Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
NTNU Business School

Ørjan Mikal Antonsen

The sales seasonality premium

A replication of Gustavo Grullon, Yamil Kaba, and Alexander Nùñez-Torres's seasonality study: "When Low Beats High: Riding the sales seasonality premium."

Master's thesis in Financial Economics
Supervisor: Snorre Lindset
June 2021

NTNU
Norwegian University of
Science and Technology

Ørjan Mikal Antonsen

The sales seasonality premium

A replication of Gustavo Grullon, Yamil Kaba, and Alexander Nùñez-Torres's seasonality study: "When Low Beats High: Riding the sales seasonality premium."

Master's thesis in Financial Economics
Supervisor: Snorre Lindset
June 2021

Norwegian University of Science and Technology
Faculty of Economics and Management
NTNU Business School

NTNU
Kunnskap for en bedre verden

Preface

The thesis concludes my master's degree in financial economics at the Norwegian University of Science and Technology. I want to thank my supervisor Snorre Lindset, for providing me with solid feedback and guidance throughout the writing process of this thesis. I also want to thank my study colleagues during my years as a student.

Trondheim, June 2021.

Ørjan Mikal Antonsen

Abstract

In this thesis, I replicate the research by Grullon et al. (2020), who find seasonal patterns in stock returns to be counter-seasonal. Applying the same long-short investment strategy as them, I also conclude this in the thesis. With an annual excess return of 7.5% when buying firm stocks in low season and short them in high season, based on a period from 1972-2017. Further testing shows that the effect is more substantial amongst larger firms. Testing the volatility shows that the firms which have low and high season are the ones who create the seasonality effect, compared to firms with stable sales. The seasonality effect also increases through the decades. Using a factor created on sales by Grullon et al. (2020) shows that the sales effect has a high Sharpe ratio and a low standard deviation. Removing other seasonality effects shows that this seasonality effect is not related. Lastly, I test some of the hypotheses by Grullon et al. (2020) in order to find an explanation for the premium. Such as real option and leverage theory, where I confirm that they are some part of the explanation, similar to the findings of Grullon et al. (2020).

Sammendrag

I denne oppgaven replikerer jeg forskningen av Grullon et al. (2020), som finner at sesongmønstre i avkastningen er mot syklisk. Ved å bruke den samme lang-kort investeringsstrategien som dem, konkluderer jeg også dette i oppgaven. Med en årlig meravkastning på 7,5 % når man går lang i aksjer i lavsesong og kort i aksjer i høysesong, basert på en tidsperiode fra 1972-2017. Videre testing viser at effekten er sterkere blant større selskaper. Testing av volatilitet viser at firmaene som har mest svingninger i lav og høysesong er de som skaper sesongeffekten, sammenlignet med firmaer som har stabilt salg ut året. Sesongeffekten blir også sterkere det siste tiåret. Ved å bruke en faktor basert på salg av Grullon et al. (2020) som viser at salgseffekten har en høy Sharpe og en lav standardavvik. Å fjerne andre sesongeffekter viser at denne sesongeffekten ikke er relatert til disse. Til slutt tester jeg noen av hypotesene av Grullon et al. (2020) for å finne en forklaring på premien, som real opsjon og leverage teori, der jeg bekrefter at de er en del av forklaringen, i likhet med funnene til Grullon et al. (2020).

Contents

1 Introduction

Several studies have attempted to explain the seasonal effect on stock returns. A seasonal effect, in short, is an attempt to find periodic patterns in a company's stock returns based on the seasons. Heston and Sadka (2008) find evidence of seasonality effect through stock returns for the previous year, while Chang et al. (2016) points to earnings announcement and previously announced quarterly results. Both papers find a seasonality effect and evidence that the effect is not counter-seasonal. Unlike the other papers, Grullon et al. (2020) uses sales to show that firms in low sales season beat firms in high sales season in stock returns. The sale is how much sales a firm generates in a period. A firm with a steady sale will generate 25% of its sales quarterly (100% annually). Grullon et al. (2020) show that going long position in firms that are in low sales season and going short position in firms that are in a high sales season yields an excess return of 8.4% annually from 1972-2016 when risk adjusting with the Fama-French five-factor model. Their result shows that the seasonality effect is counter-seasonal. The reason for choosing sales to find the seasonality effect is because Grullon et al. (2020) states that it is beyond the firm's control to affect this variable, contrary to the Heston and Sadka (2008) and Chang et al. (2016). Furthermore, they also test several hypotheses for this phenomenon. Their main results point towards the real options theory, leverage, and investors' inattention as the leading causes.

In this thesis, I replicate the work of Grullon et al. (2020), I use the same methodology and robust testing. Due to time and data limitations, I am not able to replicate all of their results. As in the original study, I confirm that firms' low sales season beats firms in the high sales season in return. Using value-weighted returns when testing for the capital asset pricing model, the annual excess return is 7.5%. The excess return controlling for the Fama-French three-factor model is 8.784% annually. Finally, controlling for the Fama-French five-factor model, the excess return is 8.172%. Further, I am testing for the firm size effect to see if small-capitalization firms do not cause the seasonality effect since the return of the value-weighted portfolios is higher than for the equal-weighted.

Creating 5x5 portfolios, sorting first for size, and then within those five portfolios, sorted again on sales, I find that the effect is more significant among larger firms to rule out the small firm effect. I then test whether the firms with the highest oscillation in the sales variable are the firms that generate the excess returns, which is what the results are suggesting. Next, the portfolios are sorted into different decades and compared with the market capitalization changes of the firms, which shows that the excess return of sales seasonality increases with firm size through the decades. The highest excess return and firm size are in the last decade from 2008-2017 with a return of 16.08%. By using a factor for sales created by Grullon et al. (2020), I test if there is a correlation between the Fama-French factors. According to the results, there is no correlation. The sales factor outperforms the Fama-French factors when testing the mean, standard deviation, and Sharpe ratio. Including additional controls in a Fama-Macbeth regression again shows that the small firm effect is not present and the seasonality effect becomes more substantial. Then following Grullon et al. (2020), to show that the results are not affected by the seasonality effect from Heston and Sadka (2008), I remove their seasonality effect from the sample to test whether it affects the sales seasonality excess return. The same goes for the earnings announcement effect by Chang et al. (2016). Both tests show that the sales seasonality effect is still present after removing their seasonality effect from the sample. Finally, as by Grullon et al. (2020), I attempted to find out why this premium exists. Using proxies for investment, leverage, and financial variables such as size, book-to-market, profitability, firm fixed effects, year dummies, and the firm's advancement for the next season. The result shows that the real option and leverage effect are in play as in Grullon et al. (2020). I omitted Investor inattention from the empirical research of this thesis because of limitations to time and data, and further research would be to try and replicate this part of the study by Grullon et al. (2020).

2 Literature review

In this chapter, I present different literature behind seasonal patterns in stock returns.

Heston and Sadka (2008) show that a firm's return cycle is the same every month of the year, for each year. For example, if a firm has a high return in the summer, then it will most likely have a high return the next summer, and so on. Their results are also not counter-seasonal, unlike Grullon et al. (2020)'s method. However, the important difference between these studies is that Heston and Sadka (2008) uses a measure based on the return to test for the seasonal effect.

Chang et al. (2016), use the earning announcement test to test for seasonality effects. They argue that stock prices have higher returns when they are in the earnings announcement season if they have historically high returns during this period. The method Chang et al. (2016) uses to test for the seasonality effect is to construct portfolios where they go long in high season stocks and short in low season stocks, unlike Grullon et al. (2020), who goes short on the high season stock and long on the low season stock.

The original paper by Grullon et al. (2020) attempts to test the seasonality premium on stocks, and their findings show that the optimal seasonality strategy is to go long on firms that are out of season and short firms in season. Unlike other seasonal studies that show the opposite in comparison regarding seasonality strategy, they then test different hypotheses behind the reason for their finding. Their results are based on the firms' sales variable, arguing that they use sales because it is outside the firm's control and cannot directly affect it. Thus, making it the dominant variable when sorting for sales seasonality (Grullon et al., 2020).

Grullon et al. (2020) sorted into monthly portfolios from 1972 to 2017, and their results lead to an annual excess return of 8.4% when adjusting for risk using both the Capital Asset Pricing Model and Fama-French three and five factors models. They also show that the annual excess return has increased over the decades, with the annual excess return in the first decade of 1978-1988 being 8.4%. Compared to the last decade from 2008-20017, the annual excess return is 14.52%. They use volatility to show that the excess return generated is more robust for firms that fluctuate the most in quarterly sales. Moreover, their results contradict the small firm effect theory, showing that the premium is more substantial for larger firms. Furthermore, Grullon et al. (2020) looks for evidence that other factors do not also explain sales seasonality, and their results are that the sales seasonality effect has the highest Sharpe ratio and low standard deviation. To be sure that the previously mentioned seasonality effects do not affect returns, they remove both the seasonality effect of Chang et al. (2016), and that of Heston and Sadka (2008) from the sample, with Grullon et al. (2020) finding that their seasonality effect is unrelated and still present.

Finally, Grullon et al. (2020) attempts to explain this economic phenomenon by testing several hypotheses. The first theory they test is Berk et al. (1999)'s real options theory, according to which returns should be lower during high season because they use the growth to invest in assets. Their results show that this is part of the explanation. The next theory is that the firm reduces leverage during the high season, which reduces systematic risk, for which the results also provide evidence. The last and final theory comes from the model by Merton (1987), which states that when the market neglects stocks, they should yield a higher return. The reason is that investors who hold these stocks are exposed to higher idiosyncratic risk because their portfolio will have an overweight of these stocks. Therefore, a higher risk premium is required. Finally, Grullon et al. (2020) uses proxies to test investors' level of attention to stocks to see if investors undervalue (overvalue) stocks that have seasonally low (high) sales. The result by Grullon et al. (2020) points towards the Merton (1987) Incomplete Market Theory and low investor attention during the low-season.

3 Theory

Grullon et al. (2020) tests several hypotheses to try to explain the high excess return that results from seasonal sorting. In this chapter, I present the theory behind the hypothesis.

Starting with the small firm effect, Banz (1981) tests returns using the capital asset pricing model developed by Sharpe (1964). In the paper, Banz (1981) finds a correlation between firm size and returns. This finding is consistent with the risk-return trade-off theory formalized by Markowitz (1952), in which he states that the only way to earn additional return is to take additional risk. Thus, according to Banz (1981), since size represents a firm's risk, a small firm should expect an additional return. Therefore, Grullon et al. (2020) tests for the possibility that the small firm effect drives the excess returns. Their result and this thesis conclude that this is not the case by constructing portfolios based on size and sales seasonality.

The following effect Grullon et al. (2020) tests is the multiplier effect, which states that investors demand higher returns when firms are in high season due to the increase in systematic risk during that time. Both this thesis and Grullon et al. (2020) results contradict this theory, suggesting that other factors drive the premium.

One explanation for this phenomenon may be that firms increase their assets and reduce their leverage during high season. Therefore, Grullon et al. (2020) tests Berk et al. (1999)'s real options theory, which states that as firms use their growth to invest in assets during high sales season, their systematic risk decreases, and so do their returns. Grullon et al. (2020) also tests for the leverage effect, where an increase in leverage leads to higher risk and premium. This thesis shows that this is counter-seasonal, where firms use their income during the high season to reduce leverage, thus reducing risk and expected return during that period. Evidence after testing these two effects points to this as a possible explanation for the seasonality premium in both this thesis and Grullon et al. (2020) study. Meaning that the leverage and real option effect shows that investing in assets and reducing risk can produce a counter-seasonal effect, but the results are not large enough to explain all of the seasonality effects on the returns, only some.

4 Method

In this chapter, I present the methodology that I have used in order to replicate the study by Grullon et al. (2020).

4.1 Portfolio formation

Following Bali et al. (2016) I explain the theory behind the portfolio formation.

4.1.1 Univariate portfolio formation

Univariate portfolio construction means sorting portfolios by a variable X attempting to explain the cross-sectional relationship in the outcome variable Y [1]. After the four steps in Bali et al. (2016), the first step begins with the calculation of breakpoints. Breakpoints are the values that determine what portion of the sample, based on the sort variable, goes into each of the portfolios. The breakpoints are periodic and will be different for each period, but to show an example, the table below shows the mean breakpoints from 1972 to 2017. So for the first portfolio, all firms with a sales variable below 17.5% are sorted into the first portfolio. Then for the second portfolio, all firms

[1] This method cannot control for other effects (Bali et al., 2016).

with sales between 17.5% and 21.9% are sorted in, and so on. Each month I calculate new breakpoints, and the sample is sorted again by the variable X.

