



SCHOOL OF
ECONOMICS AND
MANAGEMENT

A most unusual tool for trading, YouTube.

A study that evaluates the possibility of using YouTube to make abnormal returns on small-cap stocks.

Bachelors Thesis in Economics NEKH02

Author: Sharif Radwan

Supervisor: Nagihan Mimioglu

Lund University, School of Economics and Management

Date of completion: 11/1-2021

Acknowledgements

I would like to thank Nagihan Mimioglu for her expert advice and I sincerely hope that I have been a worthy, first student at Lund University that she has supervised in the journey of writing a paper.

Abstract

Making money quickly and easily is something that most of us would without a doubt want to do. The process of doing this is however harder than most people imagine. YouTube has long been a platform for sharing ideas and creating entertainment content, but is it possible to use this platform to make money on the stock market? The goal of this study is to evaluate if there is a possible relationship between the abnormal returns of the stocks that are recommended on YouTube videos, and the sentiment behind the videos. Using data that measures the sentiment of the stocks that are recommended to the audiences of the channels, several regressions were conducted to test for a relationship. It became evident that such a relationship did not exist. Though this holds true, other findings, such as the individual cases of extreme abnormal returns turned out to be of great interest in this study. This study also faced a few issues that may have had some implications on the models.

Keywords: YouTube, Regression, Sentiment, Abnormal Returns, Small Cap.

Table of Contents

1. Introduction	5
2. Literature Review	7
3. Data	12
3.1. Data Overview and Collection	12
3.2. Variables	14
3.3. Data Processing	14
4. Methodology	16
5. Empirical Results	18
5.1. Extreme Cases	19
5.2. Distribution of Returns	19
6.1. Limitations	22
6.2. Alternative Explanation of Results	23
7. Conclusion	25
References	27
Appendix	30

1. Introduction

Being able to predict the future movements of stocks is a skill that some people would pay huge amounts of money to master. But the approach to how this could be made is very different. There is a whole variety of different ways to speculate on the stock market; this could be through fundamental analysis on a stock, or technical analysis. For most traders, sentiment trading is not very known, or at least still a fresh topic worth discussing and investigating. The research around this topic is usually very advanced and covers multiple areas in finance but also in data science and deals with machine learning (another term for AI) and neural networks (Bollen et al., 2011). Usually, this kind of research is used by bigger hedge funds and so the normal masses find it hard to use it and apply it. Data amounts collected for these kinds of studies tend to be very big, therefore huge computational power is usually needed to process all of it (Nguyen et al., 2015). To the best of my knowledge, there is no research using *youtube.com* as a data collection platform of sentiment and using sentiment as an exogenous variable to determine if it has any statistically significant effects on the return of a stock or not. The natural research question that arises is thus: **Can YouTube be used to identify price trends of small cap stocks on the American stock market?**

The general objective of this study is to determine if there is a statistically significant relationship between the abnormal returns of small cap stocks and variables that can be found in various YouTube videos. It will be determined whether it is possible to obtain consistent profits and hence “beat” the market or not. Another goal of this study is to enlighten the general mass and institutions about a possible relationship between the abnormal returns and variables from a website of which its following is only growing by the day.

This paper can assist and help retail investors and institutional investors gain better understanding about where the price-trend of a stock could potentially go. This will eventually result in better decision-making. I also believe that this research will give way to the importance of the irrational behavior of the common retail investor and consequently prove how the incorrect prices of a company are not really of great importance as it is set out to be according to the most influential research on this topic.

This research will thus set out to challenge the general idea of the Efficient Market Hypothesis.

There were not any statistically significant results that to some degree would portray a good and reliable relationship. I believe that this is due to a big list of problems that are presented in the “Discussion” section, below.

2. Literature Review

The broader topic of being able to predict future stock prices or future stock price directional moves is very widely studied. From the well-known efficient market hypothesis (EMH) that sets out to prove that changes in stock price are mere random walks, to the more modern technical analysis that entails that not all information is reflected (at least directly) in the price of a stock. What has not been very much researched though and experimented with is the influence of social media on stock prices. There are certainly a notable number of critical studies in this area, but most of these either depend on platforms that may or may not encounter different biases and problems (most notably *twitter.com*).

In a more general note, the popular EMH was first introduced in the 1960s by Fama (1991; Fama et al., 1969). Fama describes how stock prices already reflect all available information. This information includes news, articles, etc. Therefore, change in the stock market is merely a result of updated news and information. This then builds upon the notion of randomness of news and information which leads us to the “random walk” pattern that was presented by Burton G. Malkiel (1999). Though this theory outlines a popular view on prediction of trends in the stock market, there have been several attempts of contradiction.

The first of these is the general attempt of proving that the stock market does not in fact follow a random walk, but instead could be predicted to some degree. (Bollen et al., 2011; Qian & Rasheed, 2007; Vu et al., 2012). Other popular attempts of disproving the EMH are fundamental analysis and technical analysis. Fundamental analysis builds upon company specific information such as price to earnings ratio, and other metrics that determine the strength and the health of the company. In addition to this, it also uses macroeconomic metrics such as GDP, economic cycles, interest rates and employment numbers to determine the future price of a stock (Nguyen et al., 2015). Technical analysis on the other hand is based upon the historical prices of the company stock. According to the notion of technical analysis, history repeats itself and therefore one can construct a comprehensive model of prediction using different patterns, averages, and even mathematical occurrences to predict outcomes of the price of a stock (*ibid.*).

The sphere of behavioral finance has also tackled the EMH. The psychology of humans is at the center of inefficient pricing in the stock market. Different phenomena such as mob

psychology and hysteria are used to prove the irrationality of retail investors (Aliber & Kindleberger, 2017). This irrationality could be one of the reasons as to why different anomalies such as the size effect and the momentum effect have appeared. Shiller (2003) also expresses his thoughts about the “sharp contradiction” of behavioral finance to the EMH. Shiller goes on to explain how all the anomalies (including the popularly known January effect and the day-of-the-week effect) are not as important or vastly contradictory in regards to the notion of excess volatility. Excess volatility seems to be a result of “sunspots”, “animal spirits” or mass psychology (ibid.). Lots of research has hence been conducted to test this hypothesis; a couple of these are by Campbell and Shiller (1988) and West (1988). These, along with publications such as *Beyond Greed and Fear: Understanding the Behavioral finance and the psychology of investing* (Shefrin, 2002) laid a solid foundation to what came to be a plethora of research in the newly formed field of behavioral finance.

To introduce the notion of psychology and sociology and how it affects finance, an interesting story about the first time in history the irrational behaviors of investors created some sort of bubble that eventually burst (much like some “pump and dump” stocks) must be presented. This real-life event is often referred to as the “Tulip Mania” and was introduced by the famous book, *Memoirs of Extraordinary Popular Delusions* (1850). This event is set out in Holland in the 1630s and it entails how the feedback model works. This model is of great importance in the sphere of behavioral finance as it lays a solid foundation for the common phenomena that sometimes take place in the markets. The event that took place describes how the value of tulips at that time grew to an outstanding level in a very short period of time. This was because of the “word of mouth” effect that began with some investors having success in the market and making money out of the tulips that they bought and sold. This often also meant that investors usually confirmed each other's theories and agreed upon the future bullishness of the prices (Barberis, et al., 2003). This eventually led to public attention and resulted in a feedback loop where the demand was increasing at a rapid speed. The loop built upon the belief that the future prices would be higher and higher, leading to more and more investors buying tulips until prices reached unsustainable levels. This bubble eventually burst, and prices corrected themselves (ibid.).

