Evidence of Predictable Behavior of Security Returns: A Partial Replication of Jegadeesh [1]

Stephanie Li and Sohum Patnaik

January 28, 2015

[1] The data packages used were provided by David Kane, a portfolio manager at Hutchin Hill Capital. David Kane, Yuanchu Dang, and the Econ 20 class all provided us with incredible support during the process of making this paper. The code used to calculate the results can be found on GitHub at https://github.com/sohumpatnaik/finalProject.

Abstract

In this paper we perform a partial replication of Jegadeesh's paper (1990) by looking for evidence that past total return data is able to predict future behavior of individual securities. We first split up the top 1500 stocks in our core data set into five quintiles by cap size and find a positive serial correlation for monthly lagged returns (13.04%) as well as for twelve-month lagged returns (14.02%). This correlation is affected by unusual trading activity in January. We ran a backtest to find a pattern between past returns and future returns and found that individual securities that had high returns in the past had the highest returns in the future. Finally, we extended our replicated findings by running another backtest to find a pattern between adjusted volumes and future returns, finding an inverse relationship. The differences in our findings and those of Jegadeesh's (1990) can be explained by increased market efficiency since 1990 and the greater liquidity of stocks in our data set compared to those of the data set he used.

Introduction

During the 1970s, the efficient market hypothesis became the popular opinion of the stock market by “academics, many investment professionals, and corporate managers” (Gray and Carlisle 9). The hypothesis claims that the market is always efficient because all information about companies is public, and thus the price and value of stocks cannot be strategically manipulated to form a portfolio that earns higher returns than a portfolio with simply a random group of stocks. Two of the most famously successful investors of the late twentieth century, Warren Buffett and Edward Thorp, have investigated and rejected this hypothesis. Academics have also presented evidence with mixed results of the validity of market efficiency.“Lo and MacKinley (1988) reported positive serial correlation in weekly returns”, while “French and Roll (1986) reported significant negative serial correlation in daily returns” (Jegadeesh 881). These early papers paved the way for further examination of serial correlation of past returns and future returns.

The paper we replicated, *Evidence of Predictable Behavior in Security Returns* by Narasimhan Jegadeesh, does just this but over longer intervals of time (using one-month lagged returns and twelve-month lagged returns). He finds a significant negative first-order serial correlation for monthly returns, but a “particularly strong” positive first-order serial correlation for twelve-month lags (Jegadeesh 882). This was one of the first papers to provide evidence of momentum in the stock market, and more broadly, the ability to use a quantitative edge to beat the market. It shows that when an individual security has positive returns in the past twelve months, it is likely to continue to have positive returns in the future. On the other hand, Jegadeesh (1990), French and Roll (1986), and Lo and MacKinley (1988) showed collectively that returns using data from shorter periods of time (monthly, daily, and weekly data) are not reliable predictors of future return behavior since no consistent positive or negative correlations were found.

In our paper, we prove that momentum exists in the stock market by finding positive serial correlations over one-month and twelve-month lag periods and analyzing their economic significance by total returns.

Data and Methods

For our analysis we used the ws.data, secref, and yearly data packages provided by David Kane of Hutchin Hill Capital that detailed US stock information of over 3000 large cap companies from 1998-2007. We cleaned these data packages by first trimming the data to only include stocks in the top 1500 market cap sometime during 1998-2007. These stocks were most influential to market behavior and more worth investigating than the stocks not in the top 1500 market cap, particularly due to their higher liquidity. We then looked at data for each year separately by plotting the standard deviation of total daily returns and also analyzing the recorded top total returns in the data. With these two metrics, we eliminated companies with a total daily return greater than 200% that the plots and data clearly indicated. This is because a total recorded return greater than 200% is so unlikely and abnormal that the data was likely recorded incorrectly. Finally, we removed company data for companies that did not have data for the full 1998-2007 time period (implying bankruptcy in the middle of the period or later entry into the market). The remaining data comprised the core data we used for our analysis.