TABLE 1: Mean breakpoints created from sales seasonality in the sample. The sample uses non-financial stocks from CRSP starting January 1972 to December 2017.

Portfolio	1	2	3	4	5	6	7	8	9	10
Mean_sales_breakpoints	0.175	0.219	0.231	0.239	0.245	0.251	0.257	0.264	0.277	0.333

For the first portfolio, I am placing variables X below the lowest breakpoint. For the second portfolio, variables X are placed between the first and second breakpoints, and so on. Using n_p as the number of portfolios, then the number of breakpoints needed depends on the number of portfolios, hence the number of breakpoints is given by $n_p - 1$. From Bali et al. (2016), the value of the breakpoints k for period t is given by

$$B_{k,t} \; for \; k \in \{1, 2, ..., n_p - 1\}. \tag{1}$$

Since the breakpoints are to be determined by the percentage of the cross-sectional distribution of the variable X at time t, we can let p_k be the percentage that determines the kth breakpoint. It means that the breakpoints are calculated as the percentage of the sort value in the sample at time t. Thus, the breakpoint is given by

$$B_{k,t} = Pctl_{p_k}(\{X_t\}), \tag{2}$$

where $Pctl_{p_k}(Z)$ is the percentage of Z and the X is the sort variable at time t (Bali et al., 2016). For the univariate portfolio in this study, we divide the number of portfolios into ten, as this replicates the study by Grullon et al. (2020). It means that the portfolio breakpoints are at 10, 20, 30,...,90 percent of the sort variables. Most studies split into two, five, or ten portfolios because the number of portfolios increases the amount of noise when trying to calculate the true mean of the sample (Bali et al., 2016). In this study, the portfolio with the highest number of shares has 273,307 shares over 552 months, where the lowest has 272,223 shares. Therefore, I can argue that sample size is not a problem in this study when using ten portfolios.

Now that the first step about breakpoints is covered, the second step is about portfolio formation. Bali et al. (2016) shows that in each period, stocks get sorted into portfolios based on the sort value X relative to the breakpoints B. Then the k portfolio at time t can be defined as

$$P_{k,t} = i | B_{k-1} \leq X_{i,t} \leq B_{k,t}, \tag{3}$$

where i is the number of entities of the sort value $X_{i,t}$ that the portfolio k holds in time period t. B_{k-1} is the lowest breakpoint and $B_{k,t}$ is the highest. The observations with the lowest sorting variable X from the sample are placed in the first portfolio and those with the highest value are placed in the last portfolio. The value of X increases with the number of portfolios (Bali et al., 2016).

The third step is to compute the average value variable Y for each portfolio in each period t. This value is defined as

$$\bar{Y}_{k,t} = \frac{\sum_{i \epsilon p_{k,t}} W_{i,t} Y_{i,t}}{\sum_{i \epsilon p_{k,t}} W_{i,t}}, \tag{4}$$

where the variables are weighted against another variable W, most commonly market capitalization is used for weighting. However, if the portfolios are not weighted but equal weighting is used, then $W_{i,t} = 1 \forall i, t$, which is the same as weighting each sample $\frac{1}{n_t}$. Therefore, for the equally weighted portfolio, this can be rewritten as follows

$$\bar{Y}_{k,t} = \frac{sum_{i \epsilon p_{k,t}} W_{i,t} Y_{i,t}}{\sum_{i \epsilon p_{k,t}} W_{i,t}} = \sum_{i}^{N} Y_{i,t} * \frac{1}{n_t} \; if \; W_{i,t} = 1 \forall i, t. \tag{5}$$

An important factor is that there could be missing values of Y in most cases and according to Bali et al. (2016), the mean should be calculated for the places where the variables are present in the sample. Bali et al. (2016) also states that equal weighting sorting is to be used to capture the phenomena for the average stock and that the value-weighted portfolio is more appropriate when it comes to stocks. Since the returns of the equal-weighted portfolio are most likely driven by the small firms as we know from Banz (1981) that they are riskier and should provide higher returns. The low cap stocks are also more expensive to trade, while the high cap stocks have more liquidity and are therefore less expensive. Normally, the equally weighted portfolios cannot realize the returns because of such high transaction costs. Therefore, the value-weighted portfolio is the most important in this paper. Replicating the study by Grullon et al. (2020), another portfolio also must be calculated because this is the main subject of this research. The difference between the two portfolios is the difference portfolio. In this study, it is the difference between the low-sales portfolio and the high-sales portfolio. From Bali et al. (2016), the differential portfolio is estimated as

$$\bar{Y}_{diff,t} = \bar{Y}_{n_{p,t}} - \bar{Y}_{1,t}. \tag{6}$$

The last step consists of interpreting the cross-sectional relationship between the sort variable X and the outcome Y. The hypothesis test is to show whether the calculated mean is statistically different from zero. If the variable is not statistically insignificant, a cross-sectional relationship exists between the sort variable X and the outcome. Researchers often reject a significance level above 5%, so in this study, the hypotheses that are not significant at the 5% significance level are also rejected (Bali et al., 2016).

4.1.2 Bivariate-dependent portfolio formation

The next part of this thesis is explaining how to create bivariate-dependent portfolios. Different from the univariate, where it is sorting on one variable X to explain the outcome Y, the bivariate sorts on two variables X_1 and X_2. X_1 is only used as a control variable and the main goal is to explain the relationship between X_2 and Y (Bali et al., 2016).

There are now two dependent variables from the previous one, but the four steps are in the same order. But, when calculating the breakpoints it is now important to decide and separate between the control variable X_1 and the test variable X_2. The first sort variable is calculated is equal to the univariate, where n_p is the number of portfolios and $n_p - 1$ is the number of breakpoints. Then the first breakpoints are defined as

$$B1_{j,k,t} = Pctl_{p_k}(\{X_t\}), \tag{7}$$

where the j indicates the group sorted by variable X_1 and k for the second variable X_2. Now that the breakpoints are created, for each amount of n_{p1}. The next part is to create the second breakpoints based on the value of X_2 and this is done for each of the portfolios n_{p1} separately. Therefore, inside each of the portfolios created and sorted by the previous breakpoints, new breakpoints are made. This means that for each group of n_{p1} there are $n_{p2} - 1$ breakpoints. From Bali et al. (2016) the second sort value is therefore defined

$$B2_{j,k,t} = Pctl_{p2_k}(\{X_{2t}|B1_{j-1,t} \leq X_{1t} \leq B1_{j,t}\}). \tag{8}$$

A quick note is that not all research uses a full sample of the sorting variable when creating the breakpoints. For instance, Grullon et al. (2020) uses only NYSE stock when creating size breakpoints and full sample when they create the breakpoints for sales seasonality. This is done in the bivariate-portfolio formation sorting for size and sales seasonality, testing for the small firm effect. Doing this is no problem as long as there are enough variables in X_1 to form n_{p2} portfolios when sorting for the second variable X_2 according to Bali et al. (2016).

The next step is to form the portfolios by first sorting the sample into portfolios based on the first breakpoint using sort variable X_1. Then in each of those portfolios, re-sort again based on the second breakpoints using the second sort variable X_2. For instance, if we want to create 2x3 portfolios, the first breakpoint is calculated to create the first two portfolios. Then for the first portfolio sorted on X_1, two new breakpoints are made, creating another three portfolios in the first, sorted on the second variable

X_2. Effectively creating six portfolios, or $n_{p1} * n_{p2}$ portfolios as from Bali et al. (2016). Further, they define the k portfolio at time t

$$P_{j,k,t} = \{i|B1_{j-1,t} \leq X1_{i,t} \leq B2_{j,t}\} \cap \{i|B2_{j,k-1,t} \leq X2_{i,t} \leq B2_{j,k,t}\}. \tag{9}$$

The average portfolio values, as from before, can be calculated equal or value-weighted. Therefore, the equation from Bali et al. (2016) for the average value of the outcome variable for the portfolio $P_{j,k,t}$ is

$$\bar{Y}_{j,k,t} = \frac{\sum_{i \in p_{j,k,t}} W_{i,t} Y_{i,t}}{\sum_{i \in p_{j,k,t}} W_{i,t}}. \tag{10}$$

As previously, when calculating equal-weighted portfolios, $W_{i,t} = 1 \forall\ i$ and t. Further, when calculating the differential portfolios for the n_{p1} portfolios, that is sorted on X_1, the difference between the highest and lowest values of the second sorting variable X_2 is taken. Thus, the equation from Bali et al. (2016) is

$$\bar{Y}_{j,diff,t} = \bar{Y}_{n_{j,p2,t}} - \bar{Y}_{j,1,t}, \tag{11}$$

this is not done for n_{p2} as it is unclear how to interpret the outcome. And lastly the average value of $\bar{Y}$ across all of the portfolios sorted on X_2 and inside the jth portfolio sorted on X_1

$$\bar{Y}_{j,Avg,t} = \frac{\sum_{k=1}^{n_{p1}} \bar{Y}_{j,k,t}}{n_{p2}}. \tag{12}$$

From Bali et al. (2016) the interpretation is that the statistically significant differences indicate a cross-sectional relation between X_2 and Y after controlling for X_1. Therefore, the analysis of this part focuses on the different portfolios sorted on the second variable X_2. The hypothesis is the same as previous, that the mean is statistically significant from zero, down to a significant level of 5%.

4.2 Sorting and weighting variables

The value-weighted returns in this paper use the same weighting as Grullon et al. (2020), which is the market capitalization for the firms. Following Bali et al. (2016) they calculate the market capitalization for each firm with the formula [2],

$$Market\ capitalization_t = Shares\ outstanding_t \ * \ price_t. \tag{13}$$

Further, the seasonality of the sales sorting variable that Grullon et al. (2020) uses is calculated by scaling the sales in quarter q of year t by the annual sales in year t. The sales seasonality variable SEA is a quarterly variable. Grullon et al. (2020) also removes samples with negative quarterly and annual sales, as well as if there are missing data for one quarter in a given year. Grullon et al. (2020) only includes firms with sales between 95% to 105% of total annual sales. Then the sales seasonality variable is as follows

$$SEA_{qt} = \frac{SALES_{qt}}{ANNUALSALES_t}. \tag{14}$$

To reduce the effect of extreme values, Grullon et al. (2020) takes the average values of SEA_{qt} from t-3 to t-2, calling the variable for $AVGSEA_{qt}$, which is mainly used as the sorting variable when trying to find the cross-sectional relation between sales seasonality and the excess return. From Grullon et al. (2020) they state that firms generating 25% sale quarterly through the year are firms with stable sales, and the firms above 25% sales for a quarter is in high season and low season if they are below.

As stated by Grullon et al. (2020), it is important to test if the sales volatility affects the sales seasonality premium. The reason behind this is because two stocks might happen to go from the lowest decile to the highest decile when sorting with AVGSEA. While they both are in the same decile portfolios, they might have quite different volatility. For instance, if firm A and B have 5% sales in low season, firm A might have 35% in

[2] In the data sample, the shares outstanding and price are calculated using absolute values, because some of the prices are negative, and negative market capitalization does not make any sense (Bali et al., 2016).

season, and firm B might have 40%. Therefore, they might create different seasonality premiums. To capture this effect, the portfolios are sorted into deciles based on the variable $SEARANGE_{qt}$ from the portfolio sorts file by Grullon et al. (2020). SEARANGE is the difference of the maximum and minimum values of $AVGSEA_q$ over the four last quarters (Grullon et al., 2020)

$$SEARANGE_{qt} = \max_{n=0\ to\ 3}(AVGSEA_{q-n}) - \min_{n=0\ to\ 3}(AVGSEA_{q-n}). \tag{15}$$

4.3 Capital asset pricing model and alpha

Introduced by Sharpe (1964), the capital asset pricing model is an equilibrium factor model that assumes that the market is efficient, and the returns are only affected by systematic risk. Where R_p is the return of the portfolio, the $R_f$is the risk-free rate, and the $R_m - Rf$ is the expected return on the market portfolio. The β is the risk on the asset compared to the market, and a negative beta means that the asset correlates negatively with the market. The capital asset pricing model predicts expected return, and the equation is

$$r_p - r_f = \beta^{mkt}(R_m - R_f) + \epsilon. \tag{16}$$

To find the excess return of the securities, I run a regression, defined from Berk and DeMarzo (2017) as

$$r_p - r_f = \alpha + \beta^{mkt}(R_m - R_f) + \epsilon, \tag{17}$$

where the left side of the regression is the excess return, α is the intercept of the regression, and $\beta^{mkt}(R_m - R_f)$ is the expected return. The α measures the performance of the portfolio against the market. Sharpe (1964) restricts the CAPM regression, where $\alpha = 0$. The reason for the restriction on α is because Sharpe (1964) assumes a perfect market. Thus, the return of the efficient portfolio cannot be higher than the return of the market portfolio. A positive alpha indicates that the portfolio outperforms the market portfolio. To test the seasonality strategy by Grullon et al. (2020), I run a regression to see if the alpha is statistically different from zero.