This feedback model can be constructed upon some other theories as well, such as over-representativeness (Tversky & Kahneman, 1974). In the Tversky & Kahneman (ibid.) paper, an experiment is carried out where subjects are presented with different people of different

attributes and personalities that are described to them. The subjects of the experiment are then to guess the professions of these individuals. The overwhelming majority of subjects usually guessed professions that tend to fit as closely as possible to the description, without considering the rarity of some of these professions (ibid.). Niel, Hirschleifer & Subramanyam (1998) have shown that “biased self-attribution” could also be viable as an effect that results in a feedback loop. This pattern of human behavior was presented by Daryl Bem (1965) in an experiment where individuals link positive events to their own high ability and link negative events to bad luck or self-sabotage. The positive events would confirm the actions of the investor, whereas the negative events would be of bad luck or simply the initial awareness that the result would be bad.

Shiller (2003) asserts and stresses the fact that “nothing could be more absurd than to claim that everyone knows how to solve complex stochastic optimization models” (p.96). A conclusion must be drawn that most retail investors of course do not solve these kinds of optimization problems before they buy or sell stocks. The claim that institutional investors also solely rely upon advanced models is perhaps also somewhat unrealistic.

What has been mentioned above is certainly in contradiction to the EMH and challenges it in a very fundamental way; it can this be concluded that the sentiments of investors do affect the movement of stocks. The EMH cannot entirely be wrong though, this is something that Shiller (2003) notes and in the long run, the EMH will probably hold some truth to it, but in the short- to medium term, the story is certainly different.

Currently, sentimental analysis has not been such a huge research topic regarding stocks, although it has been found to contribute with a great number of applications in restaurant reviews, product reviews, and movie reviews (Liu & Zhang, 2012). Sentimental analysis can be used to derive a public opinion about a thing or a topic. Schumaker and Chen (2009) have for instance used the “bag of words” method to classify financial news as good or bad. Zhang, Fuehres and Gloor (2011) used Twitter instead, and were able to take advantage of the “mood” of each tweet that they collected data on. This enabled them to classify each tweet in “mood-clusters” such as “hope” and “fear”. One of the results was the negative correlation between the amount of “mood” tweets and the performance of the NASDAQ and other indices. Unsurprisingly however was the positive correlation to the VIX index.

Another interesting study that was done by a group of researchers began with choosing the website *seekingalpha.com* for their study. This was particularly interesting because lots of biases and problems that arose in previous studies were to a lesser degree found at that particular platform. The study explains how authors on the website got rewarded for “good” articles and that moderators (unlike other social media platforms) exist on the website. It was also described how there could be certain monetary fines and bans from the platform if misconduct was apparent. Consequently, it could be argued that it is in everyone’s interest to be honest and give good opinions on the platform. The authors conclude that their findings are “economically” meaningful regarding the relationship between sentiment and stock returns (Chen, Hailiang, et al., 2014).

In conclusion it is evident that there is some research about this topic that, granted, is less known, but still gives a good glimpse into the relationship between sentiment and stock returns. The authors have all used different ways of classifying the sentiment from the different platforms, these include the “bag of words” method, or the “number of posts”, or the “mood” of different sentences. Then they have connected this to the abnormal stock returns in a certain period of time (usually a short one) in order to see if there is any form of relationship between the two variables. Different conclusions have been made. This is normally due to the different methods used, the different amounts of data or the types of data used from different sources on the internet. It is obvious that there is more room for further research on this topic.

The only big social media platform that has not been used or even mentioned in all past literature (to my knowledge) that is only getting bigger, especially in the topic of stocks is *youtube.com*. I believe that this platform will erase some of the problems that *seekingalpha.com* had. Not because there are moderators, but because there is a system of “rating” in this platform. This is done through liking a video and subscribing to a channel. Of course, the “youtubers”, as they are called, do also have incitements to produce videos and valuable stock recommendations in order to gain “Likes”, subscribers and achieve traffic on their videos to eventually make money through ads. Not only does this cover problems addressed by the study mentioned, but it also makes it much easier for us to conduct an experiment in a certain time span and filter the most viewed and liked videos

from the most subscribed-to gurus, so that the “best” recommendations of stock picks from the audience’s most beloved “youtubers” can be gathered.

3. Data

3.1. Data Overview and Collection

To meet the general, agreed upon “requirements” for running inference, a sample that is randomly selected and thus independent must be gathered. For this, I have chosen the year of 2020 to gather the data required. A big argument could of course be made as to the returns of 2020 being extremely distorted and biased due to the covid-19 pandemic, coupled with the turbulent political sphere of that year. This is of course totally understandable and if I were to decide, I would not have chosen the year of 2020 for this study. But the problem with data before 2020 is that there were almost no channels at all uploading videos on a regular basis that were recommending small cap stocks. This community of investors also grew substantially 2020. As mentioned before, retail investors currently account for a total of 25% of the American stock markets, as opposed to the low numbers of the year before (Dong et al., 2020).

As for the data, it is obvious that from about March until December, the number of videos that recommended different stocks for short-term holding increased. The data is thus very sufficient and can hence be used for subsequential analysis among other things. The problem with this data is that it is not easily obtainable. This is due to YouTube not offering a comprehensible API as Twitter, for instance, does. Also, due to the need of programming skills (most notably in Python), I could not gather vast amount of data in a short time. The only solution for this was for me to manually gather all the data needed. I proceeded with randomly choosing three months to collect the data from. These months turned out to be June, September, and November. As for the videos, it was impossible to filter them in a way that would yield all of them in a single search. This led me to (again) randomly choose five YouTube channels that at the time had the same kind of audience whilst still having very close amounts of subscribers each. These YouTube channels turned out to be “Ziptrader”, “Matthew Huo”, “JT Wealth”, “The StockWatch” and “My Financial Friend”. The time constraints also made it very hard to collect huge amounts of data and it was very tedious to have to watch video after video and gather the amount of “Likes”, “Dislikes”, “Views”, and “Sum-of-Repetitions”.

The aim of the “Sum-of-Repetitions” variable was of course to better explain if a certain stock could generate higher returns if mentioned more than once. A case could be made for the collection of “Subscribers”. This variable basically entails the audience of the “youtuber”, and although that may to some extent explain our regression better, I found it impossible to gather the number of subscribers of a video at the time of its release. Additionally, this variable could not be estimated easily as most of the channels that were selected for this study had enormous growth during the months of the latter half of 2020. As will be explained later, another variable that was gathered yielded very insignificant results and impacted the regressions very negatively. The variable “Subscribers” had unfortunately similar traits. As for the abnormal returns of each stock, three variables were defined. The first one being the return of the stock two days after the release of the video, the second one was two weeks after the release of the video and the last being one month after the release. The reason for choosing such short time frames was the fact that almost every YouTube channel told its viewers to hold the recommended stocks for short periods of time. This short period was not generally defined, but most of the suggestions were due for catalysts in the form of different company news, financial statements, and breakouts in a pattern (technical analysis). On that account, it would only be fitting to define three periods that all to some degree would give us different and defined timeframes.

All the data was gathered in Microsoft Excel, and at the end, the observations amounted to a total of 118. I also gained a valuable theoretical view of this community that would prove to come in handy in the latter part of the study. For the returns of each stock, an online application called “Stock Return Calculator” (customstockalerts, 2021) was used. This application enabled me to derive the returns of a stock in custom periods of time. As mentioned before, the return for the time periods, two days, two weeks and one month were collected. The abnormal returns also made it necessary for me to collect data on an S&P 500 ETF. I did this through Yahoo! Finance (Yahoo, 2021) and managed to collect the returns for the time periods that I needed.