Just like in the original paper, we split the stocks into five different groups based on size. The smallest firm stocks are in Group 1, while the largest firm stocks are in Group 5. We sorted these stocks into quintiles by using the market cap metric. According to Table 1, it is clear that the total returns over all lagged time intervals increased as firm size increased. This indicates that large companies that have previously had high total returns, will continue to generate high total returns. Smaller companies show lower total returns which indicate that they are not making returns at as high of a rate or volume as larger companies do and also that more negative returns may exist and more significantly affect the overall total returns.

It is also important to isolate January in our findings and compare the total returns found using all months of the year to those using only February through December. This is because of the January Effect that involves a higher volume of securities being bought (and thus higher prices and total returns) in January compared to any other month, since stockholders want to sell off stocks in December to exaggerate tax losses for the year. Although the January Effect is less prominent now than it was when the original paper was published in 1990 since markets have adjusted to the trend (become more efficient), it is still interesting to see the relatively less manipulated returns of data using only February through December.

Results

Using the data from Table 1, we now recognize the trends previously mentioned more quantitatively and in comparison to what we would have expected from Jegadeesh's paper (1990). While Jegadeesh (1990) reported a significant negative serial correlation in monthly stock returns, we found a positive correlation in monthly stock returns with an average correlation of 13.04% across all five quintiles. The correlation increased as cap size increased with a 47.20% increase from Q1 to Q5. Furthermore, while Jegadeesh (1990) found a significant positive serial correlation for twelve-month lagged returns, we found a positive correlation of 14.02% for twelve-month lagged returns, only 8% higher than the monthly lagged return correlation value. This correlation is not strong since it is nowhere close to 100%, and the correlation did not increase much even when using longer lags. However, the correlation did increase as cap size increased (Q1 to Q5) at an average of 52.25% across all lag periods. This indicates that security returns of larger companies can be better predicted using past returns than that of smaller companies.

When looking at Table 2, which measures the same correlations across all five quintiles but only using January data from 1998-2007, the average increase in correlation from Q1-Q5 is 101% (much higher than 52.25% using Table 1 data and 50% using Table 3 data). The average correlations for monthly returns and twelve-month lagged returns using just January data (.049 and .0672, respectively) were also much smaller than the average correlations using January through December data (.1304 and .1402, respectively). This suggests that the January data is a worse predictor of future returns than yearly data and even February through December data (average correlations .1364 and .1446, respectively).

Table 3 contains similar data to Table 1 (with no unusual trends like in Table 2) which is to be expected. The January Effect does not affect the data as greatly as it did in the original paper, but it still exists to a noticeable extent.

	1 month	3 month	6 month	12 month
Q 1	10.57%	11.31%	11.59%	11.97%
Q 2	12.84%	12.07%	12.46%	12.01%
Q 3	12.95%	13.08%	13.27%	13.46%
Q 4	13.27%	13.67%	13.95%	14.66%
Q 5	15.57%	17.66%	18.06%	17.93%

Table 1: The stocks in the data set are sorted into quintiles (Q1-Q5) based off of their cap size and the correlations between the lagged returns and current returns are found (using Jan-Dec data). The correlations are displayed in the table.

	1 month	3 month	6 month	12 month
Q 1	4.09%	4.56%	4.16%	5.2%
Q 2	2.68%	6.2%	4.32%	8.59%
Q 3	5.56%	3.01%	3.86%	5.55%
Q 4	3.59%	4.56%	4.05%	4.04%
Q 5	8.47%	9.7%	7.86%	10.28%

Table 2: The stocks in the data set are sorted into quintiles (Q1-Q5) based off of their cap size and the correlations between the lagged returns and current returns are found (using Jan data). The correlations are displayed in the table.

Forming a Portfolio

Since there is serial correlation among all size-based groups of stocks, we will construct portfolios to analyze the economic significance of the correlation. We want to see if there is a relation between the past total returns of a stock and the future total returns of the stock, similar to what Jegadeesh (1990) did. To build our portfolios, we look at the total returns with a one-month,

	1 month	3 month	6 month	12 month
Q 1	11.02%	11.78%	12.13%	12.43%
Q 2	13.66%	12.64%	13.1%	12.53%
Q 3	13.49%	13.77%	13.96%	13.87%
Q 4	13.96%	14.3%	14.68%	15.33%
Q 5	16.01%	18.17%	18.63%	18.25%

Table 3: The stocks in the data set are sorted into quintiles (Q1-Q5) based off of their cap size and the correlations between the lagged returns and current returns are found (using Feb-Dec data). The correlations are displayed in the table.

three-month, six-month, and twelve-month lag. There will be a portfolio for each time interval and each portfolio will be split into ten groups, P1-P10, which are sorted by past total returns. P1 comprises the group of lowest total returns with the specified month lag and takes a short position, and P10 comprises the group of highest total returns with the specified month lag and takes a long position. We will compare P1 with P10 for each portfolio, as well as P1 and P10 across all four portfolios.