4.4 Fama-French factor models

Fama and French (1993) created the Fama-French factor model is an extension to the capital asset pricing model. Adding factors such as size and value risk for the three factor model. The three factor model is defined as

$$r_p - r_f = \alpha + \beta^{mkt}(R_m - R_f) + \beta^{size} SMB + \beta^{value} HML + \epsilon. \quad (18)$$

The first part of the equation is equal to the capital asset pricing model. From Fama and French (1993) the two new variables added is the difference portfolio between small minus big firms (SMB) which is the size factor, that controls for the small firm effect. The second is the high minus low difference portfolio (HML) which is the value factor, that controls for the value stock effect. Further, in a later paper Fama and French (2015) adds two more factors to the three-factor model, making the five-factor model,

$$r_p - r_f = \alpha + \beta^{mkt}(R_m - R_f) + \beta^{size} SMB + \beta^{value} HML + \beta^{profitability} RMW + \beta^{investment} CMA + \epsilon. \quad (19)$$

First, of the Fama and French (2015) factor is the robust minus weak difference portfolio (RMW) that controls operating profitability. Finally, the last Fama and French (2015) factor is the conservative minus aggressive difference portfolio (CMA), controlling for the difference between high and low investment firms.

4.5 Fama-Macbeth

Fama and MacBeth (1973) is a two-step cross-sectional regression method used to test how some factors describes the portfolio return. Basically a robustness test, where the main goal is to see how the sales seasonality premium is affected by these other factors. The first step in the Fama-Macbeth regression is to run a normal OLS to get the beta on the right side for each portfolio

$$r_{p,t} - r_{f,t} = \alpha_{p,t} + \beta_p F_{1,t} + \beta_p F_{2,t} + \beta_p F_{3,t}, ..., + \beta_p F_{n,t} + \epsilon_{n,t}. \tag{20}$$

The second step of the regression involves taking the beta and run a regression again on the portfolio monthly returns,

$$r_{p,t} - r_{f,t} = \gamma_{p,t} + \gamma_p \beta_{1,t} + \gamma_p \beta_{2,t} + \gamma_p \beta_{3,t}, ..., + \gamma_p \beta_{i,t} + \epsilon_{i,t}, \tag{21}$$

then take the time-series average across all periods to produce the results. A significant variable in this regression means that there is a linear relationship between the risk factor β and the return of the stock. Interpreting this is straightforward, it a measure of how a unit increase (decrease) in factor risk γ, changes the expected return of the stock (Fama and MacBeth, 1973). A drawback with Fama-Macbeth regression is that it assumes in most cases that the relationship between all variables and outcome is linear.

4.6 Sharpe ratio

Sharpe (1964) created the Sharpe ratio, which is a measure of return against risk. Using a similar method from Grullon et al. (2020), where they calculate the Sharpe ratio using the return from the Fama and French (1993) factor portfolios and the returns from a factor based on sales from Grullon et al. (2020). The Sharpe ratio equation is

$$Sharpe = \frac{\overline{X}}{\sigma(X)} * \sqrt{12}, \tag{22}$$

where the X represents the return of a factor portfolio and the square root of twelve annualized monthly variables.

4.7 Controlling for other seasonality effects

Heston and Sadka (2008) uses three seasonality factors to measure the seasonal effect.

Their way of doing so is to use the previous return. Thus, calculating these returns as

$$HS12 = r_{t-12}, \tag{23}$$

this is the return for the last year of the companies. The second variable is the average three-year same month previous average return,

$$HS36 = \frac{r_{t-12} + r_{t-24} + r_{t-36}}{3}. \tag{24}$$

The last variable is the average five-year previous same month average return,

$$HS60 = \frac{r_{t-12} + r_{t-24} + r_{t-36} + r_{t-48} + r_{t-60}}{5}. \tag{25}$$

I am removing the earnings announcement seasonality from Chang et al. (2016) by using the earnings announcement months for the firms listed in the COMPUSTAT file. This is because Chang et al. (2016) seem to use a different method in calculating the earnings announcement month. However, the result of just excluding the earnings announcement month listed in the COMPUSTAT is minor compared to the result by Grullon et al. (2020).

5 Data

In this section, I present the data files, explain how I wrangled the data using the programming language R, and how I calculate a few key variables using the data files.

5.1 Importing necessary data files

Forming the portfolios requires stock price information from the center for research in security prices - CRSP. This is the same data used in the paper by Grullon et al. (2020) which contains monthly stock prices, identity information, delisting stock information, and shares outstanding. The CRPS data set can be downloaded from Wharton research data service (WRDS) going back to 1970 and forward to December 2017. Then importing the CRSP file into the programming language R, we have a sample with over 4 million observations over 576 months. The following data set, which is COMPUSTAT quarterly and annually, can also be imported from WRDS, using the same period as the CRSP file. The quarterly COMPUSTAT sample has 1 million observations, while the annual COMPUSTAT file contains almost three hundred thousand observations. I import the portfolio sorts from Grullon (2019) data website. The file contains the sales seasonality variable (SEA), average sales seasonality (AVGSEA), the sales volatility (SEARANGE), and the absolute change between SEA in year t-3 and t-2 (SEAVAR); all the variables are quarterly. Additionally, the website also has a factor based on the sales seasonality strategy (SEAF) by Grullon et al. (2020). Lastly, I import the Fama-French factors from Kenneth (2021) website.

5.2 Data cleaning and manipulation

The CRSP file contains different shares traded in the data. The two-digit code sorts make it possible to separate the securities. The first part of the code represents the type of security traded, this paper follows, Grullon et al. (2020), and thus, I sort out every share that does not start with the number 1. Shares starting with number 1 represent ordinary common shares, ordinary shares sold on exchanges with rights to profits and voting. The second digit that is filtered, again following the original paper, is both 0 and 1. Number 0 represents securities that do not find special status, and securities with number 1 do not require special status. Finally, after identifying the correct securities, every share that does not contain the share code 10 and 11 are sorted out, leaving only U.S-based common stocks in the sample.

The next part is to remove financial firms from the sample, yet again following the procedure from the Grullon et al. (2020). The reason behind removing financial firms is their unusual business model with high leverage compared to other firms, giving them an uncommon book-to-market ratio. High leverage most likely has another meaning for the financial firms but usually indicates more distress amongst other firms as stated by Fama and French (1992). Therefore, from Bali et al. (2016), firms listed as financial, insurance, and real estate with an industry code between 6000 and 6999 are removed from the sample.

The CRSP file has stocks listed on NYSE, AMEX, NASDAQ, and a few other small stock exchanges. Therefore, the next step is to filter out the stocks that do not have exchange numbers 1 or 31 and 2 or 32 and 3 or 33, which are stocks listed on NYSE, AMEX, and NASDAQ stocks, respectively. Lastly, distinguishing the data and making sure that there are no duplicates observations on the securities. Furthermore, figure 1 shows the number of securities traded on each exchange, where NASDAQ has the most traded securities out of all three exchanges.

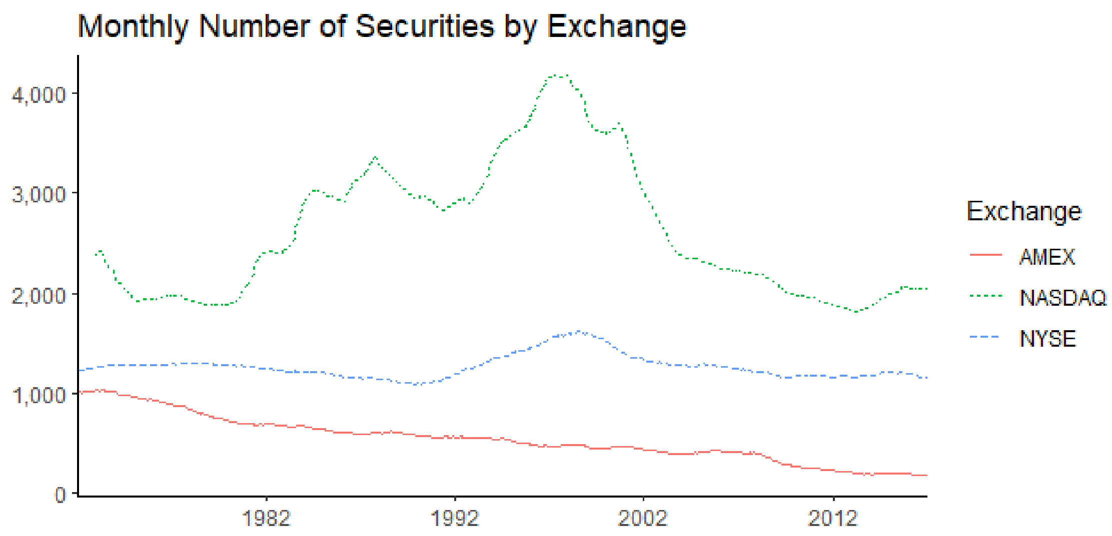

FIGURE 1: Following Bali et al. (2016) and make a graph of the monthly number of securities by Exchange.

Figure 2 shows that the total market value is the highest on the NYSE stock exchange, making it the biggest in terms of value. NASDAQ is second on, while AMEX barely is above zero. The market value has been adjusted for inflation using the 2017 dollar, as the research ends here.

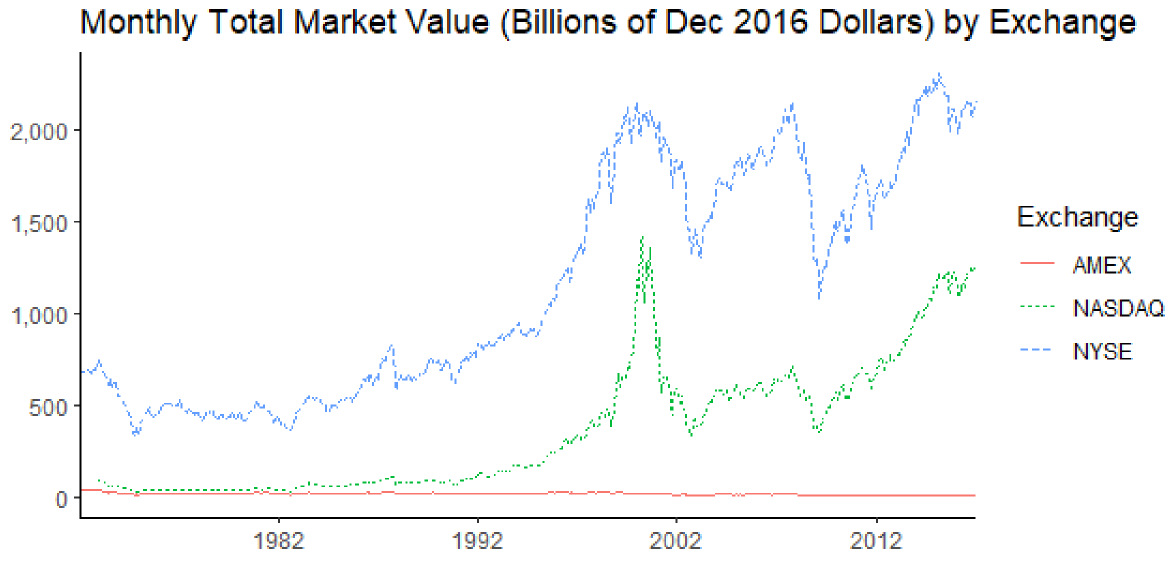

FIGURE 2: Following Bali et al. (2016) and make a graph of the monthly total market value by Exchange, adjusting for 2017 dollars.