The distribution of data from the different YouTube channels turned out to be a bit biased towards one channel that provided huge amounts of data. As **Figure 1** indicates, most of the data was gathered from a YouTube channel by the name of “The StockWatch”. This can of course be viewed as something very problematic, but I realized the problem of this a bit too late but reckoned that I would at least end up with a sizeable amount of data that would make

up for the loss of unbiasedness in the end. I could also not really choose from what channel to collect the data. This is a big flaw of the study and to be honest it is very hard and problematic to gather data this way because all these issues do not become apparent until later.

Figure 2 indicates that most videos on stock recommendations were made for the month of June. This is obviously very biased, but it was crucial to collect as much data as possible to test the hypothesis as good as possible.

3.2. Variables

The explanatory variables that were chosen for this study were:

1. “Dislikes/Views”
2. “Likes/Views”
3. “Sum-of-Repetitions”

The dependent variables were:

1. Two-Day Abnormal Returns
2. Two-week Abnormal Returns
3. 1 Month Abnormal Returns

3.3. Data Processing

Early on into the process of gathering data, I realized that using the collected variables themselves would not make for a very trustworthy or comprehensible model. It was thus obvious that the absolute values of the variables gathered would only grow and that this would not really explain the different returns. To solve this issue, I normalized every variable that I gathered. For example, instead of using the “Dislikes” and “Likes” for themselves in the model, I used the relationship between these two as a certain “grade” of how happy the viewers of the video were with the recommendations. Of course, being happy or not happy about a YouTuber recommending a stock is not the only reason that would influence whether someone would press the “Like” or “Dislike” button, but it can be assumed that most people watching such videos are not there for the sole purpose of entertainment, but instead of picking stocks that would be of benefit to them. As for the variable “views”, I reckoned that the only way I could truly normalize it was making it as some sort of “improvement factor” in the content (stock recommendations) of a YouTube channel. The way to do this was to divide

the views of a video with the number of views of the video before it. This of course did not work very well as it was later obvious that this variable showed signs of redundancy and insignificance to the inferences and the regressions that were run. To be able to easily understand why, a case can be asserted for how similar trends can seem to have a relationship, but logically not. A common example that can be found on the internet of this phenomenon is the relationship of mozzarella cheese consumption and the number of Civil Engineering PhDs awarded. In this model, a very high fit of almost 92% and results that show high statistical significance can be observed (McNeese, 2016). Does this mean that there is a relationship? In the numbers, yes, but when it comes down to logic and reasoning we cannot make such an assertion at all. For this reason, the “views-on-views” variable did not work very well as it would obviously show that the audience of a channel grew and that it in some way was linked to how good of a return the stocks that the channels recommended would be. This resulted in the problem of how the usage of the views-data could be incorporated into the regression. After some advice-seeking, this was done through removing the “Like/Dislike” ratio and instead introducing a “Like/Views” ratio and “Dislike/Views” ratio. No processing was needed for the “Sum-of-Repetitions” variable.

The abnormal returns were derived through the usage of a well-known S&P 500 ETF. The easiest way to do this was to gather the data on the returns in the same time periods of the stock returns and then taking the difference between the ETF return and the stock return for the subsequent observation. The ETF that was used is the iShares Core S&P 500 ETF (yahoo, 2021) as it is a widely used ETF.

4. Methodology

If the objective of this paper is to test whether there is a clear relationship between the YouTube sentiment and the consequent abnormal returns of different time periods, then it is easiest to first try to understand how something as abstract as the “sentiment” of the public can be measured. I have found it elaborate to use variables that already in some way measure the sentiment of the video itself. The first of these being of course the amount of “Likes” and “Dislikes” of the video, how many times the stock is repeated, and then, of course, the views of the video. The inspiration to do this was from a paper by Chen, Hailiang, et al. (2014). They managed to gather metrics that in a similar way would determine how the respective sentiment for each stock would be.

After gathering all that data manually, I will process the data in a way that can yield some useful variables. Afterwards, different regressions will be made to see if there is any relationship between a mix of different explanatory variables and different time frames of abnormal returns. For this, the statistical application “Gretl” (Cottrell & Lucchetti, 2020) will be used, as I have personally found it the easiest and fastest to use and it also provides the possibility of printing out graphs of different sorts. Whilst doing the regressions, I will report my findings in an excel file, and then afterwards, some graphs and tables will be used to showcase the results.

The way that the regressions were done was through testing all variables and how they, in different combinations of each other, can explain the different abnormal returns. Therefore, there were seven different regressions that were each run for the different time frames of the abnormal returns. There are hence 21 different regressions that are all divided in three groups. Descriptive statistics of these regressions are shown in the appendix and thus described here in the results. The different regressions are denoted by the subscripts “s”, “l” and “d”. As an example, the 2-day_{1,0,1} is a regression of the two-day abnormal return run on the “Sum-of- Repetitions” variable and the “Dislike/Views” variable. The seven possible combinations of regressions are thus:

$$(1) y_{i,t} = \alpha + SumRep x_{s,i} + \left(\frac{Like}{Views}\right) x_{l,i} + \left(\frac{Dislike}{Views}\right) x_{d,i} + \varepsilon_i$$

$$(2) y_{i,t} = \alpha + SumRep x_{s,i} + \left(\frac{Like}{Views}\right) x_{l,i} + \varepsilon_i$$

$$(3) y_{i,t} = \alpha + SumRep x_{s,i} + \left(\frac{Dislike}{Views}\right) x_{d,i} + \varepsilon_i$$

$$(4) y_{i,t} = \alpha + \left(\frac{Like}{Views}\right) x_{l,i} + \left(\frac{Dislike}{Views}\right) x_{d,i} + \varepsilon_i$$

$$(5) y_{i,t} = \alpha + SumRep x_{s,i} + \varepsilon_i$$

$$(6) y_{i,t} = \alpha + \left(\frac{Like}{Views}\right) x_{l,i} + \varepsilon_i$$

$$(7) y_{i,t} = \alpha + \left(\frac{Dislike}{Views}\right) x_{d,i} + \varepsilon_i$$

After finding an important result that made it appropriate to run a logged regression, a model was created that was based upon the best linear regression that was found. This model is presented in **Regression 9**. After running the regressions, a case will be made for whether clear, significant relationships can be discovered. This will of course be shown through different tables, some graphs, and some descriptive statistics to better visualize the data and the results. It is also asserted that the problem of heteroskedasticity will be very apparent, we will thus conduct a Breusch-Pagan test on the “best” model to test for this. Lastly, we will discuss the results and review the implications of these and how they may be tied to previous research.

5. Empirical Results

There were seven different regressions that were run on the two-day abnormal returns; these results are all presented in **Table 1**.

A conclusion can be drawn that the adjusted R^2 from all the regressions are not very high at all; the only regression that somewhat draws any attention to it is the one that only uses the “Sum-of-Repetitions” variable and the “Dislike/Views” variable as explanatory variables. This one has got the highest fit and the model is:

$$(8) y_{2-day} = 0,044 + 0,022x_s - 47,463x_d$$

The absolute values of the t-ratios of every variable do not exceed 3. The p-value for the regressions is nowhere near any significance at all, with the best one being the above-mentioned regression.

There were seven different regressions that were run on the two-day abnormal returns; these results are all presented in **Table 2**.

