Stock Return Predictability and Seasonality

Keunsoo Kim

Graduate School of Pan-Pacific International Studies, Kyung Hee University

1732 Deogyoung-daero, Giheung-gu, Yongin-si, Gyeonggi-do, 17104, Korea

keunskim@khu.ac.kr, 010-8951-9586

Jinho Byun

School of Business, Ewha Womans University

52 Ewhayeodae-gil, Seodaemun-gu, Seoul, 03760, Korea

jbyun@ewha.ac.kr, 010-3685-6101

November 2018

Comments are welcome.

Stock Return Predictability and Seasonality

Abstract

An examination of the Shiller cyclically adjusted pricing-earnings (CAPE) ratio reveals its forecasting power for 12-month CRSP equally weighted (EW) and value weighted (VW) excess returns. The 12-month EW excess returns following low CAPE ratios are, on average, 20.7% higher than those following high CAPE ratios for the period of 1927–2016. This dichotomy in the Shiller CAPE ratio has more reliable predictability than the January barometer. Previous studies report that the Halloween indicator was weak or negative in the US stock market prior to the 1950s. We find that the Halloween effect is strongly present following high CAPE ratios, even for the period of 1926–1971. Our results recommend a practical investment strategy. More specifically, if the CAPE ratio in September is lower than the 36-month median of the CAPE ratio, invest in stock markets from November to October of the following year; otherwise, invest for five months from November to March and sell in April and go away.

Keywords: Cyclically adjusted price-earnings (CAPE) ratio, January barometer, Halloween effect, Predictability, Seasonal anomaly, Asset Price Forecasts

1. Introduction

Campbell and Shiller (1988, 1998) produce a market valuation ratio referred to as the cyclically adjusted price-earnings (CAPE) ratio. The CAPE ratio is characterized as having smooth earnings which represent the average of the real earnings over the past 10 years, as Graham and Dodd (1934) suggested. Campbell and Shiller (1998) state that the CAPE ratio has a negative relationship with future 10-year real market returns. Such a negative relationship indicates the mean reversal of the CAPE ratio over the long run. However, they report that the CAPE ratio has little forecasting power for a one-year real stock market return when they measure the stock market return by using the Standard and Poor's (S&P) composite index, currently known as S&P 500 index.

On the other hand, Kothari and Shanken (1997) and Pontiff and Schall (1998) show that the book-to-market ratio of the Dow Jones Industrial Average (DJIA) has a positive relationship with the subsequent annual returns of the Center for Research in Security Prices (CRSP) equally weighted (EW) index and value weighted (VW) index. Both studies illustrate that the predictability of the DJIA's book-to-market ratio is strong for the EW returns for the entire sample period that begins with the 1920s. However, the predictability is poor in the latter sub-periods of 1963–1991 (Kothari and Shanken, 1997) and 1959–1994 (Pontiff and Schall, 1998). Pontiff and Schall (1998) report that the book-to-market ratio of the S&P index is slightly better at forecasting market returns in the post 1959 period. As they conjecture, the poor ability of the DJIA's book-to-market ratio may indicate that the DJIA is less representative of the equity market during the post-1960 period.

Robert Shiller's CAPE ratio is one of the best forecasting predictors for long-term stock market returns. If the market predictability of the CAPE ratio is superior to, or at least as good as, that of S&P's book-to-market ratio, the CAPE ratio may be able to predict one-year market returns (e.g., CRSP EW returns). By using the dichotomous approach of the CAPE ratio, this paper examines the predictive power of the CAPE ratio for one-year CRSP EW and VW returns. This paper finds that a low CAPE ratio signals higher EW and VW returns over the subsequent 12 months. The paper compares the predictability of the CAPE ratio to that of the January barometer and investigates the possible relationship between the predictability of the CAPE ratio and the Halloween indicator.

The January barometer, also known as the Other January effect, says that the positive January return forecasts a bull market over the next 11 months, while the negative January return signals a bear market [Cooper, McConnell, and Ovtchinnikov, 2006; Dzhabarov and Ziemba, 2011]. Cooper et al. (2006) report that January returns have a reliable predictive power for the following 11-month market returns for the period of 1940–2003. They show that the January barometer has been valid since 1825, but fails to predict future returns during the period of 1929–1939, which is characterized as the market crash and Great Depression period.

For the period of 1927–2016, this paper shows that the CAPE ratio has a predictive power for CRSP EW and VW excess returns, but the January barometer does not. When the sample period begins with 1940, the January barometer regains the predictive power for the EW and VW excess returns. However, our results indicate that low CAPE ratios have more reliable predictability for the EW excess returns, even for the period of 1940–2016, while the January barometer has stronger predictive power for the CRSP VW excess returns for the same period.

The Halloween indicator predicts poor stock market performance during the summer months (May–October), suggesting that investors are better off selling their stocks in late April and coming back to the market on Halloween (October 31). The long-term historical data has shown that the Halloween indicator is valid in most international stock markets (Bouman and Jacobsen, 2002; Jacobsen and Zhang, 2014) and in US stock market sectors and industries (Jacobsen and Visaltanachoti, 2009). The effect is stronger in small stocks (Dzhabarov and Ziemba, 2010) and present in the foreign currency market, credit market, and equity volatility risk premium (Andrade et al., 2013). On the contrary, Haggard and Witte (2010) and Jacobsen and Zhang (2014) find no significant Halloween effect in the US prior to 1960.

This paper confirms the validity of the Halloween indicator by showing that the average excess return is higher during the winter (November–April) than during the summer (May–October). The difference in the average return between these two periods, referred to as the Halloween effect, is statistically significant at the 1% level, in the case of EW excess returns from 1926 to 2016, and at the 5% level in the case of VW excess returns. However, when monthly excess returns from November to October are measured following either the high or low CAPE ratio in September, the Halloween effect strongly appears only following the high CAPE ratios. The Halloween effect is positive and statistically significant at the 1% level in the case of both EW and VW excess returns following the high CAPE ratios, while it is negative following the low CAPE ratios for the period of 1926–2016.

The sub-period analysis confirms that the validity of the Halloween indicator depends on the dichotomy of the CAPE ratios. Bouman and Jabobsen (2002) argue that the Halloween effect is a new puzzle that we have yet to solve. Many hypotheses (e.g., seasonal pattern of vacation, seasonal affected disorder (SAD) effect, and optimism cycle hypothesis) attempt to explain the Halloween effect. This paper challenges these hypotheses by asking why the Halloween effect depends on the preceding market valuation ratio (e.g., CAPE ratio) and provides alternative hypothesis.

The rest of this paper is organized as follows. In Chapter 2, we provide the basic results about the predictability of the CAPE ratio for CRSP EW and VW excess returns. We also propose the mispricing hypothesis to explain why the CAPE ratio have a predictive power for CRSP EW returns in the short run but not for the returns of S&P 500 index. Chapter 3 examines the predictability of the CAPE ratio and compares its predictability to that of the January barometer. Chapter 4 investigates the relationship between the predictability of the CAPE ratios and the Halloween effect. In addition, the robustness of our empirical results is examined in Chapter 4. Equity risk, the seasonal pattern of vacation, the seasonal affective disorder (SAD) effect, the optimism cycle hypothesis, and the mispricing-optimism cycle hypothesis are discussed in Chapter 5, as possible explanations for our results. Chapter 6 summarizes our main results and concludes.

2. Predictability of the CAPE Ratio

2.1 Data and Basic Results

We analyze the monthly returns of the CRSP EW and VW indices, including dividends, from January 1926 to December 2016. The CRSP market index includes equities listed on the NYSE, NYSE American, and Nasdaq markets. For most of our analyses, we use excess returns, defined as the CRSP index return minus the one-month Treasury bill (T-bill) rate. The one-month T-bill rates were obtained from Kenneth French's website. Since the one-month T-bill rates are available beginning in July of 1926, the sample period using the excess returns begins with July of 1926. The Shiller CAPE ratios and S&P 500 Index were downloaded from the website of Online Data Robert Shiller.[1]

The monthly CAPE ratio is sorted as either a high or low CAPE ratio, based on the 36-month median of the CAPE ratio. If the CAPE ratio in month *t* is higher than the median CAPE ratio from *t*-35 to *t*, month *t* is defined as a high CAPE month; otherwise, month *t* is defined as a low CAPE month. Since the classification of CAPE ratios only uses the past CAPE ratios, this dichotomous approach is free of the look-ahead bias.[2]

Figure 1 illustrates the monthly raw returns of the S&P 500 index over 24 months, following the monthly sorting of the CAPE ratio. For the sample period 1926-2016, the monthly soring produces 637 high CAPE months and 432 low CAPE months that cover December 1925 through December 2014. The monthly sorting of the CAPE ratio, with a 24-month holding period, generates the maximum of the 24-month overlapping returns in a month of the holding period.

Insert <Figure 1>

According to Figure 1, the monthly returns of the S&P index following the low CAPE months begin to outperform those following the high CAPE months, eight months after sorting the CAPE ratio. The 12-month cumulative returns following the low CAPE months average 6.36% which is a little lower than 7.37% of the 12-month returns following the high CAPE months.

---

[1] We express great thanks to Kenneth French and Robert Shiller for making their data available on websites.

[2] CAPE ratios can be divided into high or low CAPE ratios, based on the median of all CAPE ratios in the sample period. However, this classification methodology does not improve the predictive power for future one-year equity returns, although it is exposed to the look ahead bias. Siegel (2006) points out that the recent forecasts of the CAPE ratio may be too pessimistic because of the changes in the computation of GAAP earnings in 1993, which causes the upward bias of the CAPE ratio. In addition, any structural change in the economy (e.g., IT development) may fundamentally shift the level of a CAPE ratio. Such a structural change or change in calculation of earnings is less likely to influence our classification methodology of the CAPE ratios, because the classification of a high or low CAPE ratio only depends on the 36-month moving median of the CAPE ratio.

The 24-month average cumulative return following the low CAPE months is a little higher than following the high CAPE months, but their difference is only 1.8%. Thus, the CAPE ratio has little forecasting power for the returns of the S&P 500 index in the short run.

Figure 2 reports the CRSP EW monthly returns over 24 months, following low and high CAPE months. The monthly returns over 24 months following the low CAPE months are much higher than those following the high CAPE months. The average of the 24-month cumulative return following the low CAPE months is 45.7%, while the average following the high CAPE months is 19.5%. Thus, the CAPE ratio appears to have a predictive power for CRSP EW returns over 24 months.