After sorting the data using past returns to make ten portfolios, we find the forward one-month, three-month, six-month, and twelve-month total return data. A natural backtest is run to find the total returns in the past and forward one, three, six, and twelve months and is rebalanced for each time period. We find the average of the one, three, six, and twelve month forward returns per month, and the mean of all the results can be seen in Table 4.

Portfolio Performance

The results presented in Table 4 show a clear pattern between the previous one-month, three-month, six-month, and twelve-month total returns, and the future one-month, three-month, six-month, and twelve-month returns. The higher groups generally outperform the lower groups. As the time increased (amount of past data we used), the difference between the extreme deciles grew. The average difference between the extreme deciles was .24% for 1 month, .94% for 3 months, 1.77% for 6 months, and 3.77% for twelve months. The graphs (Figure 1, Figure 2, Figure 3, and Figure 4) also show that the highest group consistently outperforms lower groups, resulting in a cumulative return for the highest group that is larger than all of the other

	1 month	3 month	6 month	12 month
Group 1	0.22%	0.63%	1.75%	3.91%
Group 2	0.24%	0.51%	1.39%	2.73%
Group 3	0.25%	0.53%	1.32%	2.78%
Group 4	0.21%	0.47%	1.36%	2.85%
Group 5	0.24%	0.58%	1.3%	3.37%
Group 6	0.21%	0.67%	1.6%	3.46%
Group 7	0.22%	0.71%	1.79%	3.84%
Group 8	0.29%	0.82%	1.98%	4.24%
Group 9	0.31%	1.07%	2.33%	4.99%
Group 10	0.45%	1.57%	3.52%	7.68%

Table 4: Four backtests are run to look for a relationship between a stock's previous total returns and a stock's forward returns (use Jan-Dec data). Stocks are split into 10 groups (1-10) based off of increasing past total returns. The mean future total returns are displayed in the table. Each column shows values in which stocks were split based off of their previous specified month returns, and the mean of the forward specified month returns are displayed.

groups. This tells us that there is a relationship between the previous total returns and the future total returns of a stock. We see a core idea of momentum investing: stocks that have done well in the past will continue to do better than stocks that have done poorly in the past.

Extension

In addition to replicating methods and strategies used by Jegadeesh (1990), we extended his findings by using a new strategy and analyzing its predictive attributes for security returns. We ran a backtest to look for patterns between the adjusted volume of a stock and its future one-month, three-month, six-month, and twelve-month total returns. Like the backtest we used to obtain the results for Table 4, we similarly used a natural backtest for this procedure, so the stocks in our dataset were split into ten groups, with the stocks with the lowest volume being placed in the lowest decile and stocks with the highest volume being placed in the highest decile, and a long position was taken on the stocks in the highest group and a short position was taken on the stocks in the lowest group. The mean one-month, three-month, six-month, and twelve-month future total returns of the stocks which have been sorted based off of their adjusted volumes can be seen in Table 5.

The results were interesting to analyze since the stocks with greater volumes (in the higher groups) generally performed more poorly than the stocks with lower volumes (in the lower groups). This suggests that stocks with a higher volume are more efficient; in other words, it is harder to take advantage of inefficiencies in the market to select stocks with a high volume. It is also interesting to note that as thetime interval increases, the difference in total returns between the extreme deciles of stocks based off their volume increases. The average difference between the extreme deciles was .32% for one month, .94% for three months, 1.91% for six months, and 3.53% for twelve months. We have seen that the volume of a stock can be a powerful predictor of how the stock will perform in the future relative to other stocks, as volume is inversely related to future returns.