The adjusted R^2 from all the regression are again, not very high. In fact, the results of these regressions seem to be a bit worse than the 2-day regressions. The only regression that in some way is interesting is the one that is solely based on the “Sum-of-Repetitions” variable. Although this may somewhat portray this explanatory variable as a very effective one, it could be clearly shown through **Figure 3** that the balancing nature of the values that were observed make for a good fit, this model can therefore not be relied upon. The same phenomenon occurs in the regression for 1 month that uses “Sum-of-Repetitions” as its only explanatory variable. The p-values of all the regressions are not very significant and this is also shown through the t-ratios.

In the third type of regressions, the 1-month abnormal returns were explained with the same variables. Gathering the results from **Table 3** we can see results that are very similar to the case of the 2-day abnormal returns. The p-values are again not of any significance at all and this is also portrayed in the t-ratios of each variable.

5.1. Extreme Cases

Upon investigating the data of the different abnormal returns, it became obvious that the 1-month data had some very extreme cases of high abnormal returns. This is not visible at all in the regressions as they would somewhat disappear. Some notable examples of these tickers that all exceeded 100% return, some even 150% are \$PLG, \$CDAK and \$SOLO. This of course goes to prove the notion of “high risk, high reward” in trading stocks such as these in smaller time frames.

5.2. Distribution of Returns

The notion of very extreme cases is apparent when quickly looking at the data; this is confirmed through **Figures 4-6**. They look very similar to a chi-square distribution. This distribution is skewed mostly to the right side, where it can be seen that there are extreme cases of high positive returns. The positive skewness can be normalized though. According to Tukey (1977), a logged regression can be used to transform a model to a somewhat normalized model. The easiest way to test this is through transforming our “best” model thus far. **Regression (8)** will thus be transformed into:

$$(9) \ln y_{2-day} = 0,044 + 0,022 \ln x_s - 47,463 \ln x_d$$

This model displayed a better adj. fit of about 0.081, as opposed to the observed 0.014 of the linear model. The exogenic variables also display good significance as can be seen on Table 4. Although this is a substantially better result, it still does not yield a meaningful relation whatsoever.

5.3. Breusch-Godfrey Test

The Breusch Godfrey test was conducted on **Regression (9)**. The test yielded a p-value of 0.69 and the null hypothesis that the error variances are all equal can thus be rejected.

6. Discussion

Do the results directly disprove the hypothesis? The research encountered a number of problems that are presented in the paragraphs below. Perhaps such research cannot only be conducted using statistics, but maybe also include some sorts of theoretical framework that in a qualitative manner can test this hypothesis. I believe that the assessment of each individual return by itself would give us a deeper understanding of how the price movement of a ticker occurs after a sudden spike of sentiment occurs. This notion is hard to grasp, but these extreme values that can be observed cannot only be reflective of the fundamental value of a stock. These results should hence be used as some sort of reflection upon this subject of behavioral finances as this would only help future research to incorporate all the missing variables that have to be considered in order to create a comprehensible model that captures future price movement better.

The models that were run did not successfully capture the abnormal returns of the respective stocks. The adjusted fits, the t-ratios and p-values of the linear regression did not show any significance at all on any level whatsoever. The logged regression did however prove some sort of significance in the explanatory variables. What was more interesting was perhaps the fact that the abnormal returns of all three time periods resembled that of a Chi-square distribution, but in hindsight, this could have been a result of the possibility of more upside to a return, than downside. Some of the extreme values that were found in the return-data were also analyzed and commented upon.

The objective of this research was to evaluate if there exists a relationship between the abnormal returns of stocks that are mentioned on YouTube and the YouTube audience sentiment in relation to these stocks. What inspired me to investigate this was me realizing that my YouTube feed suddenly one day in the summer of 2020 got filled with suggestions of videos that usually recommended small cap stocks. This got me very curious as to why these kinds of videos were trending on YouTube and after watching a few of them I got interested in determining if there was a way to make any money out of this. Of course, the numerical relationship between making money in the short term and the sentiment of the stocks recommended was hard to prove and this was something that I was aware of going into this. The fact that the number of observations collected were very low, only made this task harder. But I still believe that it is sensible to analyze this relationship from a

standpoint of individual cases. The assessment of each stock and why it sometimes manages to draw the public attention and gain good momentum is very interesting. Some of these stocks even managed to create a “fanbase”, and in certain cases it could be a loyal such. A notable example of this is \$TSLA. It is astonishing to see how a company that fundamentally speaking should not be valued the way it is (Cornell & Damodaran, 2014) have gained such an upward momentum. Its market cap does not in any way reflect the fundamental value. Once a retail investor is asked why everyone is betting so much on this stock, you are always given an answer something in the lines of: “the value of this company lies in its future”. The natural question that arises out of this is how the future value of a company can be the value of the company today. This notion is in stark contrast to all classic theories that today make up most of the research of the stock markets. To further develop this argument, an interesting question to ask is how to fundamentally evaluate a company; who really decides on how to do that?

There was no notable excess volatility that could be directly observed in the models, but I believe that looking at some of the individual values could give as an important insight as to what the excess volatility could look like. The excess volatility can be explained by several factors such as mass psychology and herding (Shiller, 2003). The behavioral side of this is thus of great importance to understand how our models work. Most of the “extreme values” managed to disappear in the model and this made them very hard to observe. The over-representativeness that Tversky & Kahneman (1974) presented could to some degree also explain the feedback loop that was created and observed in the extreme values of this study. The fact that most of the retail traders believed that they could make huge abnormal gains on a stock created a feedback loop, that only made the volatility higher and higher. This goes to prove that irrationality is the main factor that drove these stocks above their so called “fundamental value”.

“Smart money” is often a term that is flaunted in this community of retail investors on YouTube. The definition of this term is basically the investments of the institutional investors (Sørensen, 2007). ARK Invest, among others, exhibits their portfolio publicly and announces any changes and new investments on a regular basis. This has gained an interesting following that to some degree have made the volatility of these stocks skyrocket. Shiller (2003) explains that “nothing could be more absurd than to claim that

everyone knows how to solve complex stochastic optimization models”, but what if you could simply follow the “smart money” and thus gain create a good portfolio that could generate abnormal returns?

The study by Chen, Hailing, et al. (2014) was very similar to this paper. Their study incorporated huge amounts of data that they could relatively easily get a hold of. They measured the sentiment through using similar techniques as in this study. The negativity and positivity of a comment would in this case be the same as a “Dislike” or “Like”. Other variables they included were the volatility of the price of a stock during a certain period of time and the number of downgrades or upgrades (recommendations). *Seekingalpha.com* was solely a stock market community and thus could not have any sort of distortions that were not related to the stocks as we could perhaps not conclude using YouTube sentiment as a predictor of stock price movement.

6.1. Limitations

This study is limited to a period between June of 2020 and November 2020. This is the result of the random choice of months that was carried out for the sample to be as randomly chosen as possible. Additionally, it is obvious that more months could have been selected (randomly), but due to time constraints and the problems of manually collecting the data, I went with only choosing three months as this would still give me sufficient data for this study. The choice of limiting the data from five YouTube channels also accounted for similar reasons mentioned above. The more, the better, of course, but the manual collection of the data was tedious, and it required me to dedicate a lot of time.

Usually, the YouTubers always make disclaimers before any video starts. They generally state that they are not financial advisors and that their opinions and analysis on the stock should not be used as the sole tools for trading. They also advise the viewers to make their own due diligence and analysis if they ever chose to trade the recommended stocks. It is obvious that this kind of analysis should go hand in hand with personal analysis and due diligence of stocks for a person to make a qualified guess on how the price of a stock will unfold. Capturing this kind of data would simply be too overwhelming, and it cannot in any way be gathered, unless some sort of machine learning algorithm is used. This data has accordingly been excluded; this study relies only upon the recommendation and analysis of

the YouTube channels. If these “youtubers” turn out to have any predictive powers whatsoever, then this should obviously be reflected in the results of this study.