Insert <Figure 2>

2.2 Mispricing Hypothesis

Campbell and Shiller (1998) argue that high (low) CAPE ratios indicate an over (under) market valuation.[3] The high or low CAPE ratio tends to move to the mean of the CAPE ratio by adjusting the prices of S&P 500 stocks over 10 years. This mean reverting of the CAPE ratio suggests that the CAPE ratio exhibits an ability to predict the subsequent S&P 500 index in the long run. However, they fail to find any predictive power of the CAPE ratio for the subsequent S&P 500 index over one year. <Figure 1> confirms that the CAPE ratio does not forecast the returns of S&P 500 index in the short run. However, <Figure 2> suggests that the CAPE ratio may be able to predict the CRSP EW returns over 24 months. Why can the CAPE ratio gain

[3] A high CAPE ratio can be justified by the high growth of future cash flows, such as dividends or earnings. However, Campbell and Shiller (1998) provide evidence that CAPE ratios are not related to the future long-term growth of the 10-year moving average of earnings.

the predictive power for the EW returns in the short run but not for the returns of S&P 500 index?

The predictability of the CAPE ratio may be related to the fact that the degrees of short sale constraints are different across stocks. The S&P 500 index is a market-capitalization index which are currently composed of 500 leading firms listed on the NYSE or Nasdaq. Since these firms are mostly well-known large companies and their stocks are actively traded, their short-sale constraints may be least among US equities. Further, negative private information has been able to be easily reflected in S&P 500 index through S&P 500 index options since 1983. Thus, S&P 500 index will be valued by integrating both optimistic and pessimistic (or rational) opinions of their fundamental values during high CAPE ratios. As a result, the prices of S&P 500 stocks may be regarded as a weighted average of two opinions.[4] When the CAPE ratio is higher, the influence of optimistic opinions may outweigh that of rational investors because new optimistic investors join the market. However, S&P 500 index may not be expected to have subsequent low returns following high CAPE ratios unless the weight of rational (or pessimistic) investors sharply increases.

In contrast, short sale constraints enable optimistic investors to overprice stocks by impeding short sales of rational investors. Thus, stocks with heavy short sale constraints more frequently get overpriced. On the other hand, stocks may not be frequently underpriced because rational investors can remove undervaluation of stocks by simply buying the stocks. Such asymmetric arbitrage ability will cause stocks with heavy short sale constraints to get more overvalued than undervalued (Miller, 1977; Jones, and Lamont, 2002; Nagel, 2005; Boehme et al., 2006;

[4] Many academic papers provide reasons why arbitrage activities of rational investors are limited in eliminating asset mispricing. Please refer to De Long et al. (1990), Campbell and Kyle (1993), and Shleifer and Vishny (1997).

Lamont, 2012; Stambaugh, Yu, and Yuan, 2015). Further, the CAPE ratio may play a role in forecasting when stocks with heavy short sale constraints tend to get overvalued. More specifically, when the CAPE ratio is high, optimistic investors are more likely to actively purchase stocks with optimistic newcommers joining stock markets. As a result, stocks with heavy short sale constraints are more likely to get overvalued when the CAPE ratio is high. The divergence between optimistic and pessimistic (or rational) opinions will be wider as the CAPE ratio is higher. However, pessimistic opinions will be eventually reflected in valuing the overvalued stocks following the high CAPE ratios, resulting in subsequent poor returns.[5] Since the EW returns consist of many stocks with heavy short sale constraints, the CAPE ratio may have a predictive power for the EW returns in the short run.

In summary, the predictive power of the CAPE ratios will be stronger for stocks with heavy short sale constraints in the short run. Likewise, the predictive power of the CAPE ratios will be stronger for CRSP EW returns than for CRSP VW returns because the EW returns are more influenced by returns on stocks with severe short sale constraints. For example, Lamont (2012) states that small stocks are generally difficult to borrow for short sale and tend to have heavy short sale constraints, compared to large stocks. By giving equal weight to every stock, the CRSP EW return is more influenced by performance of smalls stocks than that of large stocks. We refer this hypothesis to as the mispricing hypothesis that attempt to explain why the CAPE

[5] First of all, a series of information signals that the prices of these stocks are beyond their fundamental values. Second, as the divergence between negative and optimistic opinions is wider, the demand of short sale will increase given the same level of short sale constraints. Cohen et al. (2007) show that increases in shorting demand lead to poor returns. In addition, Lamont (2012) also reports that firms taking anti-shorting actions has about -2% per month of abnormal returns in the following year. This result indicates that increases in short sale constraints may delay immediate correcting of overvaluation but cannot ultimately prevent negative opinions from affecting stock prices.

ratio does not have a short-term predictive power for S&P 500 index returns but for CRSP EW returns.

2.3 Twelve-Month Predictability

We examine whether the CAPE ratio's predictability for one-year market returns is statistically reliable by measuring 12-month cumulative excess returns one month after sorting the CAPE ratio. For example, the second row of Table 1 shows the 12-month cumulative EW excess returns from January to December after the CAPE ratio is determined as high or low in November of the previous year. One month is skipped between the sorting period and the holding period to make sure of the availability of the CAPE ratios and mitigate any market microstructure bias (e.g., non-synchronous trading bias).

Table 1 confirms that the CRSP EW excess returns are much higher following low CAPE months than following high CAPE months. The return of (Low–High) is defined as a difference in the average excess returns between holding periods following low and high CAPE months. The annual EW returns of (Low–High) continue to be positive and statistically significant when the beginning month of the holding period is November through June. Thereafter, those of (Low – High) become weaker and are not statistically significant when the beginning month is July, August, and October. These results imply that the predictability of the CAPE ratio may also be related to seasonal anomalies (e.g., Halloween indicator). The annual VW returns of (Low–High) are statistically significant only when the beginning month is November or January, which suggests that the predictability of the CAPE ratio is weak for VW excess returns. This result is consistent with the results of Kothari and Shanken (1997) and Pontiff and Schall

(1998) who also show that the predictability of DJIA's book-to-market ratio is weak for CRSP VW returns.

Insert <Table 1>

3. Predictability of the CAPE Ratio and January Barometer

3.1 January Barometer and the CAPE Ratio

The January barometer, also known as the Other January effect, indicates that the market return in January has the predictive power for the rest of the year. Yale Hirsch first mentioned the January barometer in 1972 (Hirsch and Hirsch, 2008). The January barometer states that a positive market return in January signals a bull market for the remaining eleven months, while a negative market return in January signals a bear or flat market for the eleven months.

Cooper et al. (2006) investigate the predictive power of January returns over the period of 1940–2003. They find that the January barometer is a reliable predictor for the subsequent 11-month EW and VW returns. They report that the January barometer has been valid since 1825, except the period of 1929-1939, which covers the Wall Street Crash of 1929 and the Great Depression period. However, Bohl and Salm (2010) document that only 2 out of 19 countries' stock markets have the reliable January barometer and conclude that the January barometer is not an international phenomenon.

Since both the January barometer and the CAPE ratio have return predictability, we examine the predictability of both the CAPE ratio and the January barometer. Table 2 illustrates the empirical results of the following regression model:

$$r_t = \alpha + \beta_1 LOW_{t-3} + \beta_2 JBR_{t-1} + \varepsilon_t \quad \text{----- Eq (1)}$$

where r_t is the monthly excess return between February and December. LOW is a dummy variable, which is equal to one when the CAPE ratio in November of the previous year is not higher than the median CAPE ratio over 36 months; otherwise, LOW is zero. JBR is a dummy variable representing the January barometer. JBR is one when the January return is positive, and zero when the January return is non-positive. Since monthly excess returns are measured beginning with February, LOW and JBR lag a dependent variable by at least three months and one month, respectively.

Table 2 shows that LOW adds 2.68% (t=1.77) to the January EW excess return for the period of 1927–2016. The estimate of LOW is also positive in January for the CRSP VW excess returns, but is not statistically significant. In regards to the EW and VW excess returns from February to December, the coefficient estimates of JBR are positive, but are not statistically significant. Thus, the January barometer does not have any reliable predictive power for the next 11-month returns.

On the other hand, the LOW dummy variable estimate is 1.64% (t=3.41) in the case of CRSP EW returns. This indicates that the eleven-month excess returns are 18.0% higher following low CAPE months than following high CAPE months. The estimate of the intercept is –0.12% (t = –0.53), indicating negative monthly excess returns from February to December following high CAPE months. Like the EW excess returns, the CAPE ratio has a predictive power for the VW excess returns during the rest of year, although the January barometer does not. However, the predictive power of the CAPE ratio becomes weaker for CRSP VW excess returns. The estimate of LOW is 0.66% (t=1.82) of the VW excess return when LOW is only included as an independent variable. The estimate of the intercept is only 0.30% (t=1.59) for VW returns, suggesting that the average of the VW monthly returns is very low following high CAPE months.

Regarding monthly EW excess returns from January to December, the estimate of LOW is 1.72% (*t*=3.71), indicating that annual EW excess returns following low CAPE months are, on average, 20.7% higher than those following high CAPE months. In the case of VW excess returns, the estimate of LOW is 0.63% (*t*=1.86), which is still statistically significant at the 10% level. Thus, the dichotomy of the CAPE ratio has the forecasting power for EW and VW excess returns over 11 or 12 months for the period of 1927–2016, while the January barometer does not.

Insert <Table 2>

It is noteworthy that the January barometer does not have any reliable predictive power for the remaining 11-month returns during the period of 1926–2016. This result is not consistent with the results of Cooper et al. (2006), showing that the January barometer is a reliable predictor for the next 11-month returns during the period of 1940–2003. The different result may be due to the inclusion of the period of 1929–1939, when the January barometer yields –25.29% for the NYSE VW returns (Cooper et al., 2006, p. 325). The period of 1929–1939 may be sample-specific, since it is characterized as the Wall Street Crash of 1929 and the Great Depression.

We reexamine the validity of the January barometer, during the period of February 1940–December 2016. Although the predictability of the January barometer is not as strong as in Cooper et al.'s (2006) results, the estimate of the January barometer dummy remains statistically significant at the 5% level in the case of the EW returns and at the 1% level in the case of the VW returns. When we add JB and LOW together as independent variables into the regression model, both dummy variables are found to have reliable predictive powers for the

EW and VW excess returns. However, Table 2 illustrates that LOW has a more reliable predictive power for the EW excess returns, while the January barometer has a more reliable predictive power for the VW excess returns.

3.2 Sub-period Analysis

The aggregate book-to-market ratio has weak predictability for 12-month CRSP EW and VW returns in the latter sub-periods of 1963–1991 (Kothari and Shanken, 1997) and 1959–1994 (Pontiff and Schall, 1998). Since both the Shiller CAPE ratio and the aggregate book-to-market ratio reflect common valuation factors (e.g., risk, future cash flows, growth rate of cash flows), they will be highly negatively correlated with each other.[6] Thus, the CAPE ratio may also have poor predictive power in the second half of 1927–2016. In addition, Hulbert (2017) argues that the January barometer's reliability has been weaker since 1972, after Yale Hirsch first mentioned it. Thus, the sub-period analysis is required to examine the predictive power of the CAPE ratio and the validity of the January barometer, especially during the second half of 1927–2016.