The next step is to calculate the market capitalization in the sample to capture the size effect. According to Bali et al. (2016), there is an implementation issue when calculating the market capitalization. In the CRSP file, there are two price fields for monthly stocks, PRC and ALTPRC (alternative price). The PRC is the stock's closing price, and the ALTPRC is the last non-missing price overall days during the given month. ALTPRC is an actual closing price or the negative of the average of the bid and ask spread (Bali et al., 2016). Both shares outstanding (SHROUT) and alternative price have some zero values, where Bali et al. (2016) state that the interpretation of zero shares outstanding to be unclear. They also consider that if either shares outstanding or alternative price is zero, their values are missing. They also list a few ways of calculating the market cap, but following their simple step is to take the absolute value of the alternative price variable multiplied with shares outstanding and divide it by one thousand (Bali et al., 2016).

Now that I have cleaned the CRSP file and calculated key variables, the COMPUSTAT files are next. Following Fama and French (1993) first for the annual sample, keeping only the last observation for each year and firms included for two years or more. Then, filtering out financial firms again, with SIC code between 6000 to 6999. Next, the COMPUSTAT file has links that describe the connection between the CRSP and Compustat file. I will not go into detail for each of the present links, but I only choose valid links, therefore filtering out links that are not marked LU or LC and are still active at the information release time. Then, for the COMPUSTAT annual file, the variables are created by following the steps Fama and French (1993), where they are made in June at time t using information from t-1. Lastly, for the COMPUSTAT annual file is to calculate the book equity, using the steps of Bali et al. (2016), the book equity is the stockholder's equity added with deferred taxes plus investment tax credit and subtracting preferred stock redemption value if available. If this variable is not available, then the preferred stock liquidating value is subtracted. Should both not be available, then I use the preferred stock par value.

5.2.1 Proxies calculated from COMPUSTAT

When testing for the hypothesis explained previously in the theory section, this paper uses the same proxies as Grullon et al. (2020), created using the relevant variables from the COMPUSTAT file. The first proxy is the change in the total asset at quarter q relative to the total asset at q-1 (IAQ).

$$IAQ = \frac{(total\ assets)_q - (total\ asset)_{q-1}}{(total\ assets)_{q-1}} \tag{26}$$

Next is the change in current assets at quarter q relative to total assets at $q1$ (IACQ)

$$IACQ = \frac{(current\ assets)_q - (current\ assets)_{q-1}}{(total\ assets)_{q-1}} \tag{27}$$

The third proxy is the net property plant and equipment change at quarter q relative to total assets at $q1$ (PPEQ).

$$PPEQ = \frac{(net\ property\ plant)_q - (net\ property\ plant)_{q-1}}{(total\ assets)_{q-1}} \tag{28}$$

The last proxy for the investment variables is the change in inventories at quarter q relative to total assets at $q1$ (INVQ).

$$INVQ = \frac{(Inventory)_q - (Inventory)q - 1}{(total\ assets)_{q-1}} \tag{29}$$

Now that the investment variable proxies are defined, the financial variables are needed. The proxies for the financial variables in the lower panel are (i) BLchange, the change in BL (total debt divided by total asset) from $q1$ to q.

$$BLChange = \frac{(debt\ current\ liabilities)_q + (long\ term\ debt)_q}{total\ assets_q} - \frac{((debt\ current\ liabilities)_{q-1} + (long\ term\ debt)_{q-1}}{(total\ assets)_{q-1}} \tag{30}$$

Lastly, MLchange is the change in ML (total debt scaled by total debt plus the market value of equity) from q-1 to q. Also, removing firms in the lowest size quintile since the premium is small and statistically insignificant Grullon et al. (2020).

$$MLChange = \frac{(debt\ current\ liabilities)_q + (long\ term\ debt)_q}{(debt\ current\ liabilities)_q + (long\ term\ debt)_q + ((common\ shares\ outsantind)_q * (annual\ closing\ price))_q} \tag{31}$$

6 Empirical results

6.1 Sales seasonality

In this section, I will present the empirical results following the methodology and theory from the earlier sections.

6.1.1 Equal-weighted portfolio returns

Table 2 shows the cross-sectional regressions on the equal-weighted returns, the result from Grullon et al. (2020) is in parentheses for comparison. The CAPM reports a monthly excess return of 0.017% and an annual excess return of 0.20% for the portfo-

TABLE 2: Cross-sectional regression, equal-weighted returns. The equal-weighted portfolio's are created by sorting based on AVGSEA into deciles using NYSE, AMEX and NASDAQ breakpoints.The sample uses non-financial stocks from CRSP starting January 1972 to December 2017.

	CAPM										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.017	0.193	0.324**	0.384***	0.367***	0.389***	0.392***	0.375**	0.215	−0.213	0.243**
	(0.185)	(0.147)	(0.142)	(0.135)	(0.129)	(0.132)	(0.147)	(0.159)	(0.161)	(0.178)	(0.097)
Mrktf	1.184***	1.141***	1.134***	1.095***	1.097***	1.088***	1.113***	1.133***	1.158***	1.171***	0.010
	(0.047)	(0.039)	(0.041)	(0.038)	(0.041)	(0.036)	(0.039)	(0.042)	(0.043)	(0.049)	(0.027)

	Fama-French three factor risk-adjusted returns										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	−0.041	0.076	0.185**	0.265***	0.220***	0.256***	0.251***	0.235**	0.091	−0.288**	0.254**
	(0.125)	(0.086)	(0.080)	(0.074)	(0.067)	(0.072)	(0.076)	(0.094)	(0.093)	(0.118)	(0.103)
Mrktf	1.006***	1.018***	1.028***	0.984***	1.004***	0.987***	1.014***	1.017***	1.028***	1.011***	−0.006
	(0.033)	(0.028)	(0.024)	(0.021)	(0.025)	(0.021)	(0.024)	(0.029)	(0.029)	(0.038)	(0.028)
SMB	1.060***	0.889***	0.849***	0.842***	0.790***	0.805***	0.808***	0.917***	0.944***	1.002***	0.060*
	(0.051)	(0.065)	(0.049)	(0.041)	(0.052)	(0.048)	(0.047)	(0.055)	(0.062)	(0.050)	(0.034)
HML	−0.050	0.090	0.146***	0.115***	0.186***	0.131**	0.168***	0.150**	0.098	−0.017	−0.033
	(0.094)	(0.064)	(0.052)	(0.044)	(0.048)	(0.051)	(0.050)	(0.063)	(0.083)	(0.081)	(0.046)

	Fama-French five factor risk-adjusted returns										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.161	0.161	0.238**	0.285***	0.219***	0.266***	0.253***	0.279***	0.172	−0.113	0.275***
	(0.122)	(0.109)	(0.099)	(0.085)	(0.079)	(0.089)	(0.079)	(0.108)	(0.115)	(0.130)	(0.097)
Mrktf	0.961***	0.997***	1.016***	0.982***	1.008***	0.991***	1.019***	1.010***	1.009***	0.972***	−0.011
	(0.033)	(0.034)	(0.027)	(0.025)	(0.029)	(0.028)	(0.029)	(0.036)	(0.039)	(0.043)	(0.029)
SMB	0.926***	0.838***	0.812***	0.818***	0.774***	0.779***	0.787***	0.874***	0.889***	0.880***	0.046
	(0.066)	(0.062)	(0.049)	(0.041)	(0.044)	(0.045)	(0.045)	(0.058)	(0.063)	(0.064)	(0.044)
HML	−0.001	0.127**	0.159***	0.102**	0.152***	0.090*	0.127**	0.137*	0.120	0.025	−0.026
	(0.092)	(0.064)	(0.055)	(0.046)	(0.050)	(0.054)	(0.060)	(0.081)	(0.085)	(0.094)	(0.087)
RMW	−0.531***	−0.206**	−0.144**	−0.086	−0.049	−0.089	−0.066	−0.158*	−0.218**	−0.477***	−0.054
	(0.089)	(0.097)	(0.072)	(0.052)	(0.078)	(0.070)	(0.074)	(0.085)	(0.108)	(0.100)	(0.067)
CMA	−0.039	−0.055	−0.010	0.044	0.083	0.106	0.102	0.052	−0.019	−0.030	−0.010
	(0.142)	(0.116)	(0.104)	(0.084)	(0.088)	(0.099)	(0.108)	(0.143)	(0.153)	(0.177)	(0.135)

Note: *p<0.1; **p<0.05; ***p<0.01

lio in the lowest decile of sales seasonality. The highest sales seasonality decile portfolio shows a monthly excess return of -0.21% and an annual excess return of -2.55%, which is economically significant. The portfolios report a t-statistic of 0.09[3] and -1.19 respectively, thus, they are not statistically significant. Therefore, I can not reject the null hypothesis that the excess returns are different on both portfolios. The spread between the two extreme portfolios is the low-minus-high portfolio, which gives an excess monthly return of 0.24% and an annual excess return of 2.91% (3.12%), statistically significant with a t-value of 2.49 (2.88)[4].

In the following asset pricing model, the Fama-French three-factor reports a monthly excess return of -0.04% and an annual excess return of 0.49% on the lowest decile portfolio, with a t-statistic of -0.38, the excess return is not statistically significant. The highest sales seasonality decile portfolio shows a monthly excess return of -0.29% and an annual excess return of 3.45%. With a t-value of -2.43%, the finding is significant. The low-minus-high portfolio has a monthly excess return of 0.254% and an annual excess return of 3.05% (3.24%), statistically significant with a t-value of 2.45 (2.88)

The last asset pricing model to control for is risk-adjusting with the Fama-French five-factor model. The lowest decile portfolio shows a monthly excess return of 0.16%, with an excess return of 1.93% annually. A t-statistic of 1.3196 shows that excess return is not statistically different from zero. The same goes for the extreme portfolio in the highest decile, with a monthly excess return of -0.11% and an annual excess return of -1.35 %. The is return is statistically insignificant with a t-value of 0.86. Finally, the long-minus-high portfolio has a monthly excess return of 0.27% and 3.30% (3.48%) annually with a t-statistic of 2.83 (3.09), showing that it is statistically significant.

After testing the three asset pricing models, the result between the spreads is simi-

[3]The standard errors are adjusted using Newey and West (1987), for more detailed explanation see the appendix.

[4]A test with the Fama-French three-factor + momentum model, equal-weighted return shows a statistically insignificant annual alpha of 1.76%, the result is reported in table 13 in the appendix.

lar with only a 0.32% difference in the excess return between the CAPM and Fama-French five-factor risk-adjusted returns. There are no distinct patterns on the return of the portfolios between the extreme portfolios as well. The results on every model are similar to the results Grullon et al. (2020), showing that the replication on equal-weighted returns over the same number of periods is successful.

6.1.2 Value-weighted portfolio returns

Table 3 is the cross-sectional regression on the value-weighted returns. The CAPM reports a monthly excess return of 0.20 and 2.50% annually for the lowest decile portfolio. With a t-value of 1.51, the result is statistically insignificant. For the extreme portfolio in the highest decile, the monthly and annual excess return is 0.42% and 5.01%, with a t-value of 3.76; the finding is statistically significant. Finally, the long-minus-high portfolio has an excess return of 0.63% monthly and 7.51% (7.8%) annually. This return is larger than the findings in the equal-weighted returns, with a t-statistic of 3.16 (4.32); the finding is statistically significant[5].

The Fama-French three-factor model shows an excess return of 0.39% monthly and 4.67% annually for the lowest portfolio, statistically significant with a t-value of 3.16. For the portfolio in the highest decile, the excess return is -0.34% monthly and 4.10% annually, significant statistically with a t-value of 2.71. Finally, the excess return on the low-minus-high portfolio is 0.73% monthly and 8.79% (8.88%) annually, with a t-value of 4.49 (4.98); the finding is statistically significant.