Another implication that may arise is the skewness of going long on stocks. This is due to the shorting bias in the stock market. The consequence of this is that the movement of a stock is usually skewed to the upside rather than the downside (Nagel, 2005).

Retail investors currently make up about 25% of the American stock market (Dong et al., 2020). This means that retail investors do not affect prices that much and this can also turn out to be problematic to our study. This is the main reason as to why stocks that have a market capitalization rate below one billion dollars (small cap) were chosen. This (hopefully) enabled us to spot some significant moves that were mainly a result of retail investing.

Obviously, some statistical issues arose as well. These may include the consistency of the model; the data will not be that big due to time constraints and the fact that the data must be collected manually. Another problem that came to light is the notable multicollinearity between the variables as it is well known that variables such as “Views” and “Likes” will be somewhat correlated with each other. To understand this, it can be concluded that a portion of the people who view the video, press the “Like” button as well. In a reversed manner, those who press the “Like” button also view the video. The amount of “Likes” is thus contained in the “Views”. “It is not an issue of multicollinearity, but instead the degree of multicollinearity”, this notion must be considered in this paper.

6.2. Alternative Explanation of Results

The results above could likely be explained by several of the following factors:

- The first of these being the fact that one needs to gain some insight of the proper holding period of each ticker, if not, the return could end up have been measured in an incorrect way. This is unfortunately very hard to do as most of the videos that suggest these stocks only provide terms such as “short term” or “couple of days”,

or “until next week”. The only comprehensible way to do this is thus to choose some general holding periods that can be identified as “short-term”.

- Another implication of these stocks is that some of them are commonly referred to as “pump and dump” stocks. These stocks are hence not invested in for the long term, but usually subject to a sudden spike in demand and later a sudden sell off that usually takes place the same day. This only goes to prove that what moves stocks such as these, are not really underlying fundamental factors, but the actions of investors that only seek to exploit the irrationality of the markets because of different phenomena such as herding and mass psychology. This makes it even harder to gain good insight into this topic.
- What is harder than buying in these situations is usually selling. This is where the timing of the market comes in. Upon closer look at the charts of most of the stocks, it can be observed that many of them usually have sudden spikes; these spikes are usually followed by a huge downside. Being able to time the market and sell at the right moment is thus of crucial importance. Due to practical factors however, we could unfortunately not integrate this into our data.

7. Conclusion

The purpose of this study is to test if there is a significant and meaningful relationship between the abnormal returns of stocks that are recommended on YouTube and variables that were extracted from the platform and then processed to measure sentiment. This study aimed to evaluate the thesis through the usage of short-term periods to evaluate if this was possible.

In this study, 22 regressions were conducted that all involved different combinations of the data on the variables that were gathered. These regressions were put in groups of three. The first group was included regressions that tested whether the explanatory variables could explain a two-day abnormal return. The second group tried to explain the 2-week abnormal returns and lastly, the third group tried explaining the 1-month abnormal returns. Upon identifying the “best” model of these, the model was transformed into a logged such. The regressions showed overall insignificance. This was reflected in the t-ratios of the explanatory variables and in the adjusted fit of the regressions. It was therefore concluded that the null hypothesis, stating that: “sentiment can explain the abnormal returns of YouTube stock picks in the shorter time period”, could be rejected. No model of the 22 showed any significant predictive power at all.

It can be argued that because of this, this paper can give way to the proof to the Efficient Market Hypothesis. According to Shiller (2003), EMH probably holds true in the long run, but not very much in the short term. Does this research disprove that of Shiller’s? Probably not, as there were big implications and problems to this research. Some of these included the sample size, and thus the consistency of the models. Also, the YouTube channels that recommended the stocks did not usually give any information on how long the holding periods should be. Consequently, appropriate generalization of the holding periods had to be conducted. Another problem that soared through the data was the biasedness. A huge portion of the data was collected in June, and a big part of it was collected from a single YouTube channel. This does not give a fair view of how the data should be proportioned, but due to some technical and practical issues, it was the only way to collect the data necessary.

Although the models did not prove to be of any use, it was very interesting to examine the data itself. As noted, a big part of the data had some extreme abnormal returns; this proved the notion of excess volatility that Shiller (2003), among others, presented in their work.

The biggest problem with this study was with the dataset. Although there is an API that is hard to use and require programming skills, I believe that the usage of it would be of great benefit for further research on this topic. The datasets that could be collected would be very big and thus yield overall better consistency in the models. The API also makes it easier to collect the number of subscribers of every channel on the platform at different time points. Another suggestion is trying to group the different stocks in clusters of “pump and dump” stocks and “short term plays” as this kind of grouping would be useful, it is evident in the data that this could be of great benefit to the model. To further develop the model, you could also incorporate some machine learning that in some way collects data through watching the videos themselves. This would certainly need huge amounts of computational power and expert programming skills. Other research of this nature could be about the stock investing community on YouTube, I believe that lots of the retail trading that occurs does have some relationship to this platform. Understanding in what way could thus assist individuals and institutions in understanding “where” the money is going.

References

- Aliber, Robert Z., and Charles P. Kindleberger. *Manias, panics, and crashes: A history of financial crises*. Springer, 2017. p. 57.
- Barberis, Nicholas, and Richard Thaler. "A survey of behavioral finance." *Handbook of the Economics of Finance* 1 (2003): 1053-1128.
- Bem, Daryl J. "An experimental analysis of self-persuasion." *Journal of Experimental Social Psychology* (1965).
- Bollen, Johan, Huina Mao, and Xiaojun Zeng. "Twitter mood predicts the stock market." *Journal of Computational Science* 2.1 (2011): 1-8.
- Campbell, John Y., and Robert J. Shiller. "The dividend-price ratio and expectations of future dividends and discount factors." *The Review of Financial Studies* 1.3 (1988): 195-228.
- Chen, Hailiang, et al. "Wisdom of crowds: The value of stock opinions transmitted through social media." *The Review of Financial Studies* 27.5 (2014): 1367-1403.
- Cornell, Bradford, and Aswath Damodaran. "Tesla: Anatomy of a Run-up." *The Journal of Portfolio Management* 41.1 (2014): 139-151.
- Cottrell, Allin and Lucchetti, Riccardo. 2020. *Gretl*. Accessed 2020-11-10 on <http://gretl.sourceforge.net/>
- Customstockalerts.com. 2021. *Stock Return Calculator*. Accessed: 2020-12-12. <https://www.customstockalerts.com/stockReturnCalculator>
- Dong, Hang, Javier Gil-Bazo, and Raluca Ratiu. "Information Demand during the COVID-19 Pandemic." *Available at SSRN 3696334* (2020).
- Fama, Eugene F. "Efficient capital markets: II." *The Journal of Finance* 46.5 (1991): 1575-1617.
- Fama, Eugene F., et al. "The adjustment of stock prices to new information." *International economic review* 10.1 (1969): 1-21.

Finance.yahoo.com. 2021. Accessed: 2020-12-15.

<https://finance.yahoo.com/quote/IVV?p=IVV&.tsrc=fin-srch>

Nguyen, Thien Hai, Kiyooki Shirai, and Julien Velcin. "Sentiment analysis on social media for stock movement prediction." *Expert Systems with Applications* 42.24 (2015): 9603-9611.