The entire sample period is divided into two sub-periods to examine the robustness of the major results. The first sub-period is January 1927 through December 1971, covering 24 years following high CAPE months and 21 years following low CAPE months. The second sub-period is January 1972 through December 2016, covering 26 years following high CAPE months and 19 years following low CAPE months. Panel A of Table 3 shows the monthly

---

[6] The market valuation is a nominator in the case of the CAPE ratio, but a denominator in the case of the aggregate book-to-market ratio. As a result, their valuation ratios will be negatively correlated with each other.

returns for January, February through December, and January through December. The monthly returns following low CAPE months consistently exceed those following high CAPE months in both sub-periods.

The EW return of (Low–High) is 1.87% (t=2.46) from January to December during the period of 1927–1971, while it is 1.54% (t=3.03) during the period of 1972–2016. Although the magnitude of the return difference is slightly smaller during the second sub-period, its statistical significance is not weak at all during the second sub-period. Unlike the aggregate book-to-market ratio, the CAPE ratio has the reliable predictive power for EW excess returns from January to December, even in the second sub-period. However, low CAPE months do not have a reliable predictive power for January returns in the first sub-period, although January returns are higher following low CAPE months than following high CAPE months.

All VW excess returns following low CAPE months are higher than those following high CAPE months, except the January return in the first sub-period. The VW return of (Low–High) is 0.89% (t=1.65) from January to December in the first sub-period which is marginally significant at the 10% level. However, it is only 0.35% (t=0.87) in the second sub-period. In general, the predictive power of the CAPE ratio is much weaker for CRSP VW excess returns.

Insert <Table 3>

Panel B of Table 3 presents the results of the January barometer investigation for the two sub-periods. The monthly returns following positive January returns tend to be higher than following negative January returns for the rest of year. However, the January barometer has not been a reliable predictor for CRSP EW excess returns in both sub-periods and for CRSP VW excess returns in the first sub-period. The return of (Positive–Negative) is 0.05% (t=0.08) of

VW returns in the first sub-period, but 0.80% (t=1.91) of VW returns in the second sub-period. Thus, the January barometer has a reliable predictive power for VW excess returns in the second sub-period when the CAPE ratio fails to predict the next 11-month VW returns. However, the predictive power of the January barometer becomes weak during the period of 1972–2016, compared to the period of 1940–2016 in Table 2, which indicates the dwindling predictability of the January barometer.

In general, the CAPE ratio has reliable and consistent predictive power for CRSP EW excess returns, compared to the January barometer. However, the CAPE ratio does not have as reliable predictive power for VW returns as for EW returns. Although the January barometer has more reliable predictability of VW returns in the second sub-period than the CAPE ratio, its predictability has become weaker recently as Hulbert (2017) points out.

4. Predictability of the CAPE Ratio and the Halloween Indicator

4.1 Relationship between the Halloween Indicator and the CAPE Ratio

The Halloween indicator is considered another streetlore like the January barometer. It is also known as the "Sell in May" effect. The indicator predicts poor stock market performance from May to October, suggesting that investors are better off selling their stocks in late April and coming back to the market on Halloween, which is October 31. Bouman and Jacobsen (2002) investigate whether stock market returns are significantly lower during the summer (May–October) than during the winter (November–April). When the Halloween effect is defined as a difference in monthly returns between the winter and the summer, the effect is found to exist in 36 out of 37 countries in their sample. However, the effect is not statistically significant at the 10% level in 17 countries out of 37 (Bouman and Jacobsen, 2002, p. 1622).

Jacobsen and Zhang (2014) show that the Halloween effect has strengthened in recent years. Furthermore, they present that 46 out of 65 stock markets have negative excess returns during the summer. Using data for 33,348 monthly returns obtained from pooling 65 countries' stock markets, they report a –1.17% excess returns (t=–3.36) during the summer, violating the risk-return trade-off. Since their data includes all available stock markets worldwide, with full available years of stock market indices, they argue that their results are the best evidence against data mining and sample selection bias. Thus, the Halloween effect is a persistent and strong anomaly in the international stock markets.

However, the Halloween effect is not a strong phenomenon in the US market. Bouman and Jacobsen (2002) report that the Halloween indicator is statistically significant at the 10% level in the US from 1970 to 1998, but its significance is slightly less than the 10% level when the January return is excluded from the winter returns. Haggard and Witte (2010, p. 382) document that the Halloween effect does not exist in the US for the period of 1926–1953. The table 6 of Jacobsen and Zhang (2014) also show that the Halloween effect is negative in the US between 1931 and 1950. Furthermore, after analyzing sub-periods covering over 60 years and 28 countries, Jacobsen and Zhang (2014, p. 25) state that "*The Halloween effect seems to be a phenomenon that emerges from the 1960s and has become stronger over time, especially among the European countries*".

We examine the validity of the Halloween indicator following high and low CAPE months to identify the possible relationship between the Halloween effect and the predictability of the CAPE ratio. The CAPE ratio is annually classified into either a high or low CAPE ratio in

September. Following the high and low CAPE months, monthly returns are measured from November to October of the following year. Total in Table 4 indicates the average of monthly excess returns for all months (1,086 months) of the entire sample period July 1926 – December 2016. The number of months following the high CAPE months is 594 months (49 years and 6 months), while that following the low CAPE months is 492 months (41 years). Winter in Table 4 indicates the average of the monthly excess returns in the winter (November–April). Winter(J) is the winter returns excluding the January returns. Summer in Table 4 is defined as the average monthly return in the summer (May–October).

The Halloween effect is measured by the return of (Winter–Summer). Total in Table 4 illustrates that the EW return of (Winter–Summer) is 1.32% (t=3.00) per month so that the Halloween effect strongly appears. However, when the January return is excluded, the Halloween effect becomes only 0.66% (t=1.48), suggesting that the effect heavily relies on the January return in the case of EW excess returns. In the case of CRSP VW returns, however, the Halloween effect remains statistically significant (0.61%, t=1.81), even when the January effect is eliminated. In general, the Halloween effect exists in the US stock market but is not strong for the period of 1926–2016; this result is consistent with the results of earlier papers.

Table 4 also shows the monthly excess returns following low and high CAPE months. The Halloween effect completely disappears for both the CRSP EW and VW monthly returns following low CAPE months. The return of (Winter–Summer) turns out to be negative following low CAPE months. Thus, there is no Halloween effect at all following low CAPE months. The average of the CRSP EW return in the summer is 1.78% (t = 2.96), which is 0.12% higher than the winter return following low CAPE months. When the January return is excluded, the summer return is 0.91% higher than the winter return. Similarly, the average VW monthly return is 1.10% ($t = 2.65$) during the summer, which is 0.30% higher than that during the winter.

Insert <Table 4>

Table 4 illustrates that the Halloween effect strongly appears following high CAPE months. The EW and VW returns of (Winter–Summer) are 2.54% (t=5.63) and 1.43% (t=3.85) per month, respectively, following high CAPE months. Even when January returns are excluded, those of (Winter(J)–Summer) are above 3.50 times their standard errors for both EW and VW returns. The prominent Halloween effect following high CAPE months is mainly due to negative summer returns following high CAPE months. The return of (Low–High) is negligible for winter returns, including January returns, indicating that there is no significant difference in winter returns between holding periods following low and high CAPE months. However, the summer return of (Low–High) is 2.69% (t=3.89) for EW returns and 1.42% (t=2.79) for VW returns. Because of such a significant difference in the summer returns, the Halloween effect is strongly present following the high CAPE months. Further, it is noteworthy that summer returns following high CAPE months are negative for EW returns (–0.91%; t=–2.65) and for VW returns (–0.32%; t=–1.08). The negative summer returns following high CAPE months mainly contribute to reliably positive summer returns of (Low–High).

4.2 Robustness

According to Table 6 in Jacobsen and Zhang (2014), the Halloween effect is negative in the US for several sub-periods prior to 1951. Thus, it is interesting to reexamine the robustness of the Halloween effect and its relationship with the high CAPE months, particularly in the first half of 1926-2016. Table 5 reports that the Halloween effect strongly appears following high CAPE months in the first and second sub-periods which indicate the period of November 1926-October 1971 and the period of November 1971- December 2016, respectively. The EW return of (Winter–Summer) is 3.12% (*t*=4.37) per month in the first sub-period (Panel A) and 2.04% (*t*=3.55) in the second sub-period (Panel B). In the case of VW excess returns, the Halloween effect following high CAPE months becomes weaker, but their statistical significances are strong enough to reject the null hypothesis of no Halloween effect at the 1% level in the first sub-period and at the 5% level in the second sub-period.

Insert <Table 5>

The Halloween effect is weak in the case of the total months (540 months) in the first sub-period. The EW and VW returns of (Winter–Summer) are 0.71% (*t*=0.97) and 0.37% (*t*=0.69), respectively, in the first sub-period. However, the Halloween effect is reliably strong in the case of the total months (542 months) for the second sub-period; the EW and VW returns of (Winter–Summer) are 1.95% (*t*=4.10) and 0.94% (*t*=2.43), respectively. This is consistent with Jacobsen and Zhang (2014)'s result that the Halloween effect has been strong since the 1960s.

The difference in the Halloween effect between these two sub-periods is mainly due to the reversal of the Halloween effect following the low CAPE months in the first sub-period. Panel A of Table 5 shows that the Halloween effect is –1.81% (*t*=–1.40) of the EW excess returns and –1.15% (*t* =–1.29) of the VW excess returns following the low CAPE months for the first sub-period. The reversal of the Halloween effect is caused by the high summer returns following the low CAPE months, which is 2.78% (*t*=2.71) for the EW excess return and 1.64% (*t*=2.46) for the VW excess return. In contrast, the Halloween effect is strongly positive following low CAPE months for the second sub-period, mostly because of the low summer returns.

There exists a substantial difference in the summer returns between holding periods following low and high CAPE months, even in the second sub-period. More specifically, the summer EW return of (Low–High) is 1.44% ($t = 2.16$) for the second sub-period. It is interesting to note that summer returns following high CAPE months are always negative, regardless of the EW or VW returns in both sub-periods. The Halloween effect strongly emerges in the US stock market following high CAPE months, mostly because of consistent negative summer returns.

The sub-period analysis confirms the robustness of our primary result in that the Halloween effect strongly appears following high CAPE months. However, by examining the period 1970–1998, Maberly and Pierce (2004) report that the Halloween effect is driven by two negative return outliers: The Crash of October 1987 and the collapse of the hedge fund Long-Term Capital Management in August 1998. Thus, several extreme events that occurred after or around the high CAPE months could cause extremely poor returns in the summer following high CAPE months.

