JEL Classification: C58, G1, G12, G14
Keywords: night and daytime returns, market efficiency, US equity exchange-traded funds, calendar effects, multiple structural change

# Is There Seasonality in Traded and Non-Traded Period Returns in the US Equity Market? A Multiple Structural Change Approach*

João Dionísio MONTEIRO - Department of Management and Economics, NECE, University of Beira Interior, Covilha, Portugal (jdm@ubi.pt), *corresponding author*

José Luis MIRALLES-QUIRÓS - Department of Financial Economics, University of Extremadura, Badajoz, Spain (miralles@unex.es)

José Ramos PIRES MANSO - Department of Management and Economics, NECE, University of Beira Interior, Covilha, Portugal (pmanso@ubi.pt)

*Abstract*

*This paper simultaneously examines day-of-the-week, turn-of-the-month and pre- and post-holidays calendar effects in traded and non-traded daily period returns in a group of broad-index exchange traded funds (ETFs) that track the major US stock indices (S&P 500, DJIA 30, NASDAQ 100 and Russel 2000 index). Bai and Perron (1998, 2003)'s method is employed to examine stability of the significant calendar effects over time and across ETFs. Results exhibit a high instability of the significant calendar effects over the various regimes up until 2001. From 2001 onwards and until the end of 2013, only a single regime across all ETFs is identified. In this last regime, results point to the disappearance of the previous significant effects across the US equity ETFs group. Unstable observed effects could have been motivated by market-specific conditions in such short time periods. The disappearance of these effects from 2001 onwards are consistent with the nature of this asset class: these ETFs are broadly diversified portfolios with diversification of private information, with higher liquidity and lower transactions cost, which is likely to reduce potential calendar effects.*

## 1. Introduction

Empirical studies in the market efficiency field have been extensive in finance. According to Fama (1970), evidence on return anomalies does not appear to be significant and of sufficient importance to justify rejection of the efficient market hypothesis, nor does it serve as a theoretical hallmark of fundamentally profitable trading rules. In the following decades, however, a large number of studies examining return anomalies were carried out, of which calendar anomalies are the most common. Results suggest that these could be pervasive and able to be used as profitable trading rules. Some suggest, however, its decline of importance or even its disappearance.

The most investigated calendar effects include: the weekend effect, where returns are lower between Friday close and Monday close (Thaler, 1987; Abraham and Ikenberry, 1994; Pearce, 1996; Zainudin and Coutts, 1997); the day-of-the-week

*We are most grateful to the editorial team and to two anonymous referees for their helpful comments. This work was supported by NECE, a I&D unity funded by the FCT - Portuguese Foundation for the Development of Science and Technology, Ministry of Science, Technology and Higher Education.

effect, where returns tend to vary throughout the week - null or negative returns on Monday and positive on Friday - (Chang *et al.*, 1993, 1998; Pettengill, 2003); the holiday effect, where returns are higher on trading days preceding public holidays (Thaler, 1987; Lakonishok and Smidt, 1988; Ariel, 1990; Pearce, 1996; Brockman and Michayluk, 1998; Chong *et al*., 2005) and the turn-of-the month effect, where returns are higher on the last trading days of the month and the first few days of the following month (Cadsby and Ratner, 1992; McConnell and Xu, 2008). These empirical studies, however, mostly use in their analysis returns calculated as close-to-close daily prices.

As trading can be perceived as a continuous-time process during the entire day, theoretical and empirical studies have focused their interest on decomposing daily (close-to-close) returns into non-trading (close-to-open) and trading (open-to-close) periods, examining implications on trading and returns (Hong and Wang, 2000; Barclay and Hendershott, 2003; Branch and Ma, 2006; Cliff *et al*. 2008; Kelly and Clark, 2011; Berkman *et al.*, 2012; Lachance, 2015; Lou *et al*., 2015). Evidence points to the prevalence of statistically higher returns during overnight. Also, empirical studies documented that the first, the second and higher order moments of the return generating process are different over trading and non-trading periods (Cliff *et al*. 2008; Tompkins and Wiener, 2008) and that risk-adjusted overnight returns are significantly higher than risk-adjusted daytime returns (Kelly and Clark, 2011). Furthermore, given the pervasiveness of this effect, empirical studies suggest that these differences could be of economic significance through implementation of profitable trading strategies, even when incorporating realistic transaction costs (Kelly and Clark, 2011; Lachance, 2015).

The pervasiveness of this effect across empirical studies, contrary to the efficient market hypothesis, points to the need for further research. The above cited evidence motivates carrying out the present study and examine the verifiability and the robustness of this effect in areas not yet exploited, specifically in the framework of calendar effects. The present study rest on the rationale of market efficiency. The aim of this paper is to add to the field of market efficiency an analysis of calendar effects during trading and non-trading daily period returns in US equity market of exchange-traded funds (ETFs). Although less scrutinized with respect to trading and non-trading daily periods and specifically in the ETFs market, the US equity market has been extensively addressed in calendar effect studies. In this vein, as far as we know, a comprehensive and simultaneous analysis containing various calendar effects during trading and non-trading daily periods remains to be done. Likewise, the simultaneous relative strength of each effect in the calendar regression model has not yet been examined. Indeed, previous studies have not examined over time the variability of these effects, which may have sprung from regulatory changes in the trading microstructure, which occurred in the middle of the 2010s in US equity markets or from improvements in the information impounding into prices over the last decades. Most likely, if calendar effects on night and daytime returns exhibit instability over time, it may happen that the significant effects are only momentary and arise from period-specific features and would not persist. This study aims to analyze these issues.