Lastly, testing for the Fama-French five-factor model on the lowest decile portfolio shows an excess return of 0.54% monthly and 6.43% annually with a t-value of 4.10; it is statistically significant. The high extreme decile portfolio has a monthly excess return of -0.155% and -1.78% annually, but with a t-value of 1.25, this finding is not statistically significant. Finally, the long-minus-high portfolio has an excess return of

[5]A test with the Fama-French three-factor + momentum model, value-weighted return shows a statistically significant annual alpha of 4.764%, the result is reported in table 14 in the appendix.

TABLE 3: Cross-sectional regression, value-weighted returns. The value-weighted portfolio's are created by sorting based on AVGSEA into deciles using NYSE, AMEX and NASDAQ breakpoints. The sample uses non-financial stocks from CRSP starting January 1972 to December 2017.

	CAPM										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.209	0.101	0.256***	0.077	0.135**	0.155***	0.018	−0.106	−0.179**	−0.418***	0.626***
	(0.138)	(0.096)	(0.080)	(0.066)	(0.063)	(0.058)	(0.085)	(0.083)	(0.087)	(0.111)	(0.173)
Mrktf	1.168***	1.024***	0.994***	0.937***	0.903***	0.939***	0.963***	1.041***	1.039***	1.070***	0.096*
	(0.047)	(0.025)	(0.024)	(0.021)	(0.018)	(0.024)	(0.024)	(0.017)	(0.025)	(0.044)	(0.050)

	Fama-French three factor risk-adjusted returns										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.389***	0.175**	0.283***	0.098	0.149**	0.161***	0.039	−0.084	−0.144	−0.342***	0.732***
	(0.123)	(0.089)	(0.082)	(0.071)	(0.059)	(0.056)	(0.086)	(0.078)	(0.095)	(0.126)	(0.163)
Mrktf	1.035***	0.979***	0.983***	0.943***	0.923***	0.954***	0.963***	1.026***	1.018***	1.005***	0.029
	(0.030)	(0.024)	(0.028)	(0.022)	(0.018)	(0.024)	(0.023)	(0.016)	(0.030)	(0.041)	(0.046)
SMB	0.329***	0.086***	0.004	−0.075***	−0.135***	−0.088**	−0.044	0.031	0.038	0.192***	0.137**
	(0.050)	(0.029)	(0.040)	(0.027)	(0.037)	(0.040)	(0.035)	(0.032)	(0.064)	(0.049)	(0.054)
HML	−0.446***	−0.174***	−0.059*	−0.034	−0.008	0.005	−0.040	−0.051	−0.080	−0.195***	−0.251***
	(0.077)	(0.034)	(0.034)	(0.041)	(0.043)	(0.033)	(0.040)	(0.036)	(0.049)	(0.070)	(0.080)

	Fama-French five factor risk-adjusted returns										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.536***	0.241***	0.258***	0.009	0.057	0.078	−0.066	−0.062	−0.098	−0.148	0.681***
	(0.137)	(0.087)	(0.083)	(0.059)	(0.062)	(0.056)	(0.081)	(0.079)	(0.111)	(0.118)	(0.166)
Mrktf	0.991***	0.957***	0.988***	0.972***	0.956***	0.989***	0.995***	1.015***	0.997***	0.944***	0.047
	(0.033)	(0.024)	(0.026)	(0.018)	(0.020)	(0.022)	(0.024)	(0.018)	(0.027)	(0.032)	(0.047)
SMB	0.265***	0.066**	0.024	−0.043	−0.117***	−0.091**	−0.00001	0.036	0.046	0.115***	0.150**
	(0.038)	(0.032)	(0.032)	(0.027)	(0.031)	(0.036)	(0.042)	(0.032)	(0.051)	(0.042)	(0.058)
HML	−0.333***	−0.103**	−0.061	−0.123**	−0.130***	−0.144***	−0.130***	−0.001	0.015	−0.023	−0.311***
	(0.051)	(0.043)	(0.052)	(0.049)	(0.041)	(0.041)	(0.041)	(0.036)	(0.068)	(0.062)	(0.088)
RMW	−0.279***	−0.097**	0.074	0.148***	0.105*	0.035	0.196***	0.005	−0.001	−0.347***	0.068
	(0.091)	(0.046)	(0.047)	(0.038)	(0.056)	(0.053)	(0.063)	(0.042)	(0.086)	(0.066)	(0.115)
CMA	−0.223	−0.151**	−0.006	0.185***	0.277***	0.340***	0.196***	−0.118*	−0.222**	−0.351***	0.129
	(0.146)	(0.077)	(0.080)	(0.068)	(0.081)	(0.078)	(0.071)	(0.064)	(0.113)	(0.118)	(0.169)

Note: *p<0.1; **p<0.05; ***p<0.01

0.68% monthly and 8.17% (8.4%), statistically significant with a t-value of 4.10 (4.56).

Again, after testing for all the different asset pricing models, the returns are between the spreads are similar. The findings are similar to Grullon et al. (2020) as well, and thus, concluding that replication of the value-weighted return for the same number of periods to be successful. In contrast to the equal-weighted returns, the value-weighted returns are higher. Therefore, the sales seasonality effects seem to be stronger among larger firms, indicating that this effect is not a small firm phenomenon. The results are on par with the findings of Grullon et al. (2020).

6.1.3 Size and sales seasonality

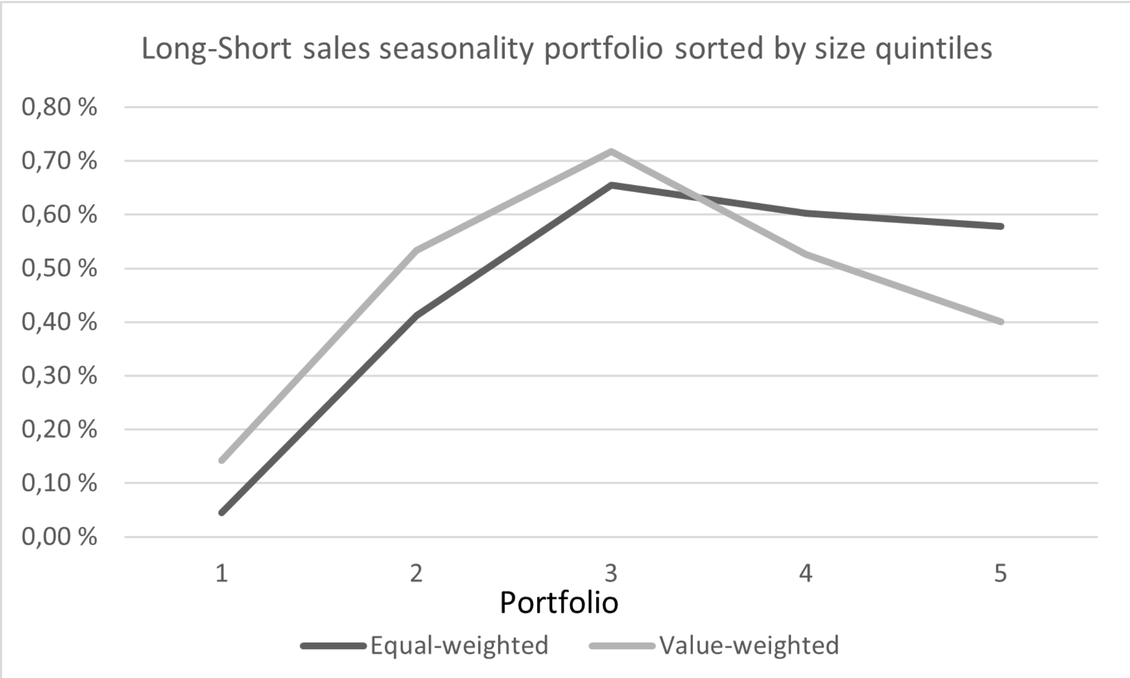

FIGURE 3: Long-Short portfolios sorted on size and sales. First, creating five portfolios sorted on market capitalization, using NYSE breakpoints. Inside each quintile, creating five new portfolios by sorting on sales using NYSE, AMEX, and NASDAQ breakpoints. The result is 25 portfolios for both equal and value-weighted returns. For each size portfolio, I make the difference between the low and the high sales portfolios. The excess return from the size portfolios is the low-minus-high portfolio reported on the X-axis for each size portfolio. The Y-axis represents excess return.

After the previous finding in the value-weighted returns and its indication, the next step will test for the relationship between size and seasonality effect. Furthermore, I am sorting the five portfolios constructed to test for this relationship on market capitalization. Five more portfolios are constructed inside each of these size portfolios, sorted independently into quintiles on the same previous sales seasonality variable, AVGSEA. The results are 5x5 portfolios sorted on size and AVGSEA to test for firm size effects, risk-adjusted with the Fama-French five-factor model. Figure 3 shows the excess return from the low-minus-high portfolio for each portfolio sorted on size, including equal and value-weighted portfolios. For the equal-weighted, the first size portfolio reports an annual excess return of 0.54% for the low-minus-high portfolio, the third portfolio of 7.86%, and the fifth portfolio has an excess return 6.93%. For the value-weighted returns, the first portfolio shows an annual excess return of 1.70%, the third with 8.61%, the last portfolio 4.80%. Grullon et al. (2020) reports a similar excess return for all the portfolios in all the quintiles. Table 4 below shows the return of the 5x5 portfolios, and as can be seen from the table, the low and high season sales portfolio generate insignificant alphas for the lowest market capitalization portfolio. The effect is stronger among larger firms. In the low sales seasonality portfolio, the firms within the highest three-size portfolios generate the highest alphas. While in the highest sales seasonality, the alpha in the second, third and fourth size portfolio is higher than others. Thus, a similar conclusion Grullon et al. (2020) can be drawn, that there is no evidence that this is a small firm phenomenon.

TABLE 4: Using the method by Bali et al. (2016) the portfolios are sorted into quntilies based on market cap, using NYSE breakpoints. Inside each of the portfolios, five new portfolios are created using NYSE, AMEX and NASDAQ breakpoints, sorted on AVGSEA. This creates 25 portfolios double-sorted on size and sales, risk-adjusted against Fama-French five-factor. The result displays the excess return for each portfolio. The sample used are non-financial firms from CRSP, dating from January 1972 to December 2017.

	Low SEA	2	3	4	High SEA
Low ME	0.0857 [0.99]	0.2387 [3.05]	0.2097 [2.72]	0.2260 [2.875]	−0.0729 [-0.764]
2	0.1431 [1.803]	0.2112 [2.827]	0.1757 [2.407]	−0.0013 [-0.019]	−0.3736 [-4.524]
3	0.3667 [3.786]	0.2102 [2.709]	−0.0579 [-0.799]	−0.0033 [-0.04]	−0.3231 [-3.708]
4	0.3822 [4.07]	0.1032 [0.134]	0.0098 [0.134]	−0.1045 [-1.439]	−0.123 [-1.327]
High ME	0.3997 [41.182]	0.1369 [2.153]	0.0828 [1.411]	−0.0472 [-0.613]	0.00825 [0.079]

6.1.4 Volatility on sales and its role in the seasonality premium

The findings when sorting for SEARANGE are in table 5, where the alpha on the first portfolio is -0.0002, t-value = 0.003, and an alpha of 0.09, with a t-value = 0.17 for the last portfolio, and thus, they are bot insignificant. The spread between those portfolios is 0.09, with a t-value of 0.13. None of these findings are signs indicating that the volatility does not affect the sales seasonality premium. To further test this, splitting the sample into two based on the median SEARANGE level, the firms below-median are the low volatility firms, and vice versa for those above the median. Panel B reports the firms below-median SEARANGE sorted into deciles based on AVGSEA. None of the findings in panel B are significant, showing evidence that firms with low variation in sales seasonality do not generate an excess return compared to those who have high and low seasons. In panel C, the lowest decile portfolio is statistically significant with an alpha of 0.46 and a t-statistic of 3.53, while the highest is not significant with an alpha of -0.30 and a t-value of 1.62. The low-minus-high portfolio has an alpha of 0.76 and statistically significant with a t-value of 3.55. This result shows that firms with high variation between sales seasonality create an excess return over those who do not. The findings are mostly on par with Grullon et al. (2020), except the highest decile portfolio in panel C, which they find to be significant down to 5% with a t-value of -2.08. Nevertheless, I find that the spread portfolio is highly significant in this thesis, and therefore the same conclusion can be drawn as Grullon et al. (2020), that the alphas are from firms with high and low seasons.