Indeed, noticeable market crashes or crises occurred when CAPE ratios were high. According to Wikipedia, the Wall Street Crash of 1929 began on October 24, 1929. September 1929 is classified as a high CAPE month (Appendix 1). The six-month EW and VW excess returns from November 1929 to April 1930 are 4.00% and 1.59%, respectively. In contrast, the six-month EW and VW excess returns from May to October in 1930 are –46.3%, and –35.1%, respectively. The Financial Crisis of 2008 began in 2007 with the subprime mortgage crisis and peaked in September 2008, when Lehman Brothers filed for bankruptcy and the US government announced the bailout of $85 billion for AIG. September 2007 is also defined as a high CAPE month. Following this month, the six-month EW and VW returns from November 2007 to April 2008 were –16.2%, and –11.0%, respectively, while those from May to October were –37.7%, and –34.8%, respectively. Did these extreme events accidently occur following high CAPE months and generate extremely low returns during the summer period May to October? If the small number of extreme events drive the Halloween effect, our results may be subject to data snooping bias.

To mitigate the impact of extreme events on the Halloween effect, we exclude the five years exhibiting the worst six-month EW returns in the summer following high CAPE months in the history of the US stock market. These high CAPE months are the Septembers of 1929, 2007, 1936, 1965, and 1997, which include the Wall Street Crash of 1929 and the Financial Crisis of 2007-2008. The exclusion of these five years, from the 49 years following the high CAPE months, decreases the Halloween effect from 2.54% ($t = 5.63$) to 2.01% ($t = 4.47$) of EW excess returns (Appendix 2). Likewise, the Halloween effect following high CAPE months decreases from 1.43% ($t = 3.85$) of VW excess returns to 1.08% ($t = 2.90$). Our main results, however, are still strong enough to reject no Halloween effect at the 1% significance level. Thus, our main empirical results are not likely to be driven by a few extreme events.

5. Possible Explanations

5.1 Equity Risk

The Halloween effect is mainly due to negative summer returns following high CAPE months, which will also contribute to low annual returns following high CAPE months. Although the risk-return trade off cannot account for the consistent negative summer returns, it is interesting to examine whether the winter months (May through October) are riskier than the summer months (November through April), following the high CAPE months. If equity risk is higher during the winter than during the summer, investors will demand higher returns during the winter. More specifically, since the higher volatility of the EW returns in the winter can expose investors to high chances of substantial losses, it can be considered higher equity risk. Thus, we examine the risk of six-month summer and winter excess returns following high CAPE months in terms of the cumulative frequency distribution.

Figure 3 illustrates the cumulative frequency distribution of the six-month EW excess returns following high CAPE months from November 1926 to October 2016. Winter (Summer) returns are defined as six-month excess returns from November to April (from May to October). The horizontal line indicates 10 classes with a 10% width, covering less than -40% through less than 50%. The vertical line indicates the corresponding cumulative frequency. The cumulative frequency is defined as the sum of the relative frequencies for all classes that are less than or equal to the given class.

Insert <Figure 3>

According to Figure 3, the cumulative frequency for the six-month summer returns is much higher than that for the six-month winter returns for every class. The summer returns appear to be riskier than the winter returns because investors have higher chances of losing money at any class of negative returns. For example, the cumulative frequency that the six-month return is less than zero is 63.3% for all summer returns following high CAPE months, while it is 22.4% for all winter returns. Similar results are obtained for the VW returns. The cumulative frequency for negative VW summer returns is 44.9%, while that for negative VW winter returns is 20.4%. Following high CAPE months, the summer excess returns appear to be riskier than the winter excess returns.[7]

5.2 Other Explanations

We propose the mispricing hypothesis to explain why the high CAPE ratio can forecast low returns of EW and VW returns in the short run. Overpricing is more likely to occur for stocks with strong short sale constraints during the period of high CAPE ratios because short sale constraints impede the short sale of rational investors. However, such overpricing tends to get corrected following high CAPE months so that the CAPE ratio has a predictive power for the returns. In addition, the mispricing hypothesis suggests that CRSP EW returns are more likely to be influenced by returns on stocks with heavy short sale constraints than CRSP VW returns, Consistent with the mispricing hypothesis, our results show that the CAPE ratio tends to have stronger predictability of CRSP EW returns than CRSP VW returns. However, this hypothesis does not account for the seasonal pattern of winter and summer returns following high CAPE months, which begs us to ask the following two questions: Why are summer excess returns consistently lower than winter excess returns following high CAPE months? Why are summer excess returns consistently negative following high CAPE months?

Several researchers attempt to explain poor summer returns. Bouman and Jacobsen (2002) report that low summer returns are related to the seasonal pattern of vacations. The summer

[7] We also examine the cumulative frequency of the annual excess returns following low and high CAPE months. The cumulative frequency for annual returns is higher following high CAPE months for every class than low CAPE months. The annual returns following high CAPE months appear to be riskier than those following low CAPE months. For an example, the cumulative frequency of having a less than zero return is 44.2% for annual EW returns following high CAPE months, while it is only 17.5% for those following low CAPE months.

period may be characterized as decreases in the number of investors because many investors sell their stocks off for liquidity purpose and go for summer vacations. Such a change in vacations during summer may increase the degree of risk aversion but decrease the risk-bearing capacity in the economy. When the remaining investors ask for a higher market risk premium, the stock prices will fall in summer. However, the stock prices will decrease in the summer only when the market continuously fails to predict the seasonal pattern of vacations and its economic effects. If investors are able to expect the seasonal pattern of the risk-bearing capacity caused by the seasonal change in vacations, their arbitrage activity forces the stock prices to fall before the summer. Summer returns will then be higher, due to the high risk compensation in the summer. Furthermore, the seasonal pattern of vacations fails to explain why summer returns are much lower following high CAPE months than following low CAPE months.

Kamstra et al. (2003) and Garrett et al. (2005) argue that the seasonal affective disorder (SAD) generates seasonal varying risk premiums. SAD is a well-known disorder in the medical and psychology literature. It causes depression in people during the fall and winter months, because of the shorter daylight hours per day. It is claimed that the depression lowers one's willingness to take risks, resulting in seasonal changes in risk premiums. Kamstra et al. (2003) presents evidence that seasonal variations in stock returns are related to the reduced length of the day in the fall and winter. Higher risk premiums are required during the winter when people become more pessimistic and the degree of their risk aversion increases. On the other hand, summer requires lower risk premiums. As a result, winter returns are higher than summer returns. Such influence of SAD on market returns is termed as the SAD effect.

However, the SAD effect does not seem to be reconciled with our major results. Since the SAD effect only depends on season and latitude, it should be applied to winter and summer returns following low CAPE months as well as following high CAPE months. However, the Halloween effect does not exist following the low CAPE months for the entire sample period of 1926-2016 because of the strong negative Halloween effect in the first sub-period. As shown in Panel A of Table 5, the summer returns are higher than the winter returns following the low CAPE months in both cases of the EW and VW excess returns. When the January effect is excluded, the Halloween effect is –2.72% (t=–2.02) of the EW excess returns following low CAPE months in the first sub-period.

When our sample period is restricted to the second sub-period 1971–2016, the Halloween effect appears even following low CAPE months. However, the most problematic issue is not the high winter returns that the SAD effect attempts to account for. It is the low summer returns following the high CAPE months. Both the EW and VW summer excess returns following high CAPE months are negative during the two sub-periods. As shown in Table 5, the summer EW excess returns are –1.03% (t=–1.88) for the first sub-period and –0.82% (t=–1.87) for the second sub-period. Consequently, the SAD effect does not help us understand why these summer excess returns are negative.

Doeswijk (2008) provides the optimism cycle hypothesis for the Halloween indicator. The optimism cycle hypothesis suggests that investors are basically optimistic. Their optimistic mindset about the prospect of the economy begins in the last quarter of the year, when they start looking forward to the next calendar year. Such an optimistic outlook is initially strong, because of their self-attribution bias.[8] However, it gradually turns into a pessimistic view over

[8] Daniel et al. (1998) introduced biased self-attribution to the financial literature from Bem (1965)'s attribution theory in psychology. Biased self-attribution is a bias of formulating investors' confidence about their belief. When investors receive information in favor of their belief, they attribute this information to reinforcement of their confidence. In contrast, when they receive information against their confidence, they attribute this information to a chance. As a result, biased self-attribution tends to initially strengthen investors' belief and delay their

several months. When the optimistic outlook is dominant among investors, stock returns will be high during the winter. However, the summer returns will poor, when investors gradually realize that their optimistic outlook does not fit in with the reality.

The optimism cycle hypothesis alone cannot justify why the Halloween effect only strongly appears following high CAPE months. However, combining the mispricing and optimism cycle hypotheses may explain the presence of the Halloween effect following high CAPE months. The mispricing hypothesis suggests that stocks with heavy short sale constraints are more likely to be overvalued during high CAPE months. Because of the optimistic outlook about the stock market and the biased-self attribution, investors continue to be optimistic in the winter, resulting in high stock returns. However, as a series of relatively pessimistic information against the optimistic outlook continues to arrive in the market over time, investors' outlook eventually turns bearish, depressing the stock prices in the summer. This hypothesis explains why the Halloween effect strongly appears and summer returns are extremely low following high CAPE months. We refer the hypothesis to as the mispricing-optimism cycle hypothesis.

In contrast, low CAPE months indicate that stocks may be undervalued, or at least are not likely to be overvalued. Since investors become optimistic around the time of a new calendar year, stock returns tend to be high at the beginning of new calendar year. However, if investors are too pessimistic about the economy during the period of low CAPE ratios, their initial optimistic prospect following low CAPE months may not be as strong as that following high CAPE months. If a series of new information about the stock market is better than the initial outlook, the summer returns could be higher than the winter returns, like during the first sub-period.

---

recognition of the reality against their belief.

However, if not, the summer returns could be lower than the winter returns, like during the second sub-period. Thus, this hypothesis can account for why the Halloween effect is not clear following the low CAPE months.

<Table 6> reports average monthly returns from November to October. Total in <Table 6> indicates average monthly returns for the entire sample periods of July 1926–December 2016 without classification of CAPE ratios. Except September and October, average EW monthly returns from May to August do not appear to be substantially lower than those from November to April. In fact, July return is the second highest EW returns next to January returns. The Halloween effect is not prominent as well in case of VW returns. Unlike the EW returns, average VW monthly return is highest in December, while July return is the second highest like the EW returns. Our result confirms that the Halloween effect is not prominent in the US for the entire period including the period 1926–1940, when January returns are excluded.

Insert <Table 6>

However, the Halloween effect strongly appears following high CAPE months. All summer EW monthly returns following high CAPE months are negative except July when the average return is 0.05% (t=0.07). September return is reliably negative which is –2.19% (t=–2.57). Although the Halloween indicator originally suggests that monthly returns begin to be low in May, our results suggest that it begins to be negative in April. When the CAPE ratio is high in September, the monthly returns tend to be positive for five months from November to March of the next year. CRSP VW monthly returns also show the similar pattern as the EW monthly returns. On the other hand, when the CAPE ratio is low in September, the monthly return tends to be low in November and December, indicating that optimistic sentiment does not begin in the last quarter. In addition, the summer returns tend to be higher than the winter returns following low CAPE months as shown in <Table 4>.