In this paper we simultaneously examine and identify the relative strength of the day-of-the-week, turn-of-the month and pre- and post-holidays calendar effects in traded and non-traded daily period returns in a set of ETFs that track the major US stock indices. Firstly, for the entire sample period, we examine for the existence of calendar effects on night and daytime returns, decomposed by the above-mentioned effects. Secondly, we employ the procedure of Bai and Perron (1998, 2003) to examine over time the existence and the stability of the significant night minus daytime return difference previously found. We evaluate whether these effects really occurred during earlier sample periods, whether they are still present, if they have diminished in magnitude or even disappeared.

Across ETFs and over the full sample period, results exhibit significant positive night returns in the last and the first trading days of the month. However, when examining stability of the significant coefficients over time, Bai and Perron (1998, 2003)'s method shows a higher variability in the significant regression coefficients since the beginning of each ETF sample period until 2001. From 2001 onwards and across ETFs, results show a decrease or even disappearance, not granting support for the existence of the effects in this sub period. These results suggest that the significant effects were motivated by sub period-specific market conditions, that markets became more efficient in information impounding into open and close prices, and that disappearance of the effects are consistent with the characteristics of this asset class: ETFs are broadly diversified portfolios with diversification of private information, with higher liquidity and lower transactions cost (Hasbrouck, 2003).

The remainder of the paper is organized as follows: In section 2, theoretical causes, predictions and empirical evidence on the behavior of returns during trading and non-trading daily periods are reviewed. Some empirical evidence of calendar effects on trading and non-trading periods is also reviewed. Section 3 presents the data, the calendar effect regression model, the structural change model and the corresponding statistical tests of Bai and Perron (1998, 2003)'s method. We present and discuss results in section 4. In section 5, summary results and some concluding remarks are given.

## 2. Literature Review

In financial markets, information flows continuously around the clock but price variations and trading are not continuous due to periodic market closure. Changes in daily transaction regimes, when markets open and close, can have important implications for the return generating process over trading and non-trading periods. Empirical studies have reported that the mean return, the trading volume, the volatility and the bid-ask spreads in general have a U-shaped pattern during the intraday trading period across developed stock markets, with these variables being high at the open and close of the market and relatively flat during the middle of the intraday trading period (Wood *et al.*, 1985; McInish and Wood, 1992; Foster and Viswanathan, 1993; Abhyankar *et al.*, 1997; Hong and Wang, 2000; Chow, *et al.*, 2004). However, less consensus exists about the behavior of the mean return during trading and non-trading daily periods.

Theoretical papers have sought to model the implications of periodic market closure for equilibrium prices (Longstaff, 1995; Hong and Wang, 2000). These models, however, suggest different predictions of the effects of periodic market closure at the first and second moments of returns. Hong and Wang (2000)'s model predicts lower returns during non-trading than in trading periods, a prediction consistent with the observed higher volatility and information flow rates during the trading period. Longstaff (1995)'s model predicts higher returns during non-trading than at trading periods to compensate liquidity providers for bearing additional risk, i.e., higher returns over non-trading periods would arise from a liquidity related non-marketability effect.

Wood *et al.* (1985) were among the first to examine return patterns around the open and the close of the market. Using high frequency of transaction data, they document the return and volatility to be unusually high at the open and close of the trading daily period. French and Roll (1986) document stock returns to be more volatile during trading than at the non-trading period, attributing the higher volatility during trading hours to the differences in information flow rates between the two periods, i.e., more private information being incorporated during the day. George and Wang (2001) examine the rate of information flow and finds that daytime information rate is about seven times higher than overnight rate. Harris (1989) document a large mean price increase before market closure and this effect is persistent across stocks and days. This movement also tends to be observed at the opening of the market. Barclay and Hendershott (2003) finds that there is less information asymmetry in the post-close than in the pre-open of the market. Their findings suggest that there will be a higher fraction of liquidity motivated trades in the post-close and a higher fraction of informed trades in the pre-open.

Concerning return patterns over trading and non-trading periods, empirical evidence is not consistent across empirical studies. French (1980) first identified the weekend effect using US daily stock returns from 1953 to 1977. French finds a weekend effect where Monday's mean return is significantly negative, while the other day-of-the-week returns are significantly positive. Rogalski (1984) examines the U.S. stock market from 1974 to 1984 to see whether the weekend effect is a closed market effect by decomposing daily close-to-close returns into a non-trading and trading return. Rogalski finds that the negative weekend return is composed of a negative Monday non-trading return (Friday close to Monday open) and a Monday trading return (Monday open to close) identical to the trading returns of other weekdays.

Cliff *et al.* (2008), using datasets of different asset classes for the period 1993-2006, perform an extensive study in US equity markets on the overnight and daytime returns. They document that the US equity premium during this decade is entirely due to overnights returns: the returns during the night being strongly positive and returns during the day being close to zero and sometimes negative. They do show that this day and night effect is found on individual equities, equity indices, ETFs and futures contracts on equity indices.

Tompkins and Wiener (2008) examine returns for five global index futures markets (S&P 500 futures, FTSE 100, DAX, CAC 40 and Nikkei 225) over traded and non-traded periods. They find significant differences between traded and non-traded period returns. For the US market the mean return is higher for the trading than in the non-traded period, with the non-traded period having significantly lower variance. For the four non-US stock markets, the non-traded period return is significantly higher than the trading period. They attribute this positive non-trading minus trading return spread to differences in regulatory risk management requirement between US and non-US equity derivative market-makers.