Liu, Bing, and Lei Zhang. "A survey of opinion mining and sentiment analysis." *Mining text data*. Springer, Boston, MA, 2012. 415-463.

Mackay, Charles. *Memoirs of extraordinary popular delusions*. Vol. 1. Lindsay and Blakiston, 1850.

Malkiel, Burton Gordon. *A random walk down Wall Street: including a life-cycle guide to personal investing*. WW Norton & Company, 1999.

McNeese, Bill. *Just Because There is a Correlation, Doesn't Mean...* BPI Consulting, 2016. Accessed 2020-12-14 on <https://www.spcforexcel.com/knowledge/basic-statistics/just-because-there-is-a-correlation>

Nagel, Stefan. "Short sales, institutional investors and the cross-section of stock returns." *Journal of financial economics* 78.2 (2005): 277-309.

Qian, Bo, and Khaled Rasheed. "Hurst exponent and financial market predictability." *IASTED conference on Financial Engineering and Applications*. 2004.

Schumaker, Robert P., and Hsinchun Chen. "Evaluating a news-aware quantitative trader: The effect of momentum and contrarian stock selection strategies." *Journal of the American Society for Information Science and technology* 59.2 (2008): 247-255.

Shefrin, Hersh. *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*. Oxford University Press on Demand, 2002.

Shiller, Robert J. "From efficient markets theory to behavioral finance." *Journal of economic perspectives* 17.1 (2003): 83-104.

Shleifer, Andrei. *Inefficient markets: An introduction to behavioral finance*. OUP Oxford, 2000.

Sørensen, Morten. "How smart is smart money? A two-sided matching model of venture capital." *The Journal of Finance* 62.6 (2007): 2725-2762.

Tukey, John W. *Exploratory data analysis*. Vol. 2. 1977.

Tversky, Amos, and Daniel Kahneman. "Judgment under uncertainty: Heuristics and biases." *science* 185.4157 (1974): 1124-1131.

Vu, Tien Thanh, et al. "An experiment in integrating sentiment features for tech stock prediction in twitter." *Proceedings of the workshop on information extraction and entity analytics on social media data*. 2012.

West, Kenneth D. "Dividend innovations and stock price volatility." *Econometrica: Journal of the Econometric Society* (1988): 37-61.

Zhang, Xue, Hauke Fuehres, and Peter A. Gloor. "Predicting stock market indicators through twitter "I hope it is not as bad as I fear"." *Procedia-Social and Behavioral Sciences* 26 (2011): 55-62.

Appendix

Figure 1. Distribution of Channels on Data

This graph illustrates the distribution of the YouTube channels that the overall data was gathered from. It can be concluded that this data is biased as the channel “The StockWatch” makes up more than 50% of the observations.

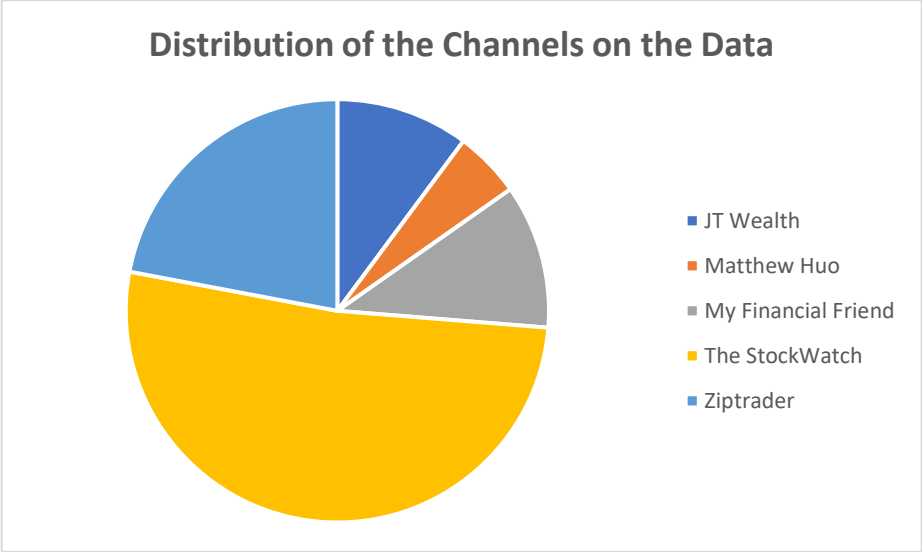


Figure 2. Distribution of Months on Data

This graph illustrates the distribution of the months that the overall data was gathered from. It can be concluded that this data is biased as the month of June makes up most of the observations.

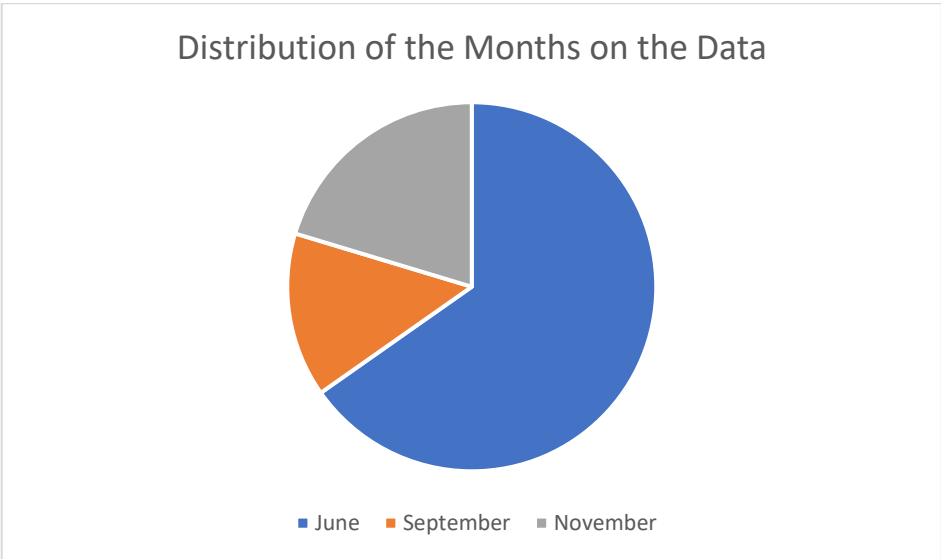


Figure 3. Regression Line

This graph illustrates the problem of the “Sum-of-Repetitions” variable. Although seemingly displaying significance and a good fit of the model, it can be concluded that this result is a direct consequence of the values balancing themselves out. This model is thus not very trustworthy.