6. Conclusions

This paper contributes to the financial literature by adding new findings about stock return predictability. First of all, an examination of the Shiller cyclically adjusted pricing-earnings (CAPE) ratio reveals its forecasting power for 12-month CRSP equally weighted (EW) and value weighted (VW) excess returns. If the CAPE ratio in November is higher than the median CAPE ratio over 36 months, November is defined as a high CAPE month; otherwise, it is a low CAPE month. the 12-month EW excess returns from January of the next year to December following the low CAPE months are, on average, 20.7% higher than those following the high CAPE months for the period of 1927–2016. The sub-period analysis also confirms the reliable predictability for the EW returns at least at the 5% significance level. On the other hand, the January barometer does not appear to have reliable predictability for CRSP EW and VW returns for the entire sample of 1927–2016, compared to the CAPE ratio.

Second, we also report new empirical results related to the Halloween effect. When the CAPE ratio is classified every September as either high or low, the Halloween effect following high CAPE months is 2.54% (t=5.63) for EW excess returns and 1.43% (t=3.85) for VW excess returns for the sample period of July 1926–December 2016. The strong Halloween effect is due to extremely low returns in the summer (May through October). The average of the EW monthly excess return is –0.91% (t=–2.65) in the summer. The sub-period analysis confirms that the Halloween effect following the high CAPE months is statistically significant for the EW excess returns at the 1% level and for the VW excess returns at the 5% level.

The dependence of the Halloween effect on the previous CAPE ratio and the negative summer excess returns following the high CAPE months cast doubt on the validity of the existing hypotheses for the Halloween effect. The seasonal change of vacation and the SAD effect do not explain why the Halloween effect only consistently appears following high CAPE months and why summer excess returns are only negative following the high CAPE months. In addition, six-month summer returns appear to be riskier than six-month winter returns; the examination of 49 years' holding periods following high CAPE months shows that a chance of losing money in terms of the CRSP EW return is 63.3% in the summer, while it is 22.4% in the summer.

We suggest the mispricing hypothesis to explain the predictability of the CAPE ratio for one-year CRSP EW and VW returns. If a high CAPE ratio tends to indicate market overvaluation, market-wide overvaluation will magnify overpricing of stocks with heavy short sale constraints because short sale constraints impede negative opinions of rational investors from affecting the prices of these stocks. The relative overpricing of these stocks will get corrected following high CAPE months. On the other hand, S&P 500 stocks do not have severe short sale constraints. Since their prices reflect negative opinions even during the period of high CAPE ratios, the CAPE ratio will not have an ability to predict S&P 500 index in the short run. In addition, the predictability of the CAPE ratio will be stronger for CRSP EW returns than for CRSP VW returns because the former will be more influenced by the returns on stocks with heavy short sale constraints, which is consistent with our results as well.

The combination of the mispricing hypothesis and optimism cycle hypothesis can also account for our empirical results related to the Halloween effect. The mispricing hypothesis suggests that the high CAPE ratio will predict the poor performance of overvalued stocks. However, the optimism cycle hypothesis implies that it depends on the seasonality when the overpricing of these stocks will get corrected. The optimism mindset of investors about the prospect of the economy begins with a new calendar year. Such seasonal optimism and the biased self–attribution of investors can delay correcting the prices of overvalued stocks. As a result, following high CAPE months, overvalued stocks tend to have poor returns in summer. However, since stocks are less likely to get overvalued during the period of low CAPE months, poor summer returns will be found only following high CAPE months.

Our empirical results contain the following practical investment strategy. If the CAPE ratio in September is lower than the 36-month median CAPE ratio, invest in stock markets from November to October of the following year; otherwise, invest for five months from November to March and then sell in April and go away.

References

Andrade, S., V. Chhaochharia, and M. E. Fuerst, 2013, "Sell in May and Go Away" Just Won't Go Away, Financial Analyst Journal 69(4), 94-105.

Barber, B. and T. Odean, 2013, The Behavior of Individual Investors, in G. Constantinides, M. Harris, and R. M. Stulz (Eds), Handbook of the Economics of Finance 2B, Amsterdam, The Netherlands: North-Holland.

Bem, D. J., 1965, An Experimental Analysis of Self-Persuasion, Journal of Experimental Social Psychology 1, 199-218.

Boehme, Rodney D., Bartley R. D., and Sorin M. Sorescu, 2006, Short-Sale Constraints, Differences of Opinion, and Overvaluation, Journal of Financial and Quantitative Analysis 41 (2), 455-487.

Bohl, M. T. and C. A. Salm, 2010, The Other January Effect: International Evidence, The European Journal of Finance 16(2), 173-182.

Bouman, S., and B. Jacobsen, 2002, The Halloween Indicator, "Sell in May and Go Away": Another Puzzle, American Economic Review 92(5), 1618-1635.

Campbell, J. Y. and A. Kyle, 1993, Smart Money, Noise Trading, and Stock Price Behavior, Review of Economic Studies 60, 1-34.

Campbell, J. Y. and R. Shiller, 1988, Stock Prices, Earnings, and Expected Dividends, Journal of Finance 43(3), 661-676.

Campbell, J. Y. and R. Shiller, 1998, Valuation Ratios and the Long-Run Stock Market Outlook, Journal of Portfolio Management 24(2), 11-26.

Cohen, L., K. B. Diether, and C. J. Malloy, 2007, Supply and Demand Shifts in the Shorting Market, Journal of Finance 62, 2061-2096.

Cooper, M. J., J. J. McConnell, and A. V. Ovtchinnikov, 2006, The Other January Effect, Journal of Financial Economics 82, 315-341.

Daniel, K., D. Hirshleifer and A. Subrahmanyam, 1998, Investor Psychology and Security Market Under- and Overreaction, Journal of Finance 53(6), 1839–1885.

DeLong, J. B., A. Shleifer, and L. Summers, and R. Waldmann, 1990, Noise Trader Risk in Financial Markets, Journal of Political Economy 98, 703-738.

Doeswijk R. Q., 2008, The Optimism Cycle: Sell in May, De Economist 156(2), 175-200.

Dzhabarov C. and W. T. Ziemba, 2010, Do Seasonal Anomalies Still Work?, Journal of Portfolio Management 36(3), 93-104.

Dzhabarov C. and W. T. Ziemba, 2011, Seasonal Anomalies, retrieved from https://ssrn.com/abstract=1760610.

Garrett, I., M. Kamstra, and L. A. Kramer, 2005, Winter Blues and Time Variation in the Price of Risk, Journal of Empirical Finance 12(2), 291-316.

Graham, B. and D. Dodd, 1934, Security Analysis, first edition. New York: McGraw-Hill.

Haggard, K. S., and H. Witte, 2010, The Halloween Effect: Trick or Treat?, International Review of Financial Analysis 19, 379-387.

Hirsch J. and Y. Hirsch, Stock Trader's Almanac, 2008, Hoboken, New Jersey: Joan Wiley & Sons.

Hulbert, M., 2017 (January 8), Sorry, the 'January Barometer' is a Market Myth, The Wall Street Journal, retrieved from https://www.wsj.com/articles/sorry-the-january-barometer-is-a-market-myth-1483931640.

Jacobsen, B. and C. Y. Zhang, 2014, The Halloween Indicator, 'Sell in May and Go Away': An Even Bigger Puzzle, Working Paper, downloaded from SSRN (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2154873)

Jacobsen, B., and N. Visaltanachoti, 2009, The Halloween Effect in U.S. Sectors, Financial Reviews 44(3), 437-459.

Jones, C. and O. Lamont, 2002, Short Sale Constraints and Stock Returns, Journal of Financial Economics 66, 207-239.

Kamstra, M., L. A. Kramer, and M. D. Levi, 2003, Winter Blues: Seasonal Affected Disorder and Stock Market Returns, American Economic Review, 93(1), 324-343.

Kothari, S. P. and J. Shanken, 1997, Book-to-Market, Dividend Yield, and Expected Market Returns: A Time-Series Analysis, Journal of Financial Economics 44, 169-203.

Lamont, O. A., 2012, Go Down Fighting: Short Sellers vs. Firms, Review of Asset Pricing Studies (2), 1-30.

Maberly, E. and R. Pierce, 2004, Stock Market Efficiency Withstands Another Challenge: Solving the "Sell in May, Buy after Halloween" Puzzle, Econ Journal Watch 1(1), 29-45.

Miller, E. M., 1997, Risk, Uncertainty and Divergence of Opinion, Journal of Finance 32, 1151-1168.

Nagel, S., 2005, Short Sales, Institutional Investors and the Cross-Section of Stock Returns, Journal of Financial Economics 78, 277-309.

Pontiff, J. and L. D. Schall, 1998, Book-to-Market Ratios as Predictors of Market Returns, Journal of Financial Economics 49, 141-160

Shleifer, Andrei, and Robert Vishny, 1997, The Limits of Arbitrage, Journal of Finance 52, 35-55.

Siegel, J. J., 2016, The Shiller CAPE Ratio: A New Look, Financial Analyst Journal 72(3), 41-50.

Stambaugh, R. F., J. Yu, and Y. Yuan, 2015, Arbitrage Asymmetry and Idiosyncratic Volatility Puzzle, Journal of Finance 70, 1903-1948.

<Table 1> Predictability of the CAPE Ratio on 12-Month Equally Weighted (EW) and Value Weighted (VW) Excess Returns: July 1926 to December 2016

The CAPE (Cyclically Adjusted Price-Earnings) ratio is annually sorted as either a high or low CAPE ratio, based on the 36-month median of the CAPE ratio. If the CAPE ratio in month *t* is higher than the median CAPE ratio from *t*-35 to *t*, month *t* is defined as a high CAPE month; otherwise, month *t* is defined as a low CAPE month. The monthly excess return is defined as the CRSP monthly return minus the one-month Treasury bill rate. The 12-month excess returns are measured one month after a high CAPE month or low CAPE month. For example, Low (High) CAPE indicates 12-month annual excess returns from January through December following the low (high) CAPE ratio in November of the previous year. (Low–High) indicates a difference in monthly excess returns between holding periods following low and high CAPE months. Beginning month in the first column refers to the month when the holding period begins. When 12-month holding period begins with January, the sample period has 50 high CAPE ratios and 40 low CAPE ratios in November. The parentheses indicate a heteroscedasticity consistent *t*-value.