Clark and Kelly (2011) compares the intraday and overnight returns on a set of US equity ETFs. Using the Sharpe ratio (SR) measure, they found the overnight SR to significantly exceed the intraday SR, implying that the premium one receives by taking on risk, is higher at night than at daytime. Berkman *et al.* (2012) examine predictions with regard to intraday patterns in retail order flow and price formation in a sample of the largest US stocks. Based on the theory of attention-based overpricing at the opening of the market, they report the existence of the night effect confined to a large US stock group and attribute the significantly positive (negative) overnight (daytime) return to the trading activity of retail investors. Qiu and Cai (2013) examine the anomaly of superior overnight returns on international stock markets. Using stock index data for thirty-two countries, they find that the anomaly exists in twenty countries including both developed and emerging markets and that the superior overnight returns are not justified by the risk-return trade off as overnight are less volatile than trading-period returns. Their results suggest that greater divergences of opinion lead to higher overnight return premiums and that short sale constrains exacerbate the anomaly. Lachance (2015), using all listed US stocks in the period 1995-2014, but not including ETFs, finds evidence that overnight returns are subject to highly persistent and positive biases in a large group of stocks. She extends the analysis to index components of each of the 23 countries of MSCI´s world index and mostly obtains similar results. Lou *et al.* (2015) show that, on average, all of the abnormal returns on momentum trading strategies occur overnight. They attribute the higher overnight return to the negative intraday pressure put by institutional investors on momentum stocks. Overall, results in the above empirical studies are not entirely consistent. Some of these points to the existence of higher overnight returns across all the individual assets of the sample, others report the effect confined to a group of assets, while others report an inverse effect.

Several arguments have been put forward to explain this night effect, namely, the timing of earning announcements, asset liquidity and investor trading heterogeneity. Early papers on the timing of earnings announcements found that companies tended to publicize good new during the market open and bad news after the market close (Patell and Wolfson, 1982). Damodaran (1989) showed that announcements of earnings and dividends made on Friday are actually more likely to contain bad news and result in subsequent negative returns during the weekend. Bagnoly *et al.* (2005) find that announcements made on Fridays, during the trading period and after the market close, are more negative than on other days of the week. Doyle and Magilke (2009) find no evidence that managers strategically choose to disclose negative information after the close of the markets or on Friday. They also find no evidence that managers decide to report “good” news before the opening of the markets. Jiang *et al.* (2012) argues that the timing of announcements changed and report that for stocks of the S&P 500, from 2004-2008, more than 95% of announcements are made overnight.

Concerning the asset liquidity issue, Amihud (2002) documents a negative relationship between various measures of liquidity and future stock returns, suggesting that increased (lower) risk or transactions costs of low (high) liquidity stocks would predict a more (less) night minus daytime return spread. Cliff *et al.*

(2008) tested the disclosure timing and the asset liquidity hypotheses and found no support to explain the positive and significant night minus day return spread observed in their study.

Investor trading heterogeneity during trading and non-trading periods is also suggested as contributing to the effect. Barclay and Hendershott (2003) report that there is a higher fraction of liquidity motivated trades in the post-close and a higher fraction of informed trades in the pre-open, being the trading in the pre-open dominated by large informed investors. They report that the pre-open period has the greatest amount of price discovery per trade. Clark and Kelly (2011) attribute the night effect to the behavior of active day (semi-professional) traders: not wanting to hold stocks over a non-trading period would push the day traders to buy at morning and sell at night. In settling and opening their positions, prices would increase by the buy and decrease by the sell pattern of the day traders. Jiang *et al.* (2012) provide evidence that firms prefer overnight announcements because trades in this period are mainly from informed investors and these trades are relied upon to convey information to the general public. Berkman *et al.* (2012) suggest that trading activity of retail investors have an important role in explaining higher overnight returns through their herding behavior in the high-attention stock group that pushes opening prices up. However, Lachance (2015) reports that the excess overnight for the overnight bias stock group disappear quickly, but not instantaneously, after the open. Lou *et al.* (2015) points to the negative intraday pressure that institutional investors put on momentum stocks.

The above evidence suggests that findings on the trading and non-trading period returns are not entirely consistent across markets, even within the same market, and that these results could be sample period-dependent and asset-specific. These empirical findings motivate the present study to further investigate night and daytime effects, examining whether in the calendar context these effects really occurred in the early years of the sample, whether they are still present and if they have diminished or even disappeared. We use an appropriate methodology to answer these questions.

The present study focusses on an ETF group that track the major US stock indexes. Although some previous studies focused on this asset group and obtained significant evidence of positive overnight return, they used a shorter sample period and different methodologies. We are not aware that a comprehensive and simultaneous analysis of night and daytime returns, across various calendar effects and examining the relative strength of each effect, was done. Also, as the night and daytime effects may be time varying, owing to the ETF specificity (trades at once a highly diversified portfolio of stocks), to improvements in information impounding and to time varying market conditions, an approach to capture variability of the significant estimated effects has not yet been applied by previous empirical studies. This study intends to fill these gaps.

## 3. Data and Methodology

### 3.1 Data

The data employed in this study are actual opening and closing daily prices from a group of four ETFs that track major US equity market indexes. The four ETFs used are the DIA (representing the Dow Jones Industrial Average 30 index), the IWM (representing the Russel 2000 index - a small-cap US companies index), the QQQ (representing the NASDAQ 100 index) and the SPY (SPYDERs - representing the S&P 500 index). ETFs allow investors to trade a basket of stocks in a single transaction. The creation and destruction features of the ETF ensure that prices on the exchange closely reflect the fair value of the underlying portfolio´s components.

The analysis of the ETFs returns offer advantages over the analysis of the indexes returns for two reasons. First, the share price of an ETF is the price for the entire portfolio, with no problem of asynchronous transactions on certain stocks in the index. Second, ETFs that track major stock market indexes are highly liquid, with very low transaction costs (bid-ask spreads) involved in the trading of these instruments. Also, two specific and useful features of ETFs are that the transaction is an in-kind trade – i.e., securities are traded for securities – and are generally more tax-efficient.

Our ETF sample was previously used by Kelly and Clark (2011) in their analysis of night and daytime SR over the sample period 1996-2006. This set of ETFs began trading on the Exchanges at different years. According to Kelly and Clark, the liquidity of these ETFs was poor during the mid-1990s and has enormously increased during this decade. The SPY started trading in 1993 but its liquidity was poor during the first-half of 1990s. For each ETF, to determine the starting point of the analysis, we follow the liquidity criterion used by Kelly and Clark[1]. Thus, SPY time series data are used from 01/02/1996 (mm/dd/yyyy), DIA time series data are used from 01/21/1998, QQQ time series data are used from 03/11/1999 and IWM time series data are used from 01/02/2001 to 01/03/2014.