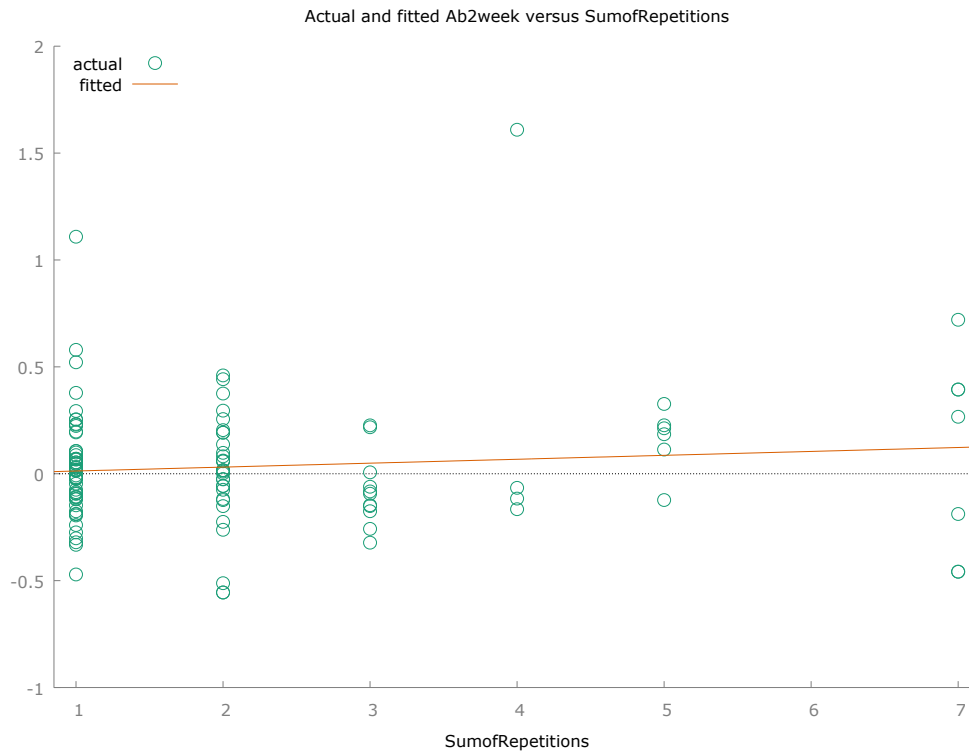


Figure 4. Distribution of 2-week Ab. Returns

This graph is an illustration of the supposed Chi-square distribution of the abnormal 2-week returns. This could be a result of the apparent, higher upside than downside. Note that the y-axis displays the percentage of the ab.return, and the x-axis displays the standard deviation of the ab.return.

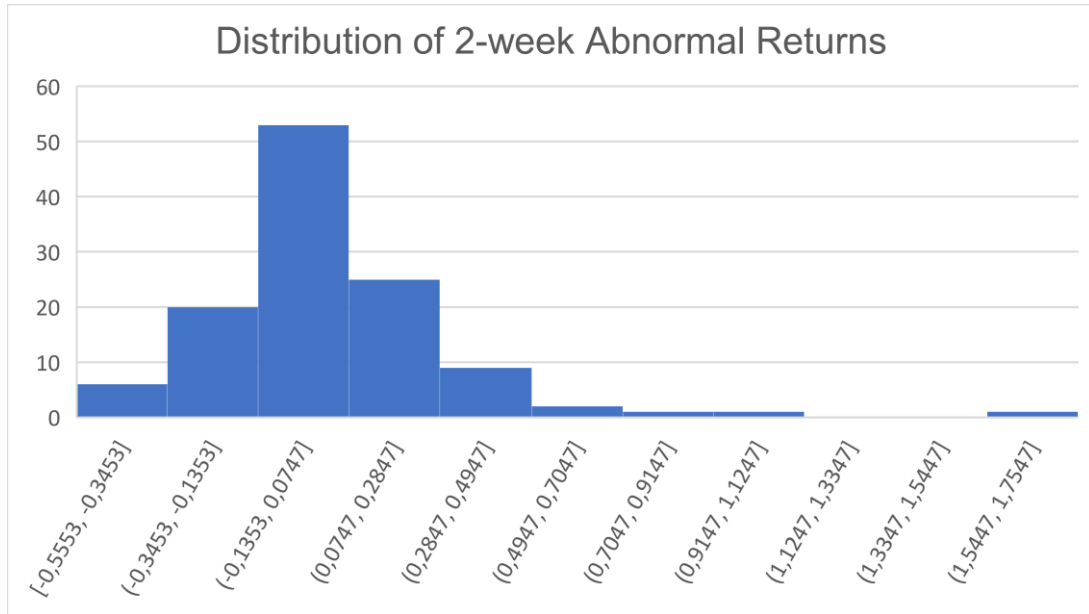


Figure 5. Distribution of 2-day Ab. Returns

This graph is an illustration of the supposed Chi-square distribution of the abnormal 2-day returns. This could be a result of the apparent, higher upside than downside. Note that the y-axis displays the percentage of the ab.return, and the x-axis displays the standard deviation of the ab.return.

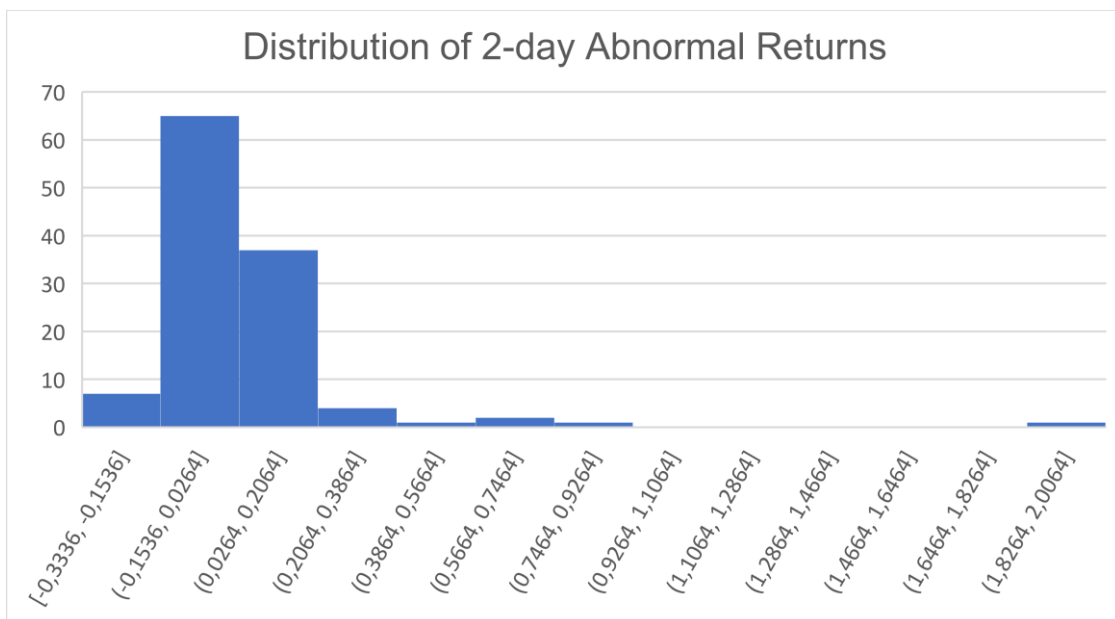
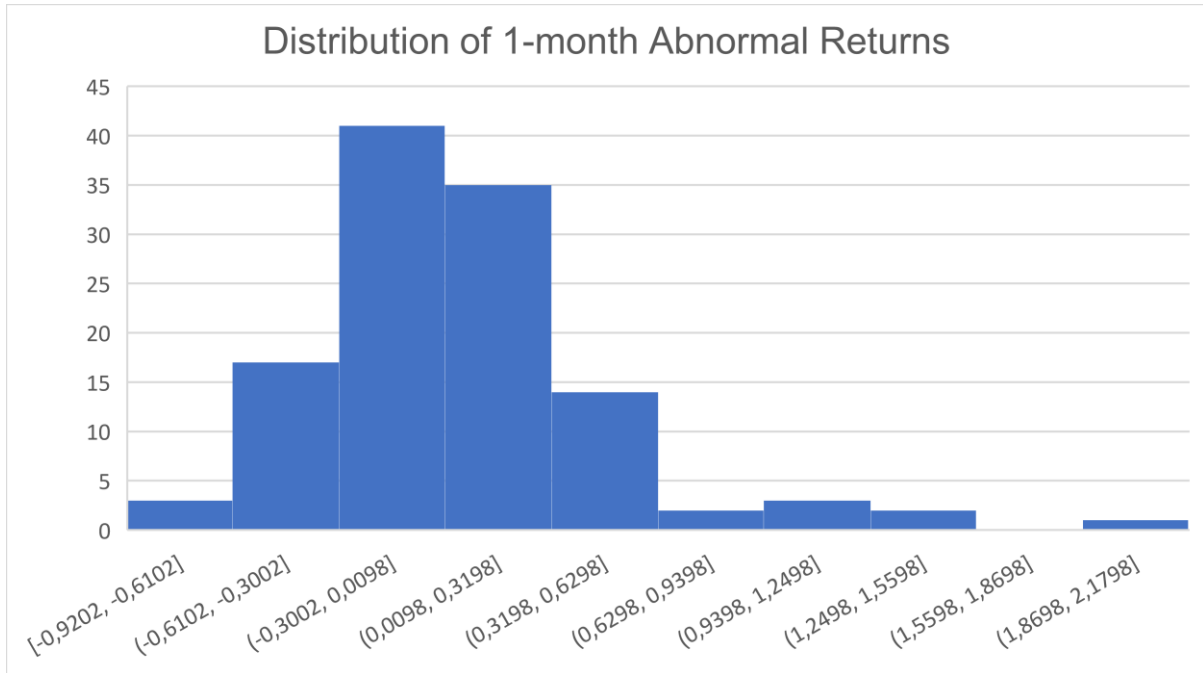