Beginning Month	12-Month EW Excess Returns (%)			12-Month VW Excess Returns (%)		
	Low CAPE	High CAPE	Low–High	Low CAPE	High CAPE	Low-High
January (50 High, 40 Low)	23.21 (4.91)	2.52 (0.76)	20.69 (3.59)	11.94 (3.48)	4.31 (1.82)	7.64 (1.83)
February (55 High, 35 Low)	22.20 (4.01)	4.64 (1.52)	17.56 (2.77)	11.06 (2.80)	5.33 (2.35)	5.73 (1.26)
March (57 High, 33 Low)	21.39 (3.19)	5.66 (1.96)	15.73 (2.16)	9.67 (2.06)	6.27 (2.95)	3.40 (0.66)
April (58 High, 32 Low)	22.48 (3.12)	5.51 (1.87)	16.98 (2.18)	10.12 (1.99)	6.20 (2.67)	3.92 (0.70)
May (58 High, 32 Low)	23.09 (3.56)	5.24 (1.83)	17.85 (1.83)	9.65 (2.08)	6.48 (2.91)	3.17 (0.62)
June (57 High, 33 Low)	22.03 (2.93)	5.44 (1.84)	16.59 (2.05)	8.91 (1.70)	6.79 (2.94)	2.13 (0.37)
July (54 High, 36 Low)	17.75 (2.41)	7.54 (2.75)	10.21 (1.30)	7.34 (1.41)	7.93 (3.74)	-0.59 (-0.10)
August (55 High, 35 Low)	16.85 (2.83)	8.38 (3.33)	8.47 (1.31)	6.36 (1.54)	8.56 (4.15)	-2.20 (-0.48)
September (51 High, 39 Low)	17.48 (4.15)	7.20 (2.63)	10.28 (2.05)	8.26 (2.72)	7.24 (3.05)	1.01 (0.26)
October (52 High, 38 Low)	16.73 (3.47)	8.00 (2.88)	8.74 (1.57)	7.76 (21.65)	7.62 (3.20)	0.14 (0.03)
November (49 High, 41 Low)	20.66 (5.35)	4.16 (1.36)	16.51 (3.35)	11.39 (4.22)	4.59 (1.86)	6.80 (1.86)
December (51 High, 39 Low)	21.61 (5.31)	4.15 (1.22)	17.46 (3.28)	11.05 (3.75)	5.16 (1.99)	5.89 (1.50)

<Table 2> Predictability of the CAPE Ratio and the January Barometer: January 1927–December 2016

When the CRSP monthly excess returns are measured as dependent variables from February to December, the following regression is tested: $r_t = \alpha + \beta_1 LOW_{t-3} + \beta_2 JBR_{t-1} + \varepsilon_t$. The excess return is defined as the CRSP monthly return minus the one-month Treasury bill rate. LOW is a dummy variable equal to one, when the CAPE ratio in November of the previous year is not higher than the median CAPE ratio over 36 months; otherwise, LOW is zero. JBR is a dummy variable that represents the January barometer. It is one when there is a positive January excess return and zero when there is a non-positive January excess return. When the monthly excess returns are measured beginning with January, the following regression is used: $r_t = \alpha + \beta_1 LOW_{t-2} + \varepsilon_t$ The parentheses indicate a heteroscedasticity consistent *t*-value.

	CRSP EW Excess Returns			CRSP VW Excess Returns		
	Intercept	LOW	JBR	Intercept	LOW	JBR
January 1927-December 2016						
Jan.	3.85 (4.29)	2.68 (1.77)		0.98 (1.56)	0.39 (0.38)	
Feb – Dec.	0.08 (0.18)		0.69 (1.33)	0.32 (1.08)		0.45 (1.24)
	-0.12 (-0.53)	1.64 (3.41)		0.30 (1.59)	0.66 (1.82)	
	-0.58 (-1.26)	1.62 (3.37)	0.62 (1.19)	0.07 (0.24)	0.63 (1.72)	0.40 (1.10]
Jan.– Dec.	0.21 (0.93)	1.72 (3.71)		0.36 (1.97)	0.63 (1.86)	
February 1940-December 2016						
Feb – Dec.	-0.26 (-0.60)		1.11 (2.38)	0.01 (0.03)		1.00 (3.18)
	-0.08 (-0.39)	1.43 (4.00)		0.33 (1.78)	0.62 (2.11)	
	-0.81 (-1.90)	1.37 (3.80]	1.01 (2.16)	-0.21 (-0.78)	0.54 (1.81)	0.95 (2.96)

<Table 3> Sub-period Analysis for the Predictability of the CAPE Ratio and January Barometer

If the CAPE ratio in November is higher than the median CAPE ratio over 36 months, November is defined as a high CAPE month; otherwise, it is a low CAPE month. The holding period is January to December following high or low CAPE month. Panel A reports the monthly excess returns following high and low CAPE months. (Low–High) indicates a difference in monthly excess returns between holding periods following low and high CAPE months. Panel B reports monthly excess returns from February to December, based on the positive and negative January excess returns. (Positive–Negative) is defined as a difference in monthly excess returns between holding periods following positive and negative January excess returns. The parentheses indicate a heteroscedasticity consistent *t*-value.

Panel A: Predictability of the CAPE ratio

Months	CRSP EW Excess Returns			CRSP VW Excess Returns		
	January 1927 – December 1971					
	Low CAPE (21 Years)	High CAPE (24 Years)	Low–High	Low CAPE (21 Years)	High CAPE (26 Years)	Low–High
Jan.	5.95 (4.18)	4.38 (2.95)	1.57 (0.76)	1.32 (1.36)	1.33 (1.53)	-0.02 (-0.01)
Feb.– Dec.	1.85 (2.60)	-0.05 (-0.13)	1.90 (2.37)	1.23 (2.48)	0.25 (0.84)	0.98 (1.69)
Jan. – Dec.	2.19 (3.28)	0.32 (0.88)	1.87 (2.46)	1.24 (2.68)	0.34 (1.21)	0.89 (1.65)
	January 1972 – December 2016					
Months	Low CAPE (19 Years)	High CAPE (26 Years)	Low–High	Low CAPE (19 Years)	High CAPE (26 Years)	Low–High
Jan.	7.17 (3.57)	3.37 (3.24)	3.81 (1.68)	1.43 (1.11)	0.66 (0.73)	0.77 (0.49)
Feb.– Dec.	1.14 (2.78)	-0.19 (-0.68)	1.33 (2.69)	0.67 (1.93)	0.35 (1.45)	0.32 (0.75)
Jan. – Dec.	1.65 (3.85)	0.11 (0.40)	1.54 (3.03)	0.73 (2.18)	0.37 (1.61)	0.35 (0.87]

Panel B: Predictability of the January Barometer

	CRSP EW Excess Returns			CRSP VW Excess Returns		
CRSP EW	Positive Jan. (29 Years)	Negative Jan. (16 Years)	Positive–Negative	Positive Jan. (26 Years)	Negative Jan. (19 Years)	Positive–Negative
February 1927 – December 1971						
Feb.– Dec.	1.03 (2.28)	0.17 (0.23)	0.86 (0.99)	0.73 (2.15)	0.67 (1.34)	0.05 (0.08)
February 1972 – December 2016						
Feb.– Dec.	0.51 (1.99)	0.01 (0.02)	0.50 (0.82)	0.82 (3.40)	0.02 (0.06)	0.80 (1.91)

<Table 4> The Halloween Indicator and the Predictability of the CAPE Ratio: July 1926–December 2016

If the CAPE ratio in September is higher than the median CAPE ratio over 36 months, September is defined as a high CAPE month; otherwise, September is defined as a low CAPE month. The holding period is one year, from November to October, following either high or low CAPE months. The sample period covers 90 years and 6 months (1,086 months): 594 months (49 years and 6 months) following high CAPE months and 492 months (41 years) following low CAPE months. Total indicates the average of all monthly excess returns for the entire sample period without classification of CAPE ratios. Winter indicates average monthly excess returns in the winter (November through April). Winter(J) means average monthly excess returns, excluding January returns. Summer indicates average monthly excess returns in the summer (May through October). The excess returns are defined as the CRSP EW or VW monthly returns minus the one-month Treasury bill rate. (Winter–Summer) measures the Halloween effect by subtracting summer returns from winter returns. The parentheses indicate a heteroscedasticity consistent *t*-value.

Monthly Returns	CRSP EW Excess Returns				CRSP VW Excess Returns			
	Total	Low CAPE	High CAPE	Low – High	Total	Low CAPE	High CAPE	Low – High
Winter	1.64 (5.83)	1.66 (3.25)	1.63 (5.57)	0.04 (0.06)	0.97 (4.65)	0.80 (2.16)	1.11 (4.91)	-0.31 (-0.72)
January	5.04 (6.73)	5.62 (4.72)	4.56 (4.84)	1.07 (0.70)	1.16 (2.32)	1.09 (1.38)	1.21 (1.91)	-0.13 (-0.13)
Winter(J)	0.97 (3.30)	0.87 (1.59)	1.05 (3.71)	-0.17 (-0.28)	0.93 (4.06)	0.74 (1.79)	1.09 (4.55)	-0.35 (-0.73)
Summer	0.31 (0.92)	1.78 (2.96)	-0.91 (-2.65)	2.69 (3.89)	0.32 (1.28)	1.10 (2.65)	-0.32 (-1.08)	1.42 (2.79)
Winter – Summer	1.32 (3.00)	-0.12 (-0.15)	2.54 (5.63)	-2.65 (-2.92)	0.65 (2.00)	-0.30 (-0.54)	1.43 (3.85)	-1.73 (-2.59)
Winter(J) – Summer	0.66 (1.48)	-0.91 (-1.11)	1.95 (4.41)	-2.86 (-3.09)	0.61 (1.81)	-0.36 (-0.61)	1.41 (3.71)	-1.77 (-2.53)

<Table 5> Sub-period Analysis for the Halloween Indicator and Return Predictability

If the CAPE ratio in September is higher than the median CAPE ratio over 36 months, September is defined as a high CAPE month; otherwise, September is defined as a low CAPE month. The holding period is one year, from November to October, following either high or low CAPE months. The first sub-period, November 1926–October 1971, covers 264 months following the low CAPE months and 276 months following the high CAPE months. The second sub-period, November 1971–December 2016, covers 228 months following the low CAPE months and 314 months following the high CAPE months. Total indicates the average of all monthly excess returns for the entire sample periods without classification of CAPE ratios. Winter indicates average monthly excess returns in the winter (November through April). Winter(J) means average monthly excess returns, excluding January returns. Summer indicates average monthly excess returns in the summer (May through October). The excess returns are defined as the CRSP EW or VW monthly returns minus the one-month Treasury bill rate. (Winter–Summer) measures the Halloween effect by subtracting summer returns from winter returns. The parentheses indicate a heteroscedasticity consistent *t*-value.