TABLE 5: Volatility of sales. Sorted into portfolios using the SEARANGE variable from Grullon et al. (2020), caputuring the volatility. Using NYSE, AMEX and NASDAQ breakpoints. Panel A uses full sample, sorting on SEARANGE. Panel B regression takes the sample below median SEARANGE and sort into decile portfolios based on AVGSEA. Panel C regression uses the sample above median SEARANGE, sorting into decile portfolios using AVGSEA.

Panel A: FF5 model, sorted on SEARANGE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(H-L)
Alpha	−0.0002	−0.013	−0.002	−0.009	0.035	0.190*	0.104	0.022	0.338**	0.093	0.095
	(0.059)	(0.077)	(0.060)	(0.063)	(0.075)	(0.100)	(0.086)	(0.089)	(0.143)	(0.104)	(0.131)
Mrktf	0.943***	1.007***	1.010***	0.985***	0.974***	0.984***	0.971***	0.993***	0.968***	0.946***	−0.003
	(0.021)	(0.019)	(0.018)	(0.017)	(0.020)	(0.023)	(0.025)	(0.028)	(0.029)	(0.025)	(0.032)
SMB	−0.084***	−0.063**	−0.014	−0.009	−0.069**	0.026	0.019	0.156***	0.118**	0.305***	−0.390***
	(0.027)	(0.030)	(0.028)	(0.034)	(0.031)	(0.041)	(0.038)	(0.050)	(0.048)	(0.036)	(0.045)
HML	−0.133***	−0.074*	−0.029	−0.068	−0.071	−0.196***	−0.012	−0.041	−0.241***	−0.293***	0.159***
	(0.040)	(0.038)	(0.040)	(0.042)	(0.048)	(0.059)	(0.059)	(0.073)	(0.057)	(0.043)	(0.060)
RMW	0.227***	0.125***	0.164***	0.168***	0.066	−0.160**	−0.008	−0.013	−0.398***	−0.417***	0.645***
	(0.035)	(0.044)	(0.043)	(0.048)	(0.061)	(0.076)	(0.065)	(0.048)	(0.092)	(0.066)	(0.079)
CMA	0.453***	0.333***	0.253***	0.139**	−0.071	−0.060	−0.242***	−0.240**	−0.323***	−0.189**	0.643***
	(0.077)	(0.049)	(0.063)	(0.066)	(0.084)	(0.077)	(0.086)	(0.094)	(0.096)	(0.079)	(0.115)

Panel B: FF5 model, below-median SEARANGE, sorted into deciles based on AVGSEA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	−0.031	0.108	0.029	0.106	0.108	0.129*	−0.056	−0.233**	−0.023	−0.087	0.056
	(0.100)	(0.107)	(0.085)	(0.085)	(0.068)	(0.074)	(0.079)	(0.105)	(0.110)	(0.095)	(0.155)
Mrktf	1.037***	1.020***	0.952***	0.978***	0.954***	0.966***	1.004***	0.993***	0.992***	0.936***	0.101**
	(0.034)	(0.027)	(0.034)	(0.022)	(0.022)	(0.025)	(0.031)	(0.029)	(0.028)	(0.023)	(0.040)
SMB	0.177***	0.005	0.004	−0.040	−0.115***	−0.071*	−0.039	−0.003	0.017	0.003	0.174***
	(0.047)	(0.044)	(0.044)	(0.032)	(0.035)	(0.043)	(0.040)	(0.037)	(0.048)	(0.036)	(0.061)
HML	−0.088	−0.006	−0.012	−0.079	−0.135***	−0.229***	−0.018	−0.011	0.002	−0.063	−0.024
	(0.075)	(0.053)	(0.064)	(0.050)	(0.050)	(0.044)	(0.050)	(0.051)	(0.068)	(0.042)	(0.090)
RMW	0.153*	0.108	0.210***	0.236***	0.117*	0.162**	0.224***	0.168**	0.221**	−0.025	0.178*
	(0.090)	(0.066)	(0.072)	(0.059)	(0.060)	(0.076)	(0.042)	(0.071)	(0.103)	(0.063)	(0.100)
CMA	0.222*	0.163	0.181	0.347***	0.365***	0.512***	0.302***	0.147	0.116	0.043	0.179
	(0.130)	(0.114)	(0.124)	(0.078)	(0.107)	(0.065)	(0.079)	(0.106)	(0.123)	(0.122)	(0.205)

Panel C: FF5 model, above-median SEARANGE, sorted into deciles based on AVGSEA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.467***	0.546***	0.438***	0.216*	−0.082	−0.040	0.064	−0.152	−0.082	−0.298	0.766***
	(0.132)	(0.177)	(0.115)	(0.112)	(0.119)	(0.117)	(0.127)	(0.116)	(0.153)	(0.183)	(0.216)
Mrktf	0.991***	0.980***	0.941***	0.981***	0.998***	1.041***	1.015***	1.074***	1.005***	0.933***	0.058
	(0.033)	(0.029)	(0.035)	(0.032)	(0.038)	(0.031)	(0.033)	(0.030)	(0.051)	(0.042)	(0.053)
SMB	0.268***	0.248***	0.075	−0.044	0.082	0.080*	0.253***	0.050	0.018	0.243***	0.025
	(0.048)	(0.046)	(0.055)	(0.048)	(0.060)	(0.048)	(0.065)	(0.071)	(0.084)	(0.065)	(0.075)
HML	−0.226***	−0.292***	−0.181**	−0.128*	−0.023	−0.057	−0.101	−0.052	0.156	−0.177**	−0.049
	(0.052)	(0.061)	(0.074)	(0.066)	(0.086)	(0.085)	(0.077)	(0.100)	(0.112)	(0.081)	(0.101)
RMW	−0.400***	−0.155	−0.139**	−0.055	−0.006	0.126*	−0.195*	−0.102	−0.219*	−0.398***	−0.002
	(0.078)	(0.114)	(0.063)	(0.080)	(0.103)	(0.075)	(0.105)	(0.107)	(0.126)	(0.098)	(0.108)
CMA	−0.163	−0.251*	−0.098	−0.141	−0.004	−0.116	−0.242**	−0.177	−0.443***	−0.294**	0.132
	(0.112)	(0.129)	(0.102)	(0.124)	(0.139)	(0.109)	(0.119)	(0.123)	(0.165)	(0.132)	(0.169)

Note: *p<0.1; **p<0.05; ***p<0.01

6.1.5 Sorting into different decades

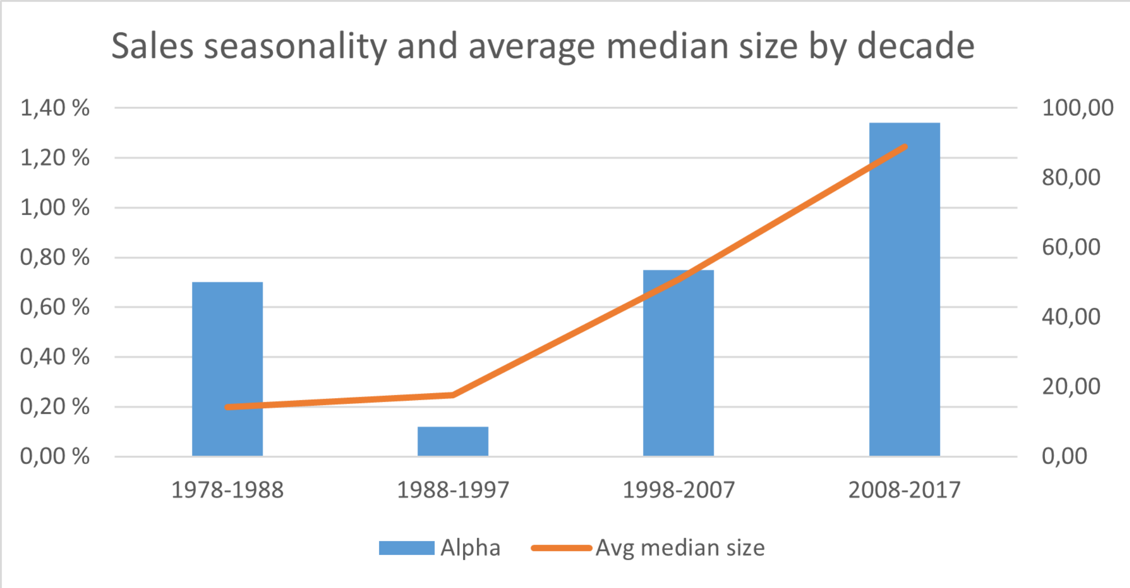

FIGURE 4: The results of the alphas through different periods and the average median firm size. The median firm size is calculated using the market equity from June in year t and is used from July to June of year t+1 using constant dollars from 1972. Then taking the median firm size each year and average it over each decade, and risk-adjusting with Fama-French five-factor (Grullon et al., 2020).

Furthermore, following Grullon et al. (2020), I calculate the return for each decade and the change in market capitalization. Shown in figure 4, the first decade from 1978-1988 shows an annual alpha of 8.40%, while the second decade shows a lower alpha of 1.44%. The annual alpha in the third decade is higher than the first, with an alpha of 9% annually. The last decade has the highest annual alpha of 16.08%. The alphas are similar to those reported by Grullon et al. (2020). Both numbers are the same, so I can draw a similar conclusion, which is that the sales seasonality premium has persisted over the decades, and it has increased as the median size of firms has increased (Grullon et al., 2020).

6.2 Controlling for other factors

6.2.1 Correlation Sharpe

TABLE 6: Pearson correlations shows the correlation between the factors. The Sharpe ratios calculated are annualized using the method from chapter 4. The sample used are non-financial firms from CRSP, dating from January 1972 to December 2017.

	Mrktf	SMB	HML	CMA	RMW	SEAF
Mrktf	1	0.235	−0.277	−0.384	−0.247	0.001
SMB		1	−0.059	−0.035	−0.377	0.025
HML			1	0.694	0.131	−0.125
CMA				1	0.043	−0.072
RMW					1	−0.027
SEAF						1

Variables	Mean%	St. dev.	Annualized Sharpe ratio
Mrktf	0.571	4.481	0.442
SMB	0.171	2.998	0.198
HML	0.361	2.907	0.430
CMA	0.310	1.960	0.548
RMW	0.284	2.286	0.430
SEAF	0.30	0.015	0.743

After analyzing the premium and its volatility, the next part is to test for correlation between the dependent and independent variables. Following Grullon et al. (2020), by using a factor for sales seasonality, I test for correlation between the variables. The sales seasonality factor is provided by Grullon et al. (2020) their website and is created by sorting into 2x3 portfolios, using median market equity and 30th and 70th percentile sales seasonality. The returns of the size portfolios and subtracting the high sales season portfolios from the low sales season portfolios for both small and big firms, the returns are value-weighted. Next, using Pearson correlation test on the five Fama-French factors, excess market return, SMB, HML, CMA, and RMW. The results reported in table 6 indicate that the SEAF variable is not correlated to any other factors, as its highest correlated with HML, with a value of -0.12, and lowest correlated to market excess return with a value of 0.001. It negatively correlated with three of the factors, HML, CMA, and RMW, and positively with excess market return and SMB.

Then following Grullon et al. (2020), to capture the performance of the factors, I calculate the mean, standard deviation, and Sharpe ratio. From table 6 the market excess return has the highest mean, with a monthly mean of 0.57% and an annual 6.85%. The SMB has the lowest, with a monthly mean of 0.17% and 2.05% annual. The HML reports a monthly mean of 0.36% and 4.33% annual. CMA has a monthly and annual mean of 0.31% and 3.72%. RMW with 0.28% and 3.40% monthly and annually. Lastly, SEAF shows a monthly mean of 0.30% and 3.60% annually. The market excess return has the highest standard deviation of all the factors with 4.48. Both SMB and HML report around 2.9 in standard deviation, CMA has 1.96, and RMW 2.29. The sales factor SEAF has the lowest standard deviation with 0.01. The market excess return has a Sharpe of 0.44, HML and RMW both have equal Sharpe with 0.43, SMB has the lowest with a Sharpe of 0.20. CMA Sharpe is 0.55, and the sales seasonality factor SEAF has the highest Sharpe ratio of 0.74.