For each ETF, we compute returns during the two daily sub-periods: the night (close-to-open prices) and daytime (open-to-close prices) returns. As in most of the analysis of daily and intra-daily financial data, we work with continuously compound return and we compute the night and daytime returns, respectively, as:

$$r_t^n = ln[P_t^o / P_{t-1}^c].100\%, \tag{1.1}$$

$$r_t^d = ln[P_t^c / P_t^o].100\%, \tag{1.2}$$

where $P_t^o$ is the ETF price level at the open of day $t$ , $P_t^c$ is the ETF price level at the close of day $t$ and $P_{t-1}^c$ is the ETF price level at the close of day $t-1$. The average returns are geometric averages and, therefore, its sign indicates whether the ETF gained or lost value during this intraday range over the sample period. For each ETF return time series, day and night returns on the various calendar effects are identified with dummy variables: by days of the week, by whether they precede or are after a public holiday, and by night and daytime returns on the last trading day of the month and on the first and the second trading day of the following month. The Monday night returns are computed as Friday close to Monday open, while Monday daytime returns are computed as Monday open-to-close. For the holiday effect, for instance, if

[1] For each ETF, data were not used from years in which the 5th percentile time of the first trade of the day is not in the first ten minutes of the trading day or the 5th percentile time of the last trade before 4 pm is not between 3:50 pm and 4:00 pm (Kelly and Clark, 2011).

a public holiday occurs on Monday, the pre-holiday night returns are computed as Thursday close to Friday open, the pre-holiday daytime returns are computed as Friday open-to-close, the post-holiday night returns are computed as Friday close to Tuesday open and the post-holiday daytime returns as Tuesday open-to-close. For the turn-of-the-month returns, for example, if January, 31th occurs on a Friday, the night and daytime returns of the last trading day of the month are computed as Thursday-close to Friday-open and Friday open-to-close, respectively. The night and daytime returns of the first and second trading days of the following month would be computed as Friday close to Monday open, Monday open-to-close, Monday close to Tuesday open and Tuesday open-to-close, respectively. Returns over the extended close after the September 11$^{th}$ 2001 tragedy were not taken into account.

### 3.2 The Calendar Effect Model

Three approaches are used to examine hypotheses of day and night returns on various calendar effects on US equity ETFs. The first involves conducting a descriptive analysis and parametric tests of night and daytime returns by day-of-the-week. For each ETF and day of the week we perform tests of equality of means and variances between night and daytime returns, using parametric tests. Parametric testing is suitable because for large samples, sample means will be normally distributed even if the underlying variables are not normally distributed, with the ratio of the two sample variances following an F-distribution. Additionally, parametric tests have more statistical power than their nonparametric counterparts. The second approach involves simultaneously examining trading and non-trading period returns over three calendar effects using a regression-based analysis. For each ETF return series, the following regression model is specified for simultaneously examining the three calendar effects and assesses the relative magnitude and significance of each effect:

$$r_t = \beta_0 + \sum_{j=1}^{4} \beta_j x_{j,t} + \sum_{k=1}^{5} \alpha_k x_{k,t} + \sum_{l=1}^{2} \gamma_l h_{l,t} + \sum_{m=1}^{2} \delta_m h_{m,t} + \sum_{n=1}^{3} \theta_n y_{n,t} + \sum_{p=1}^{3} \vartheta_p y_{p,t} + \varepsilon_t, \quad (2)$$

where $r_t$ is the return, $x_j$ is a dummy variable taking the value of one for the non-trading period return of day $j$ and zero otherwise ($j = 1,2,3,4$) (the reference category is the non-trading period return on Wednesday), $x_k$ is a dummy variable taking the value of one for the trading period of the day $k$ and zero otherwise ($k = 1,2,3,4,5$), $h_j$ is a dummy variable taking the value of one for the non-trading ($l = 1$) and trading period return ($l = 2$) of a day preceding a public holiday and zero otherwise, $h_m$ is a dummy variable taking the value of one for the non-trading ($m = 1$) and trading period return ($m = 2$) of a day following a public holiday and zero otherwise, $y_n$ is a dummy variable taking the value of one for the non-trading period return of the last trading day ($n = 1$) and the first and second trading days of the following month ($n = 2,3$) and zero otherwise, $y_p$ is a dummy variable taking the value of one for the trading period return of the last trading day ($p = 1$) and the first and second trading days of the following month ($p = 2,3$) and zero otherwise, and $\varepsilon_t$ is the error term.

$\beta_0, \beta_j, \alpha_k, \gamma_l, \delta_m, \theta_n, \vartheta_p$ are the model' parameters to be estimated. The $\beta_0$ estimate represents the non-trading period mean return on Wednesday that neither precedes nor follows a public holiday and neither occurs on the last trading day of the month nor on the first and second trading day of the following month. The $\delta_1$ parameter estimate, for instance, represents the increase on the non-trading mean return of a day following a public holiday, vis-à-vis the non-trading period mean return on Wednesday that neither precedes nor follows a public holiday and that neither occurs on the last trading day of the month nor on the first and second trading day of the following month. To examine the appropriateness of the return series for a regression-based analysis we evaluate the stationarity of the ETFs group return series conducting the null hypothesis of unit roots.

### 3.3 The Multiple Structural Change Model

The third approach involves examining over time the stability of the significant coefficients in model (2). Thus, we estimate model (2) and make inference in the context of multiple structural change model using Bai and Perron (1998, 2003)'s method. The ETF return time series, decomposed according to various calendar effects, may contain multiple structural changes, reflecting calendar effects' variability. Hereafter, are presented the method and the associated test statistics of Bai and Perron's procedure applied to our regression model (2).