	1 month	3 month	6 month	12 month
Group 1	0.4%	1.16%	2.27%	4.9%
Group 2	0.23%	0.78%	1.42%	3.24%
Group 3	0.24%	0.75%	1.38%	3.2%
Group 4	0.25%	0.75%	1.44%	3.32%
Group 5	0.3%	0.79%	1.45%	3.49%
Group 6	0.28%	0.83%	1.65%	3.66%
Group 7	0.31%	0.83%	1.63%	3.74%
Group 8	0.26%	0.83%	1.47%	3.57%
Group 9	0.2%	0.59%	1.15%	2.95%
Group 10	0.08%	0.22%	0.36%	1.37%

Table 5: Four backtests are run to look for a relationship between a stock's adjusted volume and a stock's forward returns (use Jan-Dec data). Stocks are split into 10 groups (Group 1-Group 10) based off of the increasing adjusted volumes of the stocks. The mean future total returns are displayed in the table. Each column shows values in which stocks were split based off of their adjusted volume, and the means of the forward specified month returns are displayed.

Possible Explanations

Our data yields trends that are not as pronounced as those in the original paper. For instance, while Jegadeesh (1990) found a significant positive twelve month serial correlation, we only found a positive twelve-month serial correlation that was 8% higher than that of the monthly correlation. Our mean future total returns with each specified lag increased at a slightly smaller rate over each increasing decile than that in the original paper. This can be explained by two factors. The first is that the market has become increasingly efficient in the past 25 years since the original paper was written, and it is more difficult to choose stocks at a price that deviate significantly from their value. Furthermore, the stocks we are looking at (top 1500 stocks in the ws.data) are much more liquid than the stocks looked at in the original paper which included penny stocks and other stocks of companies with questionable reliability.

Conclusion

This paper shows that returns of stocks can be predicted. We replicated Jegadeesh's findings (1990) and followed strategies similar to the ones he used in his paper. The results show a correlation between the cap size of a stock and its one-month, three-month, six-month, and twelve-month lagged return. An unusual pattern can also be seen if we isolate January in that test, which is interesting to note due to the January effect. These results were seen by splitting the stocks we gathered from ws.data, sorting the stocks into quintiles based off their cap size, and finding the average one-month forward returns of those quintiles. The stocks with a higher cap size clearly performed better than the stocks with lower cap sizes across all lag intervals.

Similarly, we have also seen that the total returns that a stock has had previously affects the total returns that the stock will have in the future. We ran a backtest on the data we gathered from ws.data, which split the stocks into deciles based off their one-month, three-month, six-month, and twelve-month previous returns. We then analyzed the total returns of those stocks in the future one, three, six, and twelve months, and found a clear pattern where stocks that had outperformed other stocks in the past continued to outperform those same stocks.

Likewise, we also extended the paper Jegadeesh had written to see if there was a pattern between the volume of a stock and the one, three, six, and twelve month returns of a stock. We found that there is an inverse relationship between the volume of a stock and the returns a stock has in the future. The lower the volume of a stock, the better it performs in the future. This may be because stocks with high volumes have more liquidity and are more efficient than stocks with low volumes, so stocks with high volumes have less inefficiencies to take advantage of. Value investing is thus more difficult, since the price of the stock more closely reflects its value.

With the findings in this paper, we can reject the notion that stocks perform randomly, for we have shown that the performance of a stock can be predicted.

References

[1] Enos, J., Kane, D., Campbell, K., Gerlanc, D., Schwartz, A., Suo, D., Colin, A., Zhao, L. (2013), backtest: Exploring portfolio-based conjectures about financial instruments, R package version 0.3-2. URL http://CRAN.R-project.org/package=backtest .

[2] Jegadeesh, Narasimhan, 1990, Evidence of Predictable Behavior of Security Returns, *The Journal of Finance*, 45, 891-898.

[3] French, Kenneth R. and Richard Roll, 1986, Stock return variances: The arrival of information and raction of traders, *Journal of Financial Economics*, 17, 5-26.

[4] Lo, Andrew W. and A. Craig MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies*, 1, 41-66.

[5] Gray, Wesley R., and Tobias E. Carlisle. *Quantitative Value a Practitioner's Guide to Automating Intelligent Investment and Eliminating Behavioral Errors Website.* Hoboken, N.J.: Wiley, 2013.