In summary, there is no correlation between the sales seasonality factor and the other factors, and thus, the sales seasonality effect does not correspond to the existing Fama-French factors. Next, the mean, standard deviation, and Sharpe ratio show that the factor based on sales outperforms all the other Fama-French factors, with the highest compensation for risk. The result in table 6 is similar to the findings reported by Grullon et al. (2020).

6.2.2 Fama-MacBeth

The first step in Fama-Macbeth is to run a cross-sectional regression, where the excess monthly return is the dependent variable (Bali et al., 2016, p.89). The independent factors are (1) sales seasonality (SEA), (2) log of the market value of equity (LogME), (3) book-to-market ratio (BM), (4) investment-to-assets (IA), (5) gross profits-to-assets (GPA), (6) momentum from t-12 to t-2 (MOM), and (7) the three-month average of the previous same month returns (HS36) (Grullon et al., 2020). The second step is to calculate time-series averages of periodic cross-sectional regression coefficients[6].

[6] In this case, the Newey and West (1987) procedure is used again with six lags.

TABLE 7: Fama-MacBeth regressions on monthly excess return against sales seasonality factor, log of market capitalization, book-to-market, investment-to-asset, gross-profits-to-assets, momentum factor and lastly HS36 from Heston and Sadka (2008). The sample is split into three, where the first part is the full sample, the second part are the firms below NYSE median, and the third are the firms above NYSE median. The sample used are non-financial firms from CRSP, dating from January 1972 to December 2017. The independent variables has been standardized.

Variables	Full sample model 1	Full sample model 2	Small firms model 1	Small firms model 2	Big firms model 1	Big firms model 2
Intercept	0.98	0.98	0.99	0.99	0.81	0.82
	[2.93]	[2.94]	[2.89]	[2.89]	[2.90]	[3.17]
SEA	-0.07	-0.08	-0.07	-0.07	-0.21	-0.20
	[-2.23]	[-2.73]	[-2.02]	[-2.32]	[-5.79]	[-6.73]
LogME		-0.18		-0.188		-0.005
		[-1.65]		[-1.63]		[-0.06]
BM		0.30		0.32		0.22
		[4.18]		[4.48]		[3.80]
IA		-0.22		-0.232		-0.1
		[-4.51]		[-4.44]		[-2.54]
GPA		0.31		0.29		0.25
		[4.86]		[4.43]		[4.41]
MOM		0.01		0.004		-0.02
		[0.27]		[0.11]		[-0.26]
HS36		0.21		0.21		0.235
		[4.97]		[4.86]		[4.57]

The regression is done in six different ways, whereas the first and second tables are the total firm sample and the other four are against small and big firms. Starting for the total sample, model 1 shows that the seasonality of the sales affects the return negatively, significant with a t-value of -2.23, confirming the previous findings that when the return increases, the sales seasonality decreases. Model 2 shows testing against all the other controls as well, where sales seasonality is negative. For the other controls, the log of the market value of equity and investments-to-assets negatively affects excess monthly return. In comparison, the book-to-market ratio, gross profit-to-asset, and HS36 affect the return positively. Momentum is not a significant variable.

For the small firm samples, the effect is still negative, with significant t-values. For the other factors, the result is almost like the previous two columns, except for a lower value for the log of the market value of equity and a higher book-to-market effect. The last two columns report for the companies above median market capitalization, where the factors have the same sign as previous some of the factors, but compared to the full and small firms sample the seasonality of the sales for the big firm has a higher impact on returns. Both of the sales seasonality variables are statistically significant for the big firms.

Lastly, the significant increases for all the samples when using additional controls, showing that the effect is more robust when including more of the firm characteristics similar to Grullon et al. (2020). Thus, this section of robust testing again confirms previous findings of the sales seasonality premium and again proves that the bigger firms have a more substantial effect on the seasonality of the sales compared to small firms, as shown by Grullon et al. (2020).

TABLE 8: The portfolios are created using NYSE, AMEX and NASDAQ breakpoints and AVGSEA to sort into decile portfolios. The value-weighted excess return of excluding the earnings announcement from the sample is reported in the table, risk-adjusted with Fama-French five factor. The sample consist of non-financial firms from the CRSP file going back to January 1972 to December 2017.

	FF5 factor, excluding earnings announcement from the sample.										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.456***	0.288***	0.188*	0.099	0.052	0.128	−0.035	0.042	−0.031	−0.150	0.602***
	(0.135)	(0.100)	(0.098)	(0.074)	(0.072)	(0.110)	(0.103)	(0.110)	(0.117)	(0.144)	(0.199)

Note: *p<0.1; **p<0.05; ***p<0.01

6.3 Excluding other seasonality effects

From the seasonal return study by Chang et al. (2016), they find that the stock returns are higher in a quarter when announcing earnings. Their sales strategy is to go long on stocks in high season and short stocks in low season. The difference is that their method focuses on the earnings announcement months, and this method focuses on sales seasonality Chang et al. (2016). Testing if this is related to their seasonality effect,Grullon et al. (2020) removes the earnings announcement months from the sample. Risk-adjusting using Fama-French five-factor, the result is reported in table 8, using value-weighted monthly returns. The lowest decile portfolio shows a monthly alpha of 0.45% and 5.47% annually, whereas the highest decile has a monthly alpha of -0.15% and 1.8% annually. The low-minus-high portfolio monthly excess return is 0.60% and 7.22% annually, which is lower than the findings in table 3 on the Fama-French five-factor model with an annual return of 8.16%. Not all the portfolios show a lower return after excluding earnings announcement months. Compared to Grullon et al. (2020). All their portfolios show a lower return, with an annual alpha on the low-minus-high portfolio of 8.4% after excluding the earnings announcement, which is equal to their original sample (Grullon et al., 2020). Nevertheless, removing the earnings announcement months from the sample still gives a significant value, both economically and statistically. Thus, I can conclude that their seasonal effect is unrelated to the findings of this study.

TABLE 9: The portfolios are created using NYSE, AMEX and NASDAQ breakpoints, sorted into decile portfolios using AVGSEA. The value-weighted excess return of excluding the return seasonality from the sample is reported in the table, risk-adjusted with Fama-French five-factor model. HS12 is the last year same month return, HS36 is the average of the last three same month return over three years. HS60 is the average of the last five same month returns over five years. The sample used are non-financial firms from CRSP, dating from January 1972 to December 2017.

	FF5 factor returns, excluding Heston and Sadka (2018)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha, No HS-12	0.559***	−0.195	−0.214**	−0.087	−0.090	−0.047	−0.017	0.230***	0.369***	−0.133	0.691***
	(0.109)	(0.133)	(0.090)	(0.097)	(0.072)	(0.065)	(0.061)	(0.071)	(0.097)	(0.113)	(0.149)
Alpha, No HS-36	0.590***	−0.195	−0.214**	−0.087	−0.090	−0.047	−0.017	0.230***	0.369***	−0.185	0.772***
	(0.115)	(0.133)	(0.090)	(0.097)	(0.072)	(0.065)	(0.061)	(0.071)	(0.097)	(0.113)	(0.156)
Alpha, No HS-60	0.522***	0.073	0.233***	0.092	0.056	0.070	−0.026	−0.113	−0.071	−0.123	0.645***
	(0.112)	(0.092)	(0.088)	(0.069)	(0.062)	(0.069)	(0.100)	(0.096)	(0.107)	(0.126)	(0.172)

Note: *p<0.1; **p<0.05; ***p<0.01

Heston and Sadka (2008) based their seasonality effect study on that the firms have a relatively high or low return in the same months as the previous year. They use a measure based on returns, where they go long on the upper decile and shorting the lowest, creating a high-minus-low portfolio. In contrast to the sales seasonality effect, which measures the effect using return, an exogenous variable beyond the firm's control (Grullon et al., 2020)[7]. The first panel in table 9[8] shows the return when excluding Heston and Sadka (2008) from the sample. Following the procedure Heston and Sadka (2008) by applying the Fama-French five-factor, using monthly value-weighted returns, first by lagging the return twelve months before the investment. Then using the breakpoints to remove the samples that lie within the lowest and highest portfolio and resort the sample on sales seasonality again to exclude the Heston and Sadka (2008) seasonality effect. The first alpha is slightly higher than before, excluding the effects of HS, showing a monthly return of 0.69% and 8.29% annually. Doing the same procedure on the second alpha, only this time using the average of the same previous three-month returns. For example, taking the return in January 1986 plus January

[7]A study by Keloharju et al. (2016) shows that the difference is significant, where the return seasonalities comes from systematic factors.

[8]Table 12 in the appendix shows the excess return when using Heston and Sadka (2008) method of sorting

1987 plus January 1988 then averaging it, using this as return seasonality in January 1989. Excluding HS36 gives a higher monthly alpha than the full sample with an alpha of 0.77% monthly and 9.26% annually. Excluding HS60, which is the same previous five-month average returns, reporting a lower alpha than the full sample, with a monthly result of 0.64% and 7.74% annually. All of the alphas in the sample are significant, and Grullon et al. (2020) report similar numbers, only slightly lower for all of the alphas. Thus, concluding that after excluding the Heston and Sadka (2008) seasonality, the sales seasonality premium is still effective[9].

6.4 The reason behind the premium

TABLE 10: Testing for seasonal patterns in investment variables and financial variables. The variables are explained and calculated in the data section. Using non-financial sample from CRSP and accounting variables from COMPUSTAT. A regression is run on the sales variable, the investment variable and the financial variable. The regressions also include log of market cap, firm fixed effect and year-time effect.

	Testing for Seasonal patterns in investment variables							
	IAQ 1973-2017		IACQ 1973-2017		PPEQ 1973-2017		INVQ 1973-2017	
SEA	0.007^{***} (0.002)		0.006^{***} (0.001)		0.001^{**} (0.0004)		-0.006^{***} (0.001)	
SEA_{q+1}		0.0012^{***} (0.002)		0.011^{***} (0.001)		0.0002 (0.0004)		0.006^{***} (0.001)

	Testing for Seasonal patterns in financial variables			
	BLchange 1973-2017		MLchange 1973-2017	
SEA	-0.004^{***} (0.001)		-0.005^{**} (0.003)	
SEA_{q+1}		0.001^{**} (0.001)		0.007^{***} (0.002)
Note:				*p<0.1; **p<0.05; ***p<0.01

Grullon et al. (2020) tests several hypotheses for a reason behind the seasonality of the sales. The first one is the real options theory hypothesis: to test corporate investments and financial decisions through the sales seasons. They test this theory against

[9]According to Grullon et al. (2020) a factor based on sales seasonality is uncorrelated with a factor based on return seasonality in an unreported analysis

several proxies for investment and leverage on the sales seasonality variable, SEA_q using panel data regression and controlling for additional variables, such as size, book-to-market, profitability, firm fixed effects, and year dummies. Additionally, they investigate the firm's advancement for a new season where they also estimate using the one-step-ahead value of the seasonality of the sales, SEA_{q+1} Grullon et al. (2020). The proxies used for investment variables reported in table 10 are (i) change in the total asset at quarter q relative to the total asset in quarter q-1 (IAQ), (ii) change in the current asset at quarter q relative to total asset in quarter q-1 (IACQ), (iii) change in net property plant and equipment from quarter q-1 to q relative to the total asset in quarter q-1 (PPEQ), and lastly, (IV) change in inventory from quarter q relative to total asset in quarter q-1 (INVQ) (Grullon et al., 2020). The proxies for the financial variables in the lower panel are (i) change in total debt divided by total asset (BLChange) and (ii) change in total debt scaled by total debt plus the market value of equity (MLChange) (Grullon et al., 2020). Also, by following Grullon et al. (2020), I remove firms in the lowest size quintile since the premium is small and statistically insignificant.