As in Bai and Perron's terminology, we hypothesize that the linear regression model (2) is a pure structural change model in the sense that all parameters can vary across the various regimes and can be expressed in matrix form as

$$\boldsymbol{R} = \overline{\boldsymbol{Z}}\boldsymbol{\delta} + \boldsymbol{E}, \tag{3}$$

where $\boldsymbol{R} = (\boldsymbol{r_1}, \boldsymbol{r_2}, \dots, \boldsymbol{r_{m+1}})'$ is the vector with the corresponding observed returns in the $m + 1$ regimes, $\overline{\boldsymbol{Z}}$ is the matrix which diagonally partitions $\boldsymbol{Z}$, the matrix with the corresponding vectors of covariates (dummy variables) at the m-partition $(T_1, T_2, \dots, T_m)$, i.e., $\overline{\boldsymbol{Z}} = diag(\boldsymbol{Z}_1, \boldsymbol{Z}_2, \dots, \boldsymbol{Z}_{m+1})$ with $\boldsymbol{Z}_i = (\boldsymbol{z}_{T_{i-1}} + 1, \dots, \boldsymbol{z}_{T_i})'$, $\boldsymbol{\delta} = (\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \dots, \boldsymbol{\delta}_{m+1})'$ are the corresponding vectors of coefficients and $\boldsymbol{E} = (\boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_2}, \dots, \boldsymbol{\varepsilon_{m+1}})'$ are the corresponding vectors of disturbances. The break dates $(T_1, T_2, \dots, T_m)$ are explicitly treated as unknown and for $i = 1,2, \dots, m$ , we have $\lambda_i = T_i/T$ with $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_m < 1$. Bai and Perron (1998) impose some restrictions on the possible values of the break dates. They define the following set for some arbitrary small positive number $\epsilon$: $\Lambda_\epsilon = \{(\lambda_1, \dots, \lambda_m); |\lambda_{i+1} - \lambda_i| \geq \epsilon, \lambda_1 \geq \epsilon, \lambda_m \leq 1 - \epsilon\}$ to restrict each break date to be asymptotically distinct and bounded from the boundaries of the sample and to assure enough observations to estimate all the sub-sample parameters. As proposed by Bai and Perron (1998), the estimation method is that based on the least-squares. For each m-partition $(T_1, T_2, \dots, T_m)$, denoted $\{T_j\}$, the associated least-squares estimate of $\boldsymbol{\delta}_j$ is obtained by minimizing the sum of squared residuals (SSR) $\sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} (r_t - \boldsymbol{z}_t' \boldsymbol{\delta}_i)^2$ . Let $\widehat{\boldsymbol{\delta}}\{T_j\}$ denote the resulting estimate, containing the vectors of coefficient estimates for each regime. Substituting it in the objective function and denoting the resulting SSR as $S_T\ (T_1, T_2, \dots, T_m)$, the estimated break dates $(\hat{T}_1, \hat{T}_2, \dots, \hat{T}_m)$ are such that

$$\left(\hat{T}_1, \hat{T}_2, \dots, \hat{T}_m\right) = \arg \min_{(T_1, T_2, \dots, T_m)} S_T\left(T_1, T_2, \dots, T_m\right), \tag{4}$$

where the minimization is taken over all partitions $(T_1, T_2, \dots, T_m)$ such that $T_i - T_{i-1} \geq [\epsilon T]$. The break point estimators are global minimizers of the objective function. The regression parameter estimates for each segment are the associated least-squares estimates at the global minimizer estimated m-partition $\{\hat{T}_j\}$, i.e., $\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{\delta}}\{\hat{T}_j\}$. To estimate the break dates and the associated regression coefficients for each segment, the efficient estimation algorithm developed in Bai and Perron (2003) is used, which is based on the principle of dynamic programming and which allows global minimizers to be obtained using a number of SSR that is of order $O(T^2)$ for any $m \geq 2$.

## 3.4 Test Statistics for Multiple Structural Changes

### 3.4.1 A Test of Structural Stability versus a Fixed Number of Changes

Bai and Perron (1998) propose the $\sup F$ type test of no structural break $(m = 0)$ versus the alternative hypothesis that there are $m = k$ breaks. Let $(T_1, T_2, \dots, T_k)$ be a partition such that $T_i = [\lambda_i T]$ $(i = 1,2, \dots, k)$. Let $\boldsymbol{R}$ be the conventional matrix such that $(\boldsymbol{R\delta})' = (\boldsymbol{\delta}_1' - \boldsymbol{\delta}_2', \boldsymbol{\delta}_2' - \boldsymbol{\delta}_3', \dots, \boldsymbol{\delta}_k' - \boldsymbol{\delta}_{k+1}')$ and $\boldsymbol{M}_x = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Let

$$F_T(\lambda_1, \dots, \lambda_k; q) = \left(\frac{T-(k+1)q-p}{kq}\right) \frac{\hat{\boldsymbol{\delta}}'\boldsymbol{R}'\left(\boldsymbol{R}(\overline{\boldsymbol{Z}}'\boldsymbol{M}_x\overline{\boldsymbol{Z}})^{-1}\boldsymbol{R}\right)^{-1}\boldsymbol{R}\hat{\boldsymbol{\delta}}}{SSR_k}, \tag{5}$$

where $SSR_k$ is the sum of squared residuals under the alternative hypothesis which depends on $(T_1, T_2, \dots, T_k)$. The $\sup F$ type test statistic is defined as $\sup F_T(k; q) = \sup_{(\lambda_1, \dots, \lambda_k) \in \Lambda_\epsilon} F_T(\lambda_1, \dots, \lambda_k; q)$. An asymptotically equivalent version could be obtained using the break dates estimates obtained from the global minimization of the SSR. The test would be defined as $\sup F_T(k; q) = F_T\left(\hat{\lambda}_1, \dots, \hat{\lambda}_k; q\right)$.