Panel A: Period of November 1926-October 1971

Monthly Returns	CRSP EW Excess Returns				CRSP VW Excess Returns			
	Total (540 M)	Low CAPE (264 M)	High CAPE (276 M)	Low − High	Total (540 M)	Low CAPE (264 M)	High CAPE (276 M)	Low − High
Winter	1.55 (3.43)	0.97 (1.24)	2.09 (4.57)	-1.12 (-1.24)	0.94 (2.81)	0.50 (0.86)	1.36 (4.07)	-0.86 (-1.28)
January	5.10 (6.73)	5.51 (3.89)	4.72 (3.12)	0.79 (0.38)	1.33 (2.05)	1.06 (1.12)	1.58 (1.79)	-0.53 (-0.41)
Winter(J)	0.83 (1.72)	0.07 (0.07)	1.57 (3.54)	-1.51 (-1.54)	0.86 (2.27)	0.39 (0.57)	1.31 (3.65)	-0.92 (-1.21)
Summer	0.83 (1.42)	2.78 (2.71)	-1.03 (-1.88)	3.81 (3.28)	0.57 (1.40)	1.64 (2.46)	-0.46 (-0.99)	2.10 (2.59)
Winter – Summer	0.71 (0.97)	-1.81 (-1.40)	3.12 (4.37)	-4.93 (-3.35)	0.37 (0.69)	-1.15 (-1.29)	1.81 (3.18)	-2.96 (-2.81)
Winter(J) – Summer	-0.00 (-0.00)	-2.72 (-2.02)	2.60 (3.69)	-5.31 (-3.50)	0.29 (0.52)	-1.26 (-1.33)	1.77 (3.02)	-3.02 (-2.72)

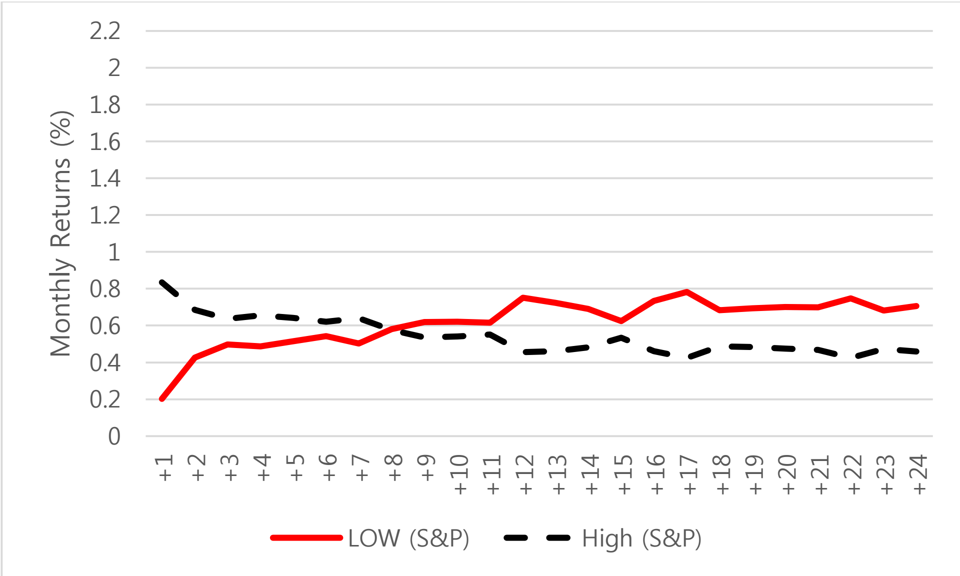
Panel B: Period of November 1971-December 2016

Monthly Returns	CRSP EW Excess Returns				CRSP VW Excess Returns			
	Total (542 M)	Low CAPE (228 M)	High CAPE (314 M)	Low − High	Total (542 M)	Low CAPE (228 M)	High CAPE (314 M)	Low − High
Winter	1.74 (5.12)	2.47 (3.95)	1.22 (3.30)	1.25 (1.72)	1.01 (3.96)	1.15 (2.68)	0.90 (2.93)	0.25 (0.47)
January	4.97 (4.62)	5.75 (2.90)	4.41 (3.80)	1.34 (0.58)	0.99 (1.30)	1.12 (0.87)	0.89 (0.98)	0.24 (0.15)
Winter(J)	1.10 (3.32)	1.81 (2.95)	0.59 (1.67)	1.22 (1.72)	1.01 (3.82)	1.16 (2.60)	0.91 (2.81)	0.25 (0.46)
Summer	-0.21 (-0.63)	0.62 (1.24)	-0.82 (-1.87)	1.44 (2.16)	0.07 (0.22)	0.47 (1.06)	-0.23 (-0.58)	0.69 (1.18)
Winter – Summer	1.95 (4.10)	1.84 (2.30)	2.04 (3.55)	-0.20 (-0.20)	0.94 (2.43)	0.68 (1.11)	1.13 (2.27)	-0.44 (-0.56)
Winter(J) – Summer	1.31 (2.79)	1.19 (1.50)	1.41 (2.50)	-0.22 (-0.23)	0.94 (2.39)	0.69 (1.10)	1.13 (2.23)	-0.44 (-0.55)

<Table 6> Average Monthly Excess Returns Following Low and High CAPE Months: July 1926–December 2016

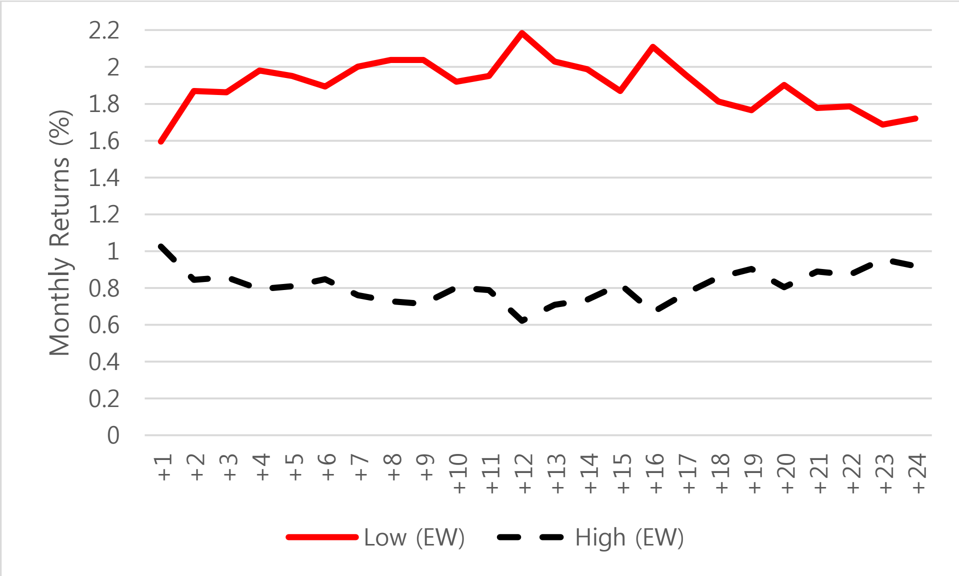
If the CAPE ratio in September is higher than the median CAPE ratio over 36 months, September is defined as a high CAPE month; otherwise, September is defined as a low CAPE month. The holding period is one year, from November to October, following either high or low CAPE months. Total indicates the average monthly return from November to October for the entire sample periods of July 1926–December 2016 without classification of CAPE ratios. Low (High) indicates the average monthly return from November to October following low (high) CAPE months. The holding period covers 492 months following low CAPE months and 594 months following high CAPE months.

Month	CRSP EW Returns (%)			CRSP VW Returns (%)		
	Total	Low	High	Total	Low	High
November	1.00 (1.51)	0.16 (0.15)	1.69 (2.02)	1.22 (2.28)	0.57 (0.70)	1.75 (2.47)
December	0.88 (1.72)	-0.12 (-0.13)	1.71 (3.48)	1.47 (3.90)	0.70 (1.07)	2.11 (5.12)
January	5.04 (6.70)	5.62 (4.66)	4.56 (4.79)	1.16 (2.31)	1.09 (1.37)	1.21 (1.89)
February	1.12 (2.08)	1.03 (1.09)	1.19 (1.96)	0.39 (0.92)	0.27 (0.37)	0.50 (0.98)
March	0.68 (1.06)	0.65 (0.50)	0.70 (1.43)	0.53 (1.03)	0.19 (0.19)	0.81 (1.76)
April	1.16 (1.33)	2.66 (1.52)	-0.09 (-0.13)	1.05 (1.56)	1.98 (1.48)	0.27 (0.51)
May	0.55 (0.57)	1.98 (1.09)	-0.65 (-0.76)	0.21 (0.36)	0.61 (0.60)	-0.12 (-0.17)
June	0.46 (0.66)	2.05 (1.72)	-0.87 (-1.14)	0.56 (1.03)	1.33 (1.48)	-0.09 (-0.13)
July	1.33 (1.75)	2.90 (2.06)	0.05 (0.07)	1.25 (2.06)	1.87 (1.64)	0.74 (1.25)
August	1.04 (1.14)	2.71 (1.55)	-0.33 (-0.43)	0.88 (1.39)	2.09 (1.96)	-0.11 (-0.16)
September	-0.79 (-0.96)	0.93 (0.63)	-2.19 (-2.57)	-1.03 (-1.67)	-0.09 (-0.09)	-1.80 (-2.57)
October	-0.74 (-0.99)	0.11 (0.10)	-1.45 (-1.40)	0.05 (0.65)	0.78 (0.85)	-0.54 (-0.58)



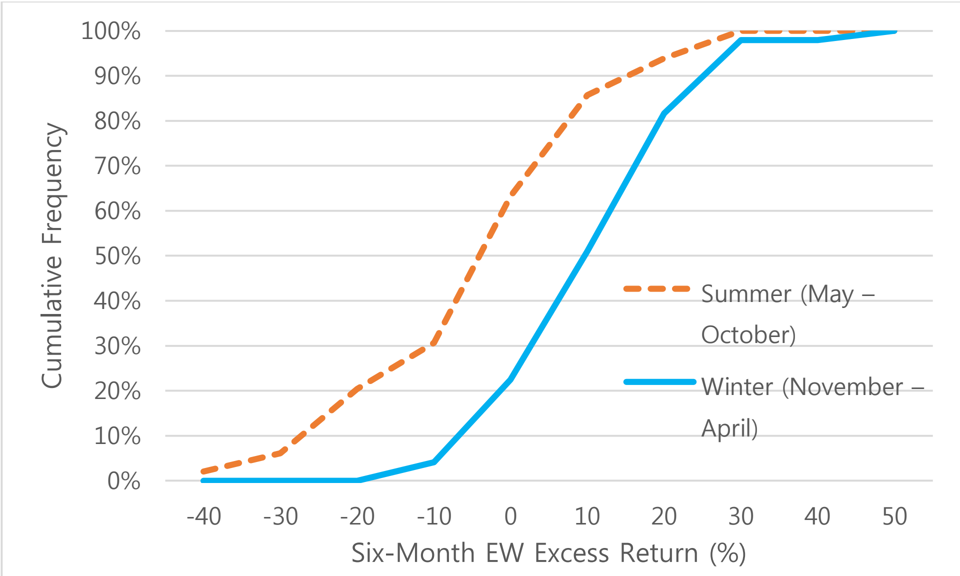
<Figure 1> S&P 500 Monthly Returns Over 24 Months Following Low and High CAPE Months: January 1926–December 2016

The CAPE ratio is monthly sorted as either a high or low CAPE ratio, based on the moving median CAPE ratio over 36 months. If the CAPE ratio in month *t* is higher (lower) than the median of the CAPE ratio from *t*-35 to *t*, month *t* is defined as a high (low) CAPE month. This figure reports S&P 500 monthly index returns over 24 months following high and low CAPE months. There are 637 high CAPE months and 432 low CAPE months from December 1925 to December 2014.