Figure 6. Distribution of 1-month Ab. Returns

This graph is an illustration of the supposed Chi-square distribution of the abnormal 1-month returns. This could be a result of the apparent, higher upside than downside. Note that the y-axis displays the percentage of the ab.return, and the x-axis displays the standard deviation of the ab.return.



The following page illustrates Tables 1-4 in a chronological manner.

* Please note that no significance was observed and as such, there is no p-value less than 0,05.

Dependent Variable	P-value (regression)	adj. R^2	constant	SumRepetitions (s)	Like/Views (l)	Dislikes/Views (d)	t-ratio (s)	t-ratio (l)	t-ratio (d)
2-day _{1,1,1}	0,295	0,006	0,073	0,021	-0,531	-46,826	1,388	-0,324	-1,523
2-day _{1,1,0}	0,496	-0,005	0,039	0,014	-0,691		0,957	-0,420	
2-day _{1,0,1}	0,164	0,014	0,044	0,022		-47,463	1,543		-1,553
2-day _{0,1,1}	0,478	-0,004	0,029		0,001	-20,719		0,469	-0,526
2-day _{1,0,0}	0,267	0,002	0,000	0,016			1,115		
2-day _{0,1,0}	0,483	-0,004	0,091		-1,116			-0,704	
2-day _{0,0,1}	0,261	0,002	0,074			-33,039			-1,129

Dependent Variable	P-value (regression)	adj. R^2	constant	SumRepetitions (s)	Like/Views (l)	Dislikes/Views (d)	t-ratio (s)	t-ratio (l)	t-ratio (d)
2-week _{1,1,1}	0,733	-0,015	0,021	0,018	-0,411	-3,208	0,987	-0,211	-0,088
2-week _{1,1,0}	0,527	-0,006	0,018	0,017	-0,422		1,014	-0,218	
2-week _{1,0,1}	0,537	-0,006	-0,002	0,019		-3,701	1,091		-0,102
2-week _{0,1,1}	0,856	-0,015	0,072		-0,942	8,011		-0,502	0,231
2-week _{1,0,0}	0,266	0,002	-0,006	0,018			1,118		
2-week _{0,1,0}	0,611	-0,006	0,082		-0,952			-0,510	
2-week _{0,0,1}	0,808	-0,008	0,023			8,427			0,243

Dependent Variable	P-value (regression)	adj. R^2	constant	SumRepetitions (s)	Like/Views (l)	Dislikes/Views (d)	t-ratio (s)	t-ratio (l)	t-ratio (d)
1-month _{1,1,1}	0,315	0,005	0,226	-0,045	-1,356	-11,719	-1,718	-0,475	-0,219
1-month _{1,1,0}	0,173	0,013	0,218	-0,047	-1,396		-1,887	-0,492	
1-month _{1,0,1}	0,189	0,012	0,151	-0,042		-13,346	-1,657		-0,251
1-month _{0,1,1}	0,735	-0,012	0,095		-0,002	-40,328		-0,001	-0,786
1-month _{1,0,0}	0,070	0,020	0,139	-0,044			-1,828		
1-month _{0,1,0}	0,986	-0,009	0,043		0,049			0,018	
1-month _{0,0,1}	0,431	-0,003	0,095			-40,327			-0,790

Dependent Variable	P-value (regression)	adj. R ²	constant	In SumRepetitions (s)	In Like/Views (l)	In Dislikes/Views (d)	t-ratio (In s)	t-ratio (In l)	t-ratio (In d)
ln 2-day _{1,0,1}	0,037	0,081	-9,384	0,549		-0,914	2,053		-2,290

Table 5. Stock Tickers

This table displays the different stocks that were recommended in the videos.

1 = '\$NERV'	11 = '\$DGLY'	21 = '\$CLVS'	31 = '\$OPTN'	41 = '\$VSTO'	51 = '\$EDUC'	61 = '\$TLGT'	71 = '\$SPH'
2 = '\$ZM'	12 = '\$WORK'	22 = '\$GME'	32 = '\$KDMN'	42 = '\$SOLO'	52 = '\$SWBI'	62 = '\$IZEA'	72 = '\$NCTY'
3 = '\$NBRV'	13 = '\$XSPA'	23 = '\$RAD'	33 = '\$LRN'	43 = '\$HYRE'	53 = '\$ARLO'	63 = '\$SINT'	73 = '\$TLRY'
4 = '\$CHMA'	14 = '\$UONE'	24 = '\$RVLV'	34 = '\$COCP'	44 = '\$DS'	54 = '\$CETX'	64 = '\$DLOC'	74 = '\$TRNE'
5 = '\$KPTI'	15 = '\$GNUS'	25 = '\$MAXR'	35 = '\$DMAC'	45 = '\$INKW'	55 = '\$JMIA'	65 = '\$MARA'	75 = '\$SRAC'
6 = '\$EVOK'	16 = '\$ETON'	26 = '\$NEXCF'	36 = '\$GEVO'	46 = '\$GMBL'	56 = '\$SPWH'	66 = '\$DFFN'	76 = '\$STNG'
7 = '\$EPZM'	17 = '\$AQST'	27 = '\$CIIC'	37 = '\$PLG'	47 = '\$MITT'	57 = '\$IDEX'	67 = '\$KTOV'	77 = '\$DHT'
8 = '\$AKTX'	18 = '\$MESO'	28 = '\$NAT'	38 = '\$CDAK'	48 = '\$PRTY'	58 = '\$ITI'	68 = '\$HCAC'	78 = '\$PLNHF'
9 = '\$ENDP'	19 = '\$FBIO'	29 = '\$SQNS'	39 = '\$GTHX'	49 = '\$SHIP'	59 = '\$BYFC'	69 = '\$EVFM'	79 = '\$NNDM'
10 = '\$MRNA'	20 = '\$LPCN'	30 = '\$KRMD'	40 = '\$GAN'	50 = '\$BGFV'	60 = '\$MVIS'	70 = '\$LCA'	80 = '\$APXT'