The limiting distribution of the test (5) depends on the presence or absence of serial correlation and heterogeneity in the residuals. When serial correlation and heterogeneity are present in the residuals, Bai and Perron (1998; 2003) suggest using the following version of the F test

$$F_T^*(\lambda_1, \dots, \lambda_k; q) = \left(\frac{T-(k+1)q-p}{kq}\right) \hat{\boldsymbol{\delta}}'\boldsymbol{R}'\left(\boldsymbol{R}\hat{\boldsymbol{V}}(\hat{\boldsymbol{\delta}})\boldsymbol{R}'\right)^{-1}\boldsymbol{R}\hat{\boldsymbol{\delta}}, \tag{6}$$

where $\hat{\boldsymbol{V}}(\hat{\boldsymbol{\delta}})$ is an estimate of the variance–covariance matrix of $\hat{\boldsymbol{\delta}}$ that is robust to serial correlation and heteroscedasticity, i.e. a consistent estimate of

$$\boldsymbol{V}(\hat{\boldsymbol{\delta}}) = \operatorname{plim} \boldsymbol{T}(\overline{\boldsymbol{Z}}'\boldsymbol{M}_x\overline{\boldsymbol{Z}})^{-1}\overline{\boldsymbol{Z}}'\boldsymbol{M}_x\boldsymbol{\Omega}\boldsymbol{M}_x\overline{\boldsymbol{Z}}(\overline{\boldsymbol{Z}}'\boldsymbol{M}_x\overline{\boldsymbol{Z}})^{-1}, \tag{7}$$

where, for a pure structural change model, $\boldsymbol{M}_x = \mathbf{I}$ and $\boldsymbol{\Omega}$ is the variance-covariance matrix of residuals incorporating the serial correlation and heterogeneity. The $F_T^*$ statistic is the conventional F-statistic for testing $\boldsymbol{\delta}_1 = \boldsymbol{\delta}_2 = \dots = \boldsymbol{\delta}_{k+1}$ against $\boldsymbol{\delta}_i \neq$

$\boldsymbol{\delta}_{i+1}$ for some $i$ given the partition $(T_1, T_2, \ldots, T_k)$. If in our estimated ETFs time series regression residuals exhibit serial correlation and heteroskedasticity, we compute the statistic (6) using the HAC Newey-West variance-covariance matrix of $\widehat{\boldsymbol{V}}(\widehat{\boldsymbol{\delta}})$.

### 3.4.2 A Double Maximum Test

In the above test it is assumed that the alternative number of break dates $m = k$ is pre-specified. However, for inference purposes the researcher may wish not to pre-specify a particular number of breaks. In cases where the number of breaks is not known, Bai and Perron (1998) introduced two tests of the null hypothesis of no structural break against an unknown number of breaks given some upper bound $M$. These are called the double maximum tests since they involve maximization both for a given $m$ and across various values of the test statistic for $m$.

This new class of tests is defined for some fixed weights $\{a_1, a_2, \ldots, a_M\}$ as

$$D \max F_T^* (M; q) = \max_{1 \leq m \leq M} a_m \sup_{(\lambda_1, \ldots, \lambda_m) \in \Lambda_\epsilon} F_T^*(\lambda_1, \ldots, \lambda_m; q). \tag{8}$$

The weights reflect the imposition of some priors on the likelihood of various numbers of structural breaks, where $a_m = c(q, \alpha, 1)/c(q, \alpha, m)$ and $c(q, \alpha, m)$ denote the asymptotic critical value of the test $\sup_{(\lambda_1, \ldots, \lambda_m) \in \Lambda_\epsilon} F_T(\lambda_1, \ldots, \lambda_m; q)$ for a significance level $\alpha$. First, they propose the equal weighted version where all weights are set equal to unity, i.e., $a_m = 1$,

$$UD \max F_T (M, q) = \max_{1 \leq m \leq M} F_T\big(\hat{\lambda}_1, \ldots, \hat{\lambda}_m; q\big), \tag{8.1}$$

where $\hat{\lambda}_1, \ldots, \hat{\lambda}_m$ are the estimates of the break points obtained from the global minimization of the SSR. As an alternative, they propose a second version of the test where the weights are defined as $a_1 = 1$ and for $m > 1$ as $a_m = c(q, \alpha, 1)/c(q, \alpha, m)$

$$WD \max F_T (M, q) = \max_{1 \leq m \leq M} \frac{c(q, \alpha, 1)}{c(q, \alpha, m)} F_T\big(\hat{\lambda}_1, \ldots, \hat{\lambda}_m; q\big). \tag{8.2}$$

Bai and Perron note that, unlike the $UD \max F_T (M, q)$ test, the value of the $WD \max F_T (M, q)$ depends on the significance level chosen since the weights depends on $\alpha$.

### 3.4.3 A Test of $l$ versus $l + 1$ Breaks

Additionally, Bai and Perron (1998) proposed a test for the null hypothesis of $l$ structural changes against the alternative that an additional change exists, labeled $\sup F_T (l + 1|l)$. The test is applied to each segment containing the observations $\big[\hat{T}_{i-1}, \hat{T}_i\big]$ $(i = 1, 2, \ldots, l + 1)$, $\hat{T}_0 = 1$ and $\hat{T}_{l+1} = T$. We conclude for a rejection in favor of a model with $(l + 1)$ breaks if the overall minimal value of the SSR (over all segments where an additional change is included) is sufficiently smaller than the SSR from the $l$ break model. The break point thus selected is the one associated with this overall minimum. The test is defined as

$$\sup F_T\,(l+1,l) = \Big\{S_T\big(\hat{T}_1,\dots,\hat{T}_l\big) - \min_{1\le i\le l+1}\inf_{\tau\in\Lambda_{i,\eta}} S_T\big(\hat{T}_1,\dots,\hat{T}_{i-1},\tau,\hat{T}_i,\dots,\hat{T}_l\big)\Big\}/\hat{\sigma}^2, \tag{9}$$

where $\Lambda_{i,\eta} = \{\tau;\hat{T}_{i-1} + (\hat{T}_i - \hat{T}_{i-1})\eta \le \tau \le \hat{T}_i - (\hat{T}_i - \hat{T}_{i-1})\eta\}$, $S_T\big(\hat{T}_1,\dots,\hat{T}_{i-1},\tau,\hat{T}_i,\dots,\hat{T}_l\big)$ is the SSR resulting from the least squares estimation from each m-partition $(T_1,\dots,T_m)$ and $\hat{\sigma}^2$ is a consistent estimator of $\sigma^2$ under the null hypothesis.

