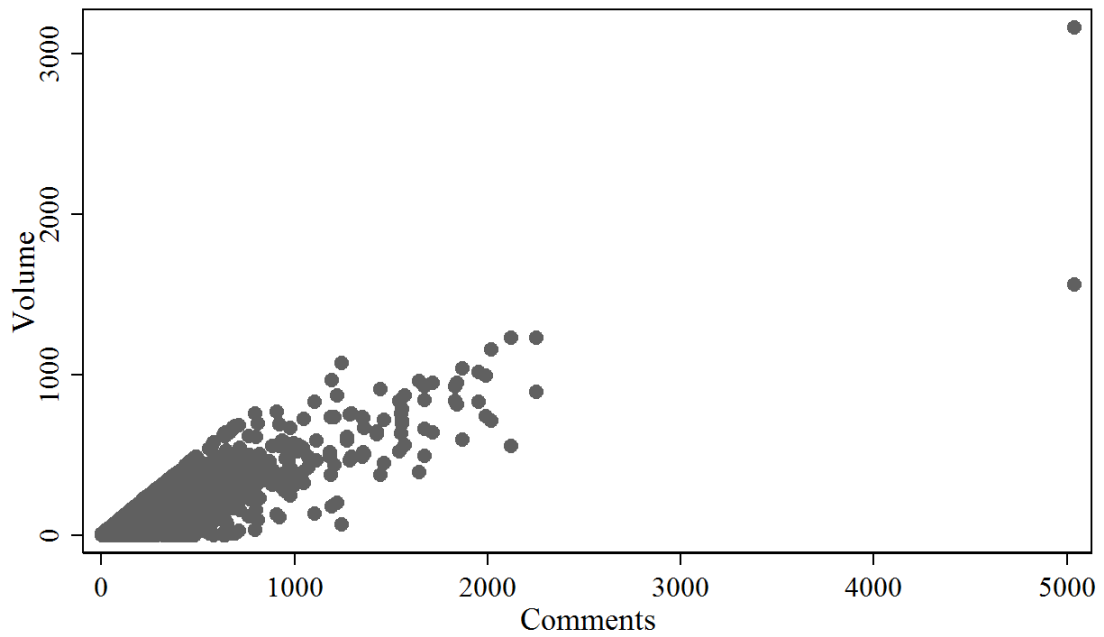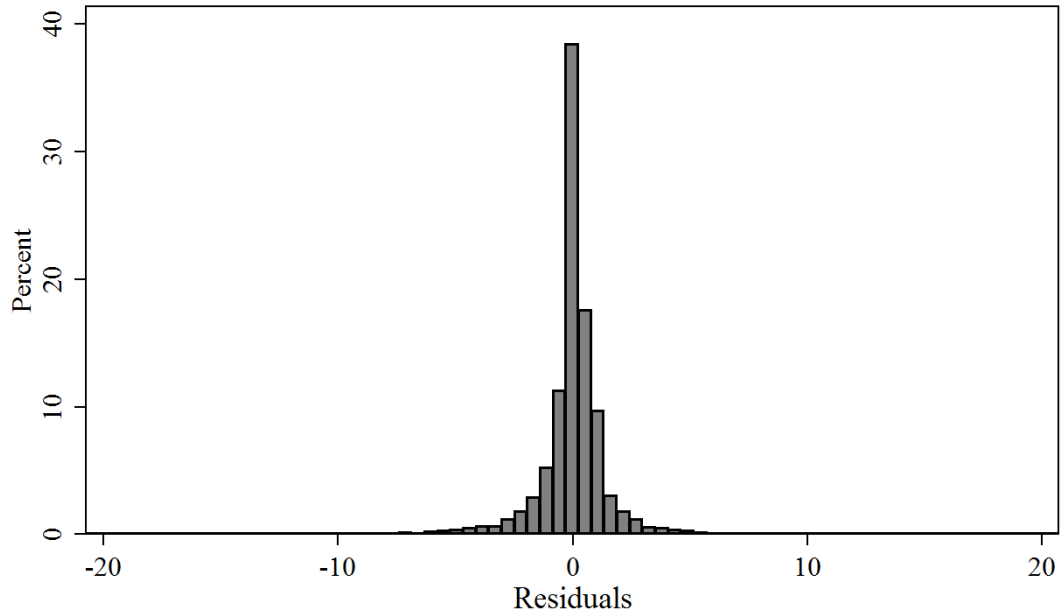
Figure 7

This figure shows the residuals from regression of Volume on Comments.

The residuals have a value of skewness of 0.16 and a kurtosis of 15.99.

Table 3.1: regressions of eq. 1a and eq.1b using data including outliers.
These regressions are used as a test of robustness.

	(Eq. 1a) Cross sectional	(Eq. 1b) Cross sectional	(Eq. 1a) Pooled	(Eq. 1b) Pooled
	Volume	Comments	Volume	Comments
Comments	0.473***		0.471***	
	(0.003)		(0.009)	
Volume		2.087***		1.813***
		(0.017)		(0.028)
Constant	0.435***	-0.874***	0.414***	0.472***
	(0.008)	(0.028)	(0.075)	(0.115)
Observations	21,858	21,858	745,381	745,381
R-squared	0.988	0.988	0.853	0.853

Robust standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Correlation of regressors

This table shows the correlations of variables used in the equations.

	Volume	Comments	MC	ROE
Volume	1			
Comments	0.874***	1		
MC	0.127***	0.138***	1	
ROE	0.005***	0.006***	0.121***	1

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1