International Review of Financial Analysis 19 (2010) 379–387

Contents lists available at ScienceDirect

International Review of Financial Analysis

The Halloween effect: Trick or treat?

K. Stephen Haggard *, H. Douglas Witte [1]

Missouri State University, 901 S. National Ave., Springfield, MO 65897 United States

ARTICLE INFO

Article history:
Received 13 November 2009
Received in revised form 18 August 2010
Accepted 15 October 2010
Available online 23 October 2010

JEL classification:
G10
G14

Keywords:
Anomalies
Market efficiency
Calendar

ABSTRACT

Research documents higher stock returns in November through April than for the rest of the year. This anomaly is known as the "Halloween effect" and results in the following trading rule: sell stocks in early May, invest in T-bills, and re-invest in stocks on Halloween. In contrast to recent studies, we show that the Halloween effect is robust to consideration of outliers and the "January effect." Additionally, we show that investing in a "Halloween portfolio" provides risk-adjusted returns in excess of buy and hold equity returns even after consideration of transaction costs.

1. Introduction

The "Halloween effect," identified by Bouman and Jacobsen (2002), is an equity return anomaly in which the months of November through April provide higher returns than the remaining months of the year. This effect, if real, is perhaps of greater interest to investors than most other anomalies because the trading rule is simple to implement with low transactions costs, making exploitation of this anomaly potentially profitable. More recent studies posit that this anomaly might be driven by outliers or is simply the "January effect" in disguise. In this study, we examine the robustness of the Halloween effect to the consideration of outliers and the January effect. We also construct mean-variance efficient portfolios to determine whether investing in a Halloween portfolio can result in risk-adjusted returns superior to those of a buy-and-hold market portfolio. Finally, we examine the impact of transaction costs on the returns to investing in a Halloween portfolio.

2. Literature review

In their seminal paper, Bouman and Jacobsen (2002) analyze stock returns across 37 countries from January 1970 through August 1998 and find a Halloween effect in 36 of these markets. This finding is remarkable in light of the adage "sell in May and go away" having appeared numerous times in the financial press before and during their sample period. Most return anomalies disappear after discovery, presumably as opportunistic traders exploit them. The effect is particularly strong in European countries and is not the result of risk differences between the May–October and November–April timeframes that delineate the Halloween effect. Bouman and Jacobsen also demonstrate the economic significance of Halloween-based investment, even when transaction costs are considered.

Bouman and Jacobsen's results for U.S. stock returns are more marginal. When the January effect is not considered, the Halloween effect attains statistical significance at the 5% level. After incorporating the January effect, significance falls just short of the 10% level they employ as a cutoff. However, the November–April period has a slightly smaller return standard deviation than the May–October period, adding to its attractiveness. Jacobsen and Visaltanachoti (2009) examine differences in the Halloween effect among U.S. stock market sectors and show that the effect is strongest for production sectors and weakest for defensive, consumer-oriented sectors.

Maberly and Pierce (2004) examine monthly U.S. stock returns over the same 1970–1998 period as Bouman and Jacobsen. By treating the October 1987 (−22.55%) and August 1998 (−15.81%) returns as outliers, the authors purport to show the dependence of the Halloween effect on these two extreme returns. The effect is diminished and not statistically significant at any conventional level in an alternative specification that controls for these two observations. The authors do not provide an objective basis for identifying exactly two outliers and do not investigate the impact of considering additional outliers. Maberly and Pierce (2003) also examine the impact of outliers on the Halloween effect in Japanese equity markets.

Galai, Kedar-Levy, and Schreiber (2008) also posit a relation between the Halloween effect and outliers. In contrast to the results of

* Corresponding author. Tel.: +1 417 836 5567; fax: +1 417 836 6224.
E-mail addresses: SHaggard@MissouriState.edu (K.S. Haggard), Witte@MissouriState.edu (H.D. Witte).
[1] Tel.: +1 417 836 8478; fax: +1 417 836 6224.

1057-5219/$ – see front matter
doi:10.1016/j.irfa.2010.10.001

Maberly and Pierce (2004), Galai et al. (2008) find that, in daily S&P 500 returns, the Halloween effect is significant only after controlling for outliers. This difference in findings might be due to analyzing daily returns versus monthly, the different time period analyzed (1980–2002 versus 1970–1998), or even the dropping of return observations from the sample. “Returns on non-consecutive days, other than weekend returns, are excluded, as they are not daily returns” (Galai et al., 2008, pp. 786–787).

Outliers are an important aspect researchers have investigated as a possible source of the Halloween effect. In this study, we utilize a formal set of econometric techniques known as robust regression to determine the size and significance of the Halloween effect after controlling for extreme returns. The advantage of this approach is that it does not require any *ad hoc* specification of the number of outliers or the number of standard deviations from the mean an observation must be before it is considered an outlier. Rather, robust regression reduces the influence extreme returns have on the ordinary least squares (OLS) estimates in rough proportion to the departure of the observation from the regression fit. Our approach eliminates the possibility of data mining in the determination of the number of outliers for which to control.

Lucey and Zhao (2008) examine U.S. stock data from 1926 to 2002 to determine the robustness of the Halloween effect to the consideration of the January effect, first identified by Wachtel (1942) and reinforced by Rozeff and Kinney (1976), in which equity returns are significantly higher in January than in other months. They find no evidence of a Halloween effect in their full sample. Using subperiod analysis, they show that neither the Halloween effect nor the January effect are consistently significant for value-weighted returns and that only the January effect is consistently significant for equal-weighted returns. The authors contend that the Halloween effect, when it does appear, might simply be the January effect in disguise.