### 3.4.4 Selection Procedure of Break Dates

Since we conclude that the estimated time series regressions contain structural changes using the $F_T^*(\lambda_1,\dots,\lambda_k;q)$, the $UD\max F_T\,(M,q)$ and the $WD\max F_T\,(M,q)$ tests, we need to determine the number of breaks using a selection procedure. To further improve the selection of the number of breaks and their locations, Bai and Perron (2003) suggests that we first look at the two double maximum tests to see if at least a structural break exists. Then, the number of breaks can be decided based upon an examination of the $\sup F_T\,(l+1|l)$ statistics obtained using the break dates estimates from a sequential global minimization of the SSR. The final number of selected breaks is thus equal to the number of rejections of the $\sup F_T\,(l+1|l)$ tests, assuming that we start with $l=0$ (no structural breaks). Bai and Perron (2003) suggest that this procedure leads to the best results and is recommended for empirical applications.

## 4. Empirical Results

Table 1 summarizes the descriptive statistics for the traded and non-traded period returns for the examined US equity ETFs, decomposed by days of the week. Across the ETF group and for every day of the week, in general, average returns are higher during the non-trading than the trading period. During the trading period and for every day of the week, except at Friday on the QQQ, average returns are not significantly different from zero. For the non-trading period, average returns on Monday and Tuesday for the SPY, on Tuesday for DIA, on Monday, Tuesday and Thursday for QQQ and on Tuesday for IWM are significantly positive. The overall non-trading return for the SPY, DIA, QQQ and IWM is also significantly positive.

The volatility of returns (as measured by standard deviation) for every day of the week and all ETFs, is higher during the trading period. This result is consistent with evidence obtained in previous studies (French and Roll, 1986; Lockwood and Linn, 1990; Cliff *et al.*, 2008) and in line with hypothesis that information flow volume is higher during the trading than the non-trading period (George and Wang, 2001).

The distributional properties of the returns series for all ETFs, days of the week and trading and non-trading periods do not appear to be normal. For almost every day of the week, and trading and non-trading periods and ETFs, return skewness is significant and there seems to exist a pattern about the sign of this parameter. Across ETFs, the skewness is negative (positive) during the trading (non-trading) period on Monday, except on QQQ, positive (negative) during the trading (non-trading) period on Friday and negative during the trading and non-trading period on Thursday, except for QQQ. There is also a common pattern on the signal of this parameter across weekdays and trading periods between SPY and DIA, exhibiting the strong correlation between these two ETFs. These results indicate the higher probability of returns lying below (above) the mean return on Friday during the non-trading (trading) period and for Monday the higher probability of returns lying below (above) the mean return during the trading (non-trading) period.

The kurtosis across ETFs, with returns decomposed by days of the week and by trading and non-trading periods, is significant, indicating leptokurtic distributions with the number of extreme returns being higher than under the normal. Finally, the Jarque-Bera statistics (values not presented), as expected, reject the null of normality of returns across ETFs, over weekdays and trading and non-trading periods.

### 4.1 Parametric Tests of the Mean and Volatility Return Differences

The last two columns of Table 1 present test results for the equality of means and variances. For each ETF and day of the week, statistical values for tests on the difference between daytime and night return are presented. For all ETFs and every day of the week, hypothesis testing of the null of equal variances are rejected at 1%, with the volatility of the non-trading being significantly lower than the volatility of the trading period return. These results are consistent with the hypothesis that the volume of information flow that occurs during the trading period is significantly higher than that observed during the non-trading period. This same result was observed by Stoll and Walley (1990) in the NYSE.

Concerning average returns, results indicate that, in general, across ETFs and days of the week, mean return differences between trading and non-trading periods are not, with few exceptions, significant. The overall night mean return for SPY, the Tuesday, the Friday and overall night returns for QQQ and Monday night return for IWM are significantly higher than the corresponding daytime average returns. These results contrast with those obtained in previous studies (Cliff *et al*., 2008; Kelly and Clark, 2011) where for these same ETFs were obtained pervasive significant differences between trading and non-trading period returns. Notice, however, that the examined sample period in Cliff *et al*. and Kelly and Clark ends in 2006 and the one used in the present study spans until the end of 2013. At first sight, it appears that the night and daytime effect previously found have significantly diminished or even disappeared, despite the pattern of a significantly lower volatility remaining during the non-trading period.