Using these proxies as the dependent variables in the regression, testing each one against the seasonality of the sales. The results are positive for the first three variables in the first panel containing investment variables, for both SEA_q and SEA_{q+1}. Both variables are negative during the sales season and positive against the one-period-ahead variable in the lower panel. The IAQ and IACQ being positive and significant both in and out of season show that the firms invest more prior and in-season than normally, which, according to Grullon et al. (2020) is consistent with the real options theory. For the PPEQ variable, the findings in this thesis are only significant when in season, contrary to Grullon et al. (2020) which is significant. Nevertheless, the PPEQ variable is statistically significant against SEA_q, which is consistent with the other findings that they invest more in season. The INVQ variable is first negative then becomes positive; this shows that the firms increase their inventory before the season and do not replenish it during the high season. Consistent with the results from Grullon et al. (2020). Both statistically significant financial variables are negative in high season, then positive before the season, which indicates that the firms pay off their debt during the high season and increase their debt before the high season, indicating that the firms are using debt to finance the investment in high season. The results are consistent with the findings Grullon et al. (2020).

7 Conclusion

This thesis tries to replicate the sales seasonality study from the paper when low beats high by Grullon et al. (2020). The evidence shows that going long low-season stocks and short high-season stock generates an value-weighted excess return of 7.5%. This result shows that the excess return is affected by firm size. Therefore, further testing shows that the effect is stronger on bigger firms and not driven by the small firm size effect. The effect still holds after running a regression on other firm characteristics and confirming again that the effect is stronger amongst bigger firms. Testing the sales factor from Grullon et al. (2020) against Fama-French factors shows that the sales factor has a low standard deviation and the highest Sharpe ratio. The sales factor is also uncorrelated with the Fama-French factors. Excluding the other seasonality events from the study still shows a strong effect on the sales seasonality effect and uncorrelated with them. Testing investment and financial variables show that some of the explanation behind the premium comes from the real option and leverage theory. The firms use their growth to invest in assets in season, reducing their exposure to risk and reducing leverage, further reducing their risk and, therefore, their return, showing that the seasonality effect is counter-seasonal. I can conclude that the replication of the seasonality phenomena study by Grullon et al. (2020) is successful, as the results are similar.

Replicating a recently published study in a top financial journal with such robust testing is difficult, and thus, a weakness of this thesis is that I am not able to replicate all of the research by Grullon et al. (2020), because of my limited access to data and time.

For instance, using the EDGAR database, Grullon et al. (2020) uses SEC filings to figure out investors' attention to stocks, showing that investors' inattention might be a cause of the high excess return. Furthermore, to test for neglected stocks Grullon et al. (2020) also creates a 2x2 portfolio sorted on size and idiosyncratic volatility, and their result shows that neglected stocks drive a big part of the premium. A further weakness as well is that I am not able to check the robustness of the result by using other measures of economic activities of the firms as Grullon et al. (2020). Therefore, further research would test for investors' inattention to stocks, testing for neglected stocks and robust test the result using other measures.

References

Bali, T. G., Engle, R. F., and Murray, S. (2016). *Empirical asset pricing: The cross section of stock returns*. John Wiley & Sons.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3–18.

Berk, J. and DeMarzo, P. (2017). *Corporate Finance ; global edition; 4th edition*. Pearson.

Berk, J. B., Green, R. C., and Naik, V. (1999). Optimal investment, growth options, and security returns. *The Journal of Finance*, 54(5):1553–1607.

Chang, T. Y., Hartzmark, S. M., Solomon, D. H., and Soltes, E. F. (2016). Being Surprised by the Unsurprising: Earnings Seasonality and Stock Returns. 30(1):281–323.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Grullon, G. (2019). Data. `http://www.ruf.rice.edu/~grullon/data.html`. (Accessed on 06/01/2021).

Grullon, G., Kaba, Y., and Núñez-Torres, A. (2020). When low beats high: Riding the sales seasonality premium. 138(2):572–591.

Heston, S. L. and Sadka, R. (2008). Seasonality in the cross-section of stock returns. 87(2):418–445.

Keloharju, M., Linnainmaa, J. T., and Nyberg, P. (2016). Return seasonalities. 71(4):1557–1590.

Kenneth, R. F. (2021). Data library. `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`. (Accessed on 13/01/2021).

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, 42(3):483–510.

Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3):703–708.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The Journal of Finance*, 19(3):425–442.

Wooldridge, J. M. (1989). A computationally simple heteroskedasticity and serial correlation robust standard error for the linear regression model. *Economics Letters*, 31(3):239–243.

Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach*. ISE - International Student Edition. South-Western.

TABLE 11: The portfolios are created using NYSE, AMEX and NASDAQ breakpoints, sorted into decile portfolios using AVGSEA. The value-weighted excess return of sorting on the return seasonality from the sample is reported in the table, risk-adjusted with Fama-French five-factor model. HS12 is the last year same month return, HS36 is the average of the last three same month return over three years. HS60 is the average of the last five same month returns over five years. The sample used are non-financial firms from CRSP, dating from January 1972 to December 2017.

	FF5 factor returns, excluding Heston and Sadka (2018)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
HS12 Alpha	−0.339**	−0.195	−0.214**	−0.087	−0.090	−0.049	−0.018	0.231***	0.369***	0.247*	0.541**
	(0.147)	(0.133)	(0.091)	(0.097)	(0.072)	(0.065)	(0.061)	(0.071)	(0.097)	(0.144)	(0.236)
HS36 Alpha	−0.316**	−0.182	−0.099	−0.177**	−0.113**	−0.062	−0.006	0.080	0.381***	0.800***	1.091***
	(0.143)	(0.153)	(0.122)	(0.074)	(0.057)	(0.074)	(0.072)	(0.079)	(0.116)	(0.135)	(0.235)
HS60 Alpha	−0.207	−0.360***	−0.211**	−0.196**	0.001	−0.010	−0.057	0.149*	0.389***	0.624***	0.824***
	(0.225)	(0.123)	(0.106)	(0.076)	(0.066)	(0.071)	(0.063)	(0.086)	(0.128)	(0.155)	(0.315)
Note:											*p<0.1; **p<0.05; ***p<0.01

TABLE 12: Cross-sectional regression, equal-weighted returns + momentum. The equal-weighted portfolio's are created by sorting based on AVGSEA into deciles using NYSE, AMEX and NASDAQ breakpoints.The sample uses non-financial stocks from CRSP starting January 1972 to December 2017.

	Fama-French three-factor + momentum, equal-weighted returns.										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.084	0.291**	0.304***	0.342***	0.338***	0.402***	0.351***	0.387***	0.292**	−0.055	0.147
	(0.137)	(0.120)	(0.105)	(0.087)	(0.082)	(0.093)	(0.101)	(0.119)	(0.120)	(0.151)	(0.114)
Mrktf	0.956***	0.982***	0.987***	0.950***	0.970***	0.943***	0.979***	0.956***	0.981***	0.956***	0.001
	(0.030)	(0.026)	(0.023)	(0.022)	(0.023)	(0.021)	(0.021)	(0.024)	(0.030)	(0.032)	(0.029)
SMB	1.079***	0.906***	0.878***	0.863***	0.835***	0.845***	0.846***	0.934***	0.973***	1.031***	0.048
	(0.053)	(0.048)	(0.046)	(0.033)	(0.041)	(0.044)	(0.038)	(0.046)	(0.045)	(0.054)	(0.043)
HML	−0.109	0.026	0.078*	0.046	0.124***	0.089*	0.107**	0.076	0.013	−0.097	−0.012
	(0.094)	(0.061)	(0.044)	(0.047)	(0.038)	(0.053)	(0.050)	(0.053)	(0.076)	(0.084)	(0.054)
MOM	−0.202***	−0.200***	−0.173***	−0.158***	−0.145***	−0.170***	−0.135***	−0.219***	−0.210***	−0.228***	0.027
	(0.062)	(0.049)	(0.039)	(0.030)	(0.023)	(0.039)	(0.031)	(0.049)	(0.047)	(0.062)	(0.038)
Note:											*p<0.1; **p<0.05; ***p<0.01

TABLE 13: Cross-sectional regression, value-weighted returns. The value-weighted portfolio's are created by sorting based on AVGSEA into deciles using NYSE, AMEX and NASDAQ breakpoints. The sample uses non-financial stocks from CRSP starting January 1972 to December 2017.

	Fama-French three-factor + momentum, value-weighted returns.										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(L-H)
Alpha	0.252*	0.203**	−0.034	0.260***	0.123	0.177*	0.044	0.063	−0.072	−0.144	0.397**
	(0.131)	(0.103)	(0.097)	(0.084)	(0.081)	(0.092)	(0.082)	(0.118)	(0.118)	(0.166)	(0.198)
Mrktf	1.063***	1.014***	1.004***	0.980***	0.924***	0.908***	0.990***	1.025***	0.999***	0.984***	0.079
	(0.034)	(0.027)	(0.027)	(0.020)	(0.026)	(0.021)	(0.027)	(0.029)	(0.038)	(0.042)	(0.051)
SMB	0.356***	0.119***	0.029	−0.026	−0.053*	−0.071**	0.051	0.129***	0.145***	0.353***	0.003
	(0.038)	(0.036)	(0.038)	(0.030)	(0.030)	(0.036)	(0.062)	(0.045)	(0.053)	(0.079)	(0.088)
HML	−0.329***	−0.160***	−0.068**	−0.055	0.014	−0.088**	−0.039	−0.215***	−0.035	−0.271***	−0.058
	(0.062)	(0.041)	(0.034)	(0.039)	(0.043)	(0.045)	(0.035)	(0.054)	(0.055)	(0.074)	(0.077)
MOM	−0.022	0.013	0.038	−0.0002	−0.006	0.004	−0.024	−0.047	0.001	−0.113	0.091
	(0.046)	(0.026)	(0.029)	(0.025)	(0.027)	(0.029)	(0.034)	(0.030)	(0.043)	(0.069)	(0.072)

Note: *p<0.1; **p<0.05; ***p<0.01

Newey and West adjustment

Newey and West (1987) developed a method that adjusts the standard errors, the reason is that time series may contain both heteroscedasticity or autocorrelation. Making the t-statistics and p-values inaccurate, therefore many researchers adjust for this in their research. If the model is

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 \hat{x}_t + \hat{e}_i. \tag{32}$$

If the model has heteroscedastic errors, it means that the standard errors increase with the function, instead of being constant (Wooldridge, 2009). This can be written as a function:

$$var(\hat{e}_i|\hat{x}_i) = f(\hat{x}_i), \tag{33}$$

where the variance increases with the function x and inflating the standard error. This is important because it violates the Gauss-Markov assumption that the ordinary least squares (OLS) estimator has the lowest amount of variance. Thus, it is not the best linear unbiased estimator (BLUE) anymore. In short, it means that there are other estimators with a lower variance that more often get closer to the true population. If the data contains autocorrelation, the covariance between two errors is not zero, inflating the standard errors. In particular $cov(\hat{e}_i, \hat{e}_s) \neq 0$ Again this makes it so that the OLS is no longer BLUE again, making other estimators better (Wooldridge, 2009). Applying Newey and West (1987) which adjust for correlation of errors over time and heteroscedasticity. Wooldridge (1989) the first step is to normally estimate the standard model regression using OLS to get the standard errors and the residuals $\hat{e}_1 ... \hat{e}_t$. Taking the residuals and run an auxiliary estimate to test for heteroskedasticity and autocorrelation. The null hypothesis is that there is constant variance in the error terms. If the variance is not constant, then proceeding to adjust for Newey and West. The formula for Newey-West standard error is:

$$se(\hat{\beta}_j) = [\frac{se(\hat{\beta}_j)}{\hat{\sigma}}]^2 * \sqrt{\hat{\delta}} \tag{34}$$

First lets say that $\hat{\xi}_t = \hat{e}_t \hat{u}_t$ then $\hat{\omega}_s = \sum_{t=s+1}^{N} \hat{\xi}_t \hat{\xi}_{t-s}$ thus, $\hat{\delta}$ can be written:

$$\hat{\delta} = \omega_0 + 2 \sum_{s=1}^{G(N)} \eta(s, G(N)) \hat{\omega}_s \tag{35}$$

The first part of the equation is the heteroskedasticity standard errors, and the second part is the autocorrelation part. Choosing the number of lags is easily calculated. FBali et al. (2016), choosing the amount of lag is done by the formula $4(T/100)^a$ where a = 2/9 [10] and in this case T = 552, this results in a value close to six and thus, both this paper and Grullon et al. (2020) uses six lag when calculating Newey-West standard errors. The Newey-West adjustment is luckily easily implemented using modern software.

[10] See Newey and West (1987) for a further discussion on why a = 2/9.

NTNU

Norwegian University of
Science and Technology