The subperiod findings of Lucey and Zhao (2008) are likely the result of the relatively short subperiods they examine, as we demonstrate in this study. Over three subperiods since 1946, their average estimate of the Halloween effect for the CRSP value-weighted index is a large 1.02% per month. However, because the sampling subperiods are small, the tests have reduced power and statistical significance is found only in the 1946–1965 period. We update the CRSP returns through 2008 and, with larger subperiods, find a Halloween effect over the most recent 55 years that it is significant and independent of the January effect. Like Lucey and Zhao, we find no evidence of a Halloween effect in the earliest subperiod we examine, 1926–1953. However, over both the 1954–1980 and the 1981–2008 subperiods, our evidence suggests a sizable Halloween effect that is independent of the January effect. This might explain the timing of the earliest reference to “sell in May and go away,” which appears in a 1964 issue of the *Financial Times*.

In this study, we examine the robustness of the Halloween effect to outliers and the January effect over the period from 1926 to 2008. We also investigate how investment in a “Halloween portfolio,” which holds equities from November to April (excluding January) and Treasury bills the remainder of the year, might improve upon the Sharpe (1966) ratio attainable using a buy-and-hold investment strategy.

3. Sample and method

We use monthly value-weighted and equal-weighted stock returns from the Center for Research in Security Prices (CRSP) over the period 1926–2008. We use the following regression model, identical to that of Lucey and Zhao (2008), in our examination:

$$R_t = \alpha + \beta_1 W_t + \beta_2 J_t + \epsilon_t \quad (1)$$

where R_t is the return on the index, W_t is the Halloween indicator, which has a value of “1” in the months from November to April and “0” otherwise, and J_t is the January indicator, which has a value of “1” in January and “0” otherwise.

In addition to using OLS regression, which is sensitive to outliers, we use the *M-estimation* techniques of Huber (1964) and Hampel (1974), which are more robust in the presence of outliers. OLS coefficient estimates are the solution to a sum of squared errors minimization problem. Each sample observation has an associated squared error, which receives the same weight, $w(e)=1$, in the following minimization, $\min_\beta \sum_{t=1}^{T} w(e_t)e_t^2$. The concept behind the *M*-estimators of Huber and Hampel is to dampen the influence of extreme errors by applying reduced weights to larger squared errors. Both estimators set $w(e)=1$ for errors up to one or more threshold levels but reduce weights for errors beyond these levels. For example, the Huber estimator has a single threshold, $k\sigma$, and sets $w(e)=1$ for $|e| \leq k\sigma$ and sets $w(e)=k\sigma/|e|$ for $|e|>k\sigma$. The parameter k in the threshold level is referred to as a “tuning constant” and it is common practice in application to estimate the “scale factor” σ, with $\hat{\sigma} = MAR/(.6745)$, where MAR is the median absolute error from OLS. Aside from the different number of threshold levels, the primary differentiating point of the Hampel estimator is that it applies a finite rejection point, beyond which the observation is classified as an outlier and given zero weight.

The estimation of the Huber and Hampel regressions can be viewed as a weighted least squares problem, minimizing the generalized sum of squared errors expression above. The solution to this minimization is given by $\hat{b} = (X'WX)^{-1}X'WY$, where W is a $T \times T$ diagonal weight matrix with elements $w_{tt}=w(e_t)$. The weights, however, depend upon the errors, the errors depend upon the estimated coefficients, and the estimated coefficients depend upon the weights. An iterative solution, in this case iteratively reweighted least squares (IRLS), must be applied.

To perform IRLS we start by first computing OLS estimates. Second, from the error terms, we calculate weights according to the weighting scheme specified by the particular *M*-estimator. Third, we solve for the new weighted least squares coefficient estimates. We repeat the second and third steps, using the error terms from the previous iteration, until the estimated coefficients converge. Upon convergence, we calculate an estimated asymptotic covariance matrix for the coefficients to determine their statistical significance (see Fox, 1997, for details).

Our selection of tuning constants is based on the work of Hoaglin, Mosteller, and Tukey (1983), who study the effect of varying these constants for many different robust estimators. In general, a smaller tuning constant provides more resistance to outliers at the expense of lower efficiency in the case of normally distributed errors. We select mid-range values for the tuning constants in both the Huber and Hampel regressions. As noted above, we use the normalized median absolute deviation as the scale factor in the Huber regressions. As is common practice, we use the median absolute deviation as the scale factor in the Hampel regressions.

We perform the OLS, Huber, and Hampel regressions for the entire U.S. sample period and for the subperiods of 1926–1953, 1954–1980, and 1981–2008. We also perform these regressions using a global sample of the same 37 countries examined by Bouman and Jacobsen (2002). To further demonstrate the impact of outliers on the Halloween effect, we use a deletion diagnostic method (Belsley, Kuh, & Welsch, 1980) to calculate the sensitivity of the estimated regression coefficient for the Halloween indicator to extreme observations.

Last, in order to assess the investment significance of the Halloween effect, we form mean-variance efficient portfolios using the method of Britten-Jones (1999). Britten-Jones (1999) demonstrates that OLS regression can be applied to determine optimal portfolio weights. The Britten-Jones framework regresses a T-vector of ones on K independent variables representing the excess returns on K risky assets (or portfolios) considered for investment from time 1 to T.