<Figure 2> CRSP EW Monthly Returns Over 24 Months Following Low and High CAPE Months: January 1926–December 2016

The CAPE ratio is monthly sorted as either a high or low CAPE ratio, based on the moving median CAPE ratio over 36 months. If the CAPE ratio in month *t* is higher (lower) than the median of the CAPE ratio from *t*-35 to *t*, month *t* is defined as a high (low) CAPE month. This figure reports CRSP EW monthly returns over 24 months following high and low months. There are 637 high CAPE months and 432 low CAPE months from December 1925 to December 2014.



<Figure 3> Cumulative Frequency Distribution of Six-Month EW Excess Returns Following High CAPE Months: November 1926–October 2016

If the CAPE ratio in September is higher than the median CAPE ratio over 36 months, September is defined as a high CAPE month. The holding period is one year from November to October following high CAPE months. The sample period includes 49 years of holding periods following high CAPE months. Winter returns are six-month EW excess returns as a percentage from November to April, while summer returns are six-month EW excess returns from May to October. The horizontal line indicates 10 classes with a 10% width, covering less than –40% through less than 50%, while the vertical line illustrates the corresponding cumulative frequency.

Appendix 1. Six-Month Cumulative Excess Returns during the Winter and Summer Following High and Low CAPE Months: November 1926–October 2016

If the CAPE ratio in September is higher than the median CAPE ratio over 36 months, September is defined as a high CAPE month; otherwise, September is defined as a low CAPE month. The holding period is one year, from November to October, following either high or low CAPE months. W and S indicate six-month winter (November to April) returns and six-month summer (May to October) returns, respectively. For example, the CAPE ratio in September 1929 is sorted as a high CAPE month. Following the high CAPE month, the six-month EW cumulative excess return are 4.0% during the winter (November 1929 to April 1930) and –46.3% during the summer (May 1930 to October 1930). For the period of 1926–2016, there are 49 years from September to October following high CAPE months and 41 years from September to October following low CAPE months. Panel A (B) shows six-month CRSP EW (VW) excess returns.

Panel A: Six-Month CRSP EW Excess Returns

EW Returns (%) Following High CAPE Months						EW Returns (%) Following Low CAPE Months					
Year	W	S	Year	W	S	Year	W	S	Year	W	S
1926	9.0	8.3	1980	7.8	-17.3	1930	-8.2	-25.8	1975	30.2	-3.9
1927	28.1	7.7	1983	-9.1	-7.6	1931	-47.8	71.2	1977	22.4	1.8
1928	11.3	-25.2	1984	8.6	-3.2	1932	38.7	59.5	1978	19.7	-0.3
1929	4.0	-46.3	1985	21.4	-7.0	1934	8.5	36.4	1979	5.9	32.8
1933	41.6	-22.3	1986	14.3	-24.5	1937	-18.4	42.4	1981	-8.1	10.2
1935	28.4	25.0	1987	13.9	-2.3	1938	-20.8	41.7	1982	37.2	-1.5
1936	11.8	-35.1	1988	5.8	-2.0	1940	-9.0	11.7	1990	32.8	6.7
1939	-1.4	-5.7	1989	-9.4	-26.1	1941	-10.2	27.6	2001	18.6	-23.4
1943	8.0	18.8	1991	14.6	-1.9	1942	46.8	6.4	2002	16.4	39.4
1944	25.3	22.3	1992	16.9	15.7	1946	-4.7	10.1	2003	11.5	2.6
1945	25.8	-25.9	1993	-4.2	0.5	1947	5.0	2.6	2004	3.1	8.8
1950	20.4	3.3	1994	2.3	11.4	1948	-9.3	11.4	2006	7.9	-2.0
1951	-0.4	3.0	1995	14.9	-2.1	1949	17.6	10.3	2008	9.3	30.4
1952	7.2	-4.3	1996	-2.1	23.5	1953	13.1	15.2	2009	25.1	1.2
1954	25.5	2.2	1997	7.5	-26.2	1957	12.2	23.4	2011	11.9	0.8
1955	13.5	-3.8	1998	17.8	0.2	1958	17.7	-1.0	2015	2.1	4.7
1956	4.3	-16.3	1999	23.3	-7.4	1960	27.6	1.3			
1959	-5.7	-2.9	2000	0.1	-6.0	1962	19.5	4.4			
1961	-3.6	-20.0	2005	15.5	-2.0	1966	35.0	18.6			
1963	4.6	8.2	2007	-16.2	-37.7	1967	15.8	16.6			
1964	12.9	4.8	2010	16.7	-14.2	1969	-34.0	-0.8			
1965	19.9	-26.5	2012	13.7	14.4	1970	33.0	-16.4			
1968	-1.4	-11.5	2013	7.8	1.4	1971	21.1	-13.3			
1972	-19.3	-2.3	2014	4.7	-8.5	1973	-16.0	-27.4			
1976	17.5	3.2				1974	30.3	3.4			

Panel B: Six-Month CRSP VW Excess Returns

VW Returns (%) Following High CAPE Months						VW Returns (%) Following Low CAPE Months					
Year	W	S	Year	W	S	Year	W	S	Year	W	S
1926	10.0	12.8	1980	1.4	-13.4	1930	-10.2	-27.1	1975	14.5	0.6
1927	19.4	8.3	1983	-6.1	1.3	1931	-47.6	33.8	1977	7.6	-4.1
1928	17.2	-9.9	1984	7.2	3.6	1932	26.9	18.1	1978	10.2	-0.7
1929	1.6	-35.1	1985	20.6	1.3	1934	8.7	28.8	1979	2.9	19.7
1933	20.5	-12.1	1986	14.5	-14.4	1937	-15.2	33.1	1981	-8.4	13.2
1935	12.1	23.4	1987	6.5	3.8	1938	-12.3	20.8	1982	20.8	-3.3
1936	0.3	-24.2	1988	8.7	6.0	1940	-10.9	6.9	1990	22.1	5.0
1939	0.3	-4.5	1989	-6.1	-12.2	1941	-19.1	23.4	2001	4.5	-18.7
1943	3.2	10.9	1991	5.1	1.3	1942	25.4	5.1	2002	5.6	17.4
1944	18.0	14.8	1992	6.1	7.8	1946	-1.3	8.5	2003	6.1	3.3
1945	17.2	-20.7	1993	-4.2	3.1	1947	4.7	5.4	2004	2.7	6.1
1950	18.2	4.7	1994	6.2	11.2	1948	-6.6	11.5	2006	6.8	4.8
1951	2.3	4.4	1995	11.3	3.9	1949	15.2	9.5	2008	-4.5	19.5
1952	3.7	1.3	1996	7.8	15.4	1953	17.4	11.6	2009	16.9	2.6
1954	21.5	6.6	1997	16.3	-6.1	1957	7.7	18.8	2011	11.3	2.1
1955	16.2	-4.6	1998	19.1	0.6	1958	13.5	-0.2	2015	0.5	4.0
1956	4.4	-12.4	1999	8.1	-3.5	1960	22.4	5.1			
1959	-5.1	-0.9	2000	-15.5	-16.5	1962	21.8	5.3			
1961	-5.1	-14.0	2005	9.4	2.4	1966	18.2	1.8			
1963	6.4	6.4	2007	-11.0	-34.8	1967	4.8	6.2			
1964	5.8	3.3	2010	16.3	-8.4	1969	-21.3	0.7			
1965	0.3	-14.0	2012	13.5	11.2	1970	23.7	-9.9			
1968	-1.3	-7.5	2013	7.5	6.6	1971	14.6	0.9			
1972	-9.6	1.1	2014	4.5	-1.5	1973	-20.9	-20.7			
1976	-1.2	-4.9				1974	17.8	1.2			

Appendix 2. Halloween Effect Following High and Low CAPE Months, Excluding 5 years with the Worst Summer Returns: July 1926–December 2016

If the CAPE ratio in September is higher than the median of the CAPE ratio over 36 months, September is defined as a high CAPE month; otherwise, September is defined as a low CAPE month. The holding period is one year from November to October following either high or low CAPE months. This table excludes five years exhibiting the worst six-month EW returns in the summer following high CAPE months. These high CAPE months are the Septembers of 1929, 2007, 1936, 1965, and 1997. The sample period covers 85 years (1,080 months): 534 months (44 years and 6 months) following high CAPE months and 492 months (41 years) following low CAPE months. Total indicates the average of all monthly excess returns for the entire sample periods of July 1926–December 2016 without classification of CAPE ratios. Winter indicates monthly average excess returns in the winter (November through April). Winter-J means monthly average excess returns excluding January returns. Summer indicates monthly average excess returns in the summer. The excess returns are defined as the CRSP EW or VW monthly returns minus the one-month Treasury bill rate. The parentheses indicate a heteroscedasticity consistent *t*-value.

	CRSP EW Excess Returns			CRSP VW Excess Returns		
	Low CAPE	High CAPE	Total	Low CAPE	High CAPE	Total
Winter	1.66 (3.25)	1.64 (5.14)	1.65 (5.67)	0.80 (2.16)	1.21 (5.17)	1.01 (4.69)
January	5.62 (4.72)	4.41 (4.38)	4.99 (6.41)	1.09 (1.38)	1.13 (1.69)	1.11 (2.16)
Winter-J	0.87 (1.59)	1.09 (3.83)	0.99 (3.26)	0.74 (1.79)	1.22 (4.96)	0.99 (4.17)
Summer	1.78 (2.96)	-0.37 (-1.11)	0.66 (1.94)	1.10 (2.65)	0.12 (0.42)	0.59 (2.36)
Winter – Summer	-0.12 (-0.15)	2.01 (4.47)	0.99 (2.22)	-0.30 (-0.54)	1.08 (2.90)	0.42 (1.28)
Winter-J – Summer	-0.91 (-1.11)	1.46 (3.32)	0.33 (0.72)	-0.36 (-0.61)	1.10 (2.88)	0.40 (1.16)