**Table 1 Descriptive Statistics of Night and Daytime Returns by Day-of-the-Week and ETF**

| *Exchange Traded Fund* | | *Number* | *Mean (%)* | *Std. dev. (%)* | *Skewness* | *Kurtosis* | *T-test for the equality of means (unequal variances)* | *F-test for the equality of two variances* |
|---|---|---|---|---|---|---|---|---|
| **SPY** | | | | | | | | |
| **Daytime** | Monday | 855 | -0.0155 | 1.1276 | -0.741 | 12.456 | -1.517 | 2.516*** |
| | Tuesday | 929 | 0.0104 | 1.1432 | 0.387 | 10.652 | -1.201 | 2.933*** |
| | Wednesday | 930 | 0.0020 | 1.0894 | 0.179 | 11.439 | -0.324 | 3.286*** |
| | Thursday | 912 | -0.0131 | 1.0829 | -0.733 | 12.007 | -0.836 | 2.935*** |
| | Friday | 908 | -0.0316 | 1.0050 | 0.102 | 5.340 | -1.335 | 1.690*** |
| | **total** | 4534 | -0.0093 | 1.0901 | -0.152 | 10.788 | -2.343** | 2.579*** |
| **Overnight** | Monday | 854 | 0.0536 | 0.7108 | 0.580 | 12.648 | --- | --- |
| | Tuesday | 929 | 0.0626 | 0.6674 | 0.217 | 11.227 | --- | --- |
| | Wednesday | 930 | 0.0152 | 0.6010 | -0.659 | 7.565 | --- | --- |
| | Thursday | 912 | 0.0215 | 0.6321 | -0.292 | 6.870 | --- | --- |
| | Friday | 908 | 0.0244 | 0.7729 | -1.870 | 25.827 | --- | --- |
| | **total** | 4533 | 0.0353 | 0.6787 | -0.522 | 15.765 | --- | --- |
| **DIA** | | | | | | | | |
| **Daytime** | Monday | 755 | 0.0194 | 1.0660 | -0.282 | 13.301 | -0.262 | 2.407*** |
| | Tuesday | 821 | 0.0018 | 1.0383 | 0.227 | 10.366 | -1.206 | 2.770*** |
| | Wednesday | 825 | 0.0096 | 1.0472 | 0.062 | 12.636 | -0.130 | 3.357*** |
| | Thursday | 810 | 0.0107 | 1.0732 | -0.677 | 12.893 | 0.110 | 3.008*** |
| | Friday | 804 | -0.0322 | 0.9701 | 0.182 | 5.662 | -0.423 | 1.751*** |
| | total | 4015 | 0.0017 | 1.0390 | -0.115 | 11.354 | -0.853 | 2.567*** |
| **Overnight** | Monday | 754 | 0.0315 | 0.6870 | 0.427 | 10.7330 | --- | --- |
| | Tuesday | 821 | 0.0528 | 0.6238 | 0.622 | 11.736 | --- | --- |
| | Wednesday | 825 | 0.0150 | 0.5715 | -0.542 | 6.591 | --- | --- |
| | Thursday | 810 | 0.0059 | 0.6188 | -0.120 | 7.583 | --- | --- |
| | Friday | 804 | -0.0140 | 0.7331 | -2.314 | 34.767 | --- | --- |
| | **total** | 4014 | 0.0181 | 0.6484 | -0.565 | 18.207 | --- | --- |
| **QQQ** | | | | | | | | |
| **Daytime** | Monday | 700 | -0.0462 | 1.6314 | 0.485 | 12.277 | -1.591 | 1.753*** |
| | Tuesday | 762 | -0.0886 | 1.8076 | -0.638 | 6.803 | -2.113** | 4.706*** |
| | Wednesday | 765 | 0.0048 | 1.8635 | 0.903 | 17.47 | -0.220 | 3.841*** |
| | Thursday | 752 | 0.0396 | 1.6841 | -0.099 | 6.178 | -0.559 | 3.315*** |
| | Friday | 749 | -0.0994 | 1.5386 | 0.108 | 7.850 | -2.130** | 2.124*** |
| | **total** | 3728 | -0.0377 | 1.7110 | 0.165 | 10.899 | -2.903*** | 1.818*** |
| **Overnight** | Monday | 699 | 0.0667 | 0.9303 | 0.523 | 9.915 | --- | --- |
| | Tuesday | 762 | 0.0637 | 0.8332 | -0.128 | 9.853 | --- | --- |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Wednesday | 765 | 0.0215 | 0.9507 | 0.198 | 9.853 | --- | --- |
| Thursday | 752 | 0.0788 | 0.9249 | 0.688 | 7.757 | --- | --- |
| Friday | 749 | 0.0457 | 1.0556 | -0.314 | 12.262 | --- | --- |
| **total** | 3727 | 0.0550 | 0.9411 | 0.157 | 10.391 | --- | --- |
| **IWM** | | | | | | | |
| **Daytime** Monday | 615 | -0.0906 | 1.3703 | -0.861 | 7.821 | -2.220** | 2.778*** |
| Tuesday | 669 | 0.0260 | 1.4458 | -0.221 | 6.891 | -0.898 | 3.237*** |
| Wednesday | 671 | 0.0365 | 1.4429 | 0.028 | 8.193 | 0.534 | 3.977*** |
| Thursday | 658 | -0.0286 | 1.4335 | -0.470 | 9.795 | -1.118 | 3.690*** |
| Friday | 657 | -0.0083 | 1.3171 | 0.587 | 9.160 | -0.462 | 1.992*** |
| **total** | 3270 | -0.0116 | 1.4034 | -0.193 | 8.397 | -1.832* | 3.013*** |
| **Overnight** Monday | 614 | 0.0525 | 0.8221 | 0.219 | 8.433 | --- | --- |
| Tuesday | 669 | 0.0835 | 0.8035 | 0.907 | 19.600 | --- | --- |
| Wednesday | 671 | 0.0032 | 0.7235 | -0.247 | 7.538 | --- | --- |
| Thursday | 658 | 0.0418 | 0.7462 | -0.035 | 7.185 | --- | --- |
| Friday | 657 | 0.0207 | 0.9331 | -0.279 | 31.131 | --- | --- |
| **total** | 3269 | 0.0402 | 0.8084 | 0.097 | 18.773 | --- | --- |

*Notes:* Night and daytime returns time series are from 02/01/1996 to 03/01/2014for SPY, 11/03/1999 to 03/01/2014 for QQQ, 21/01/1998 to 03/01/2014 for DIA and from 02/01/2001 to 03/01/2014 for IWM. Number - number of observations in each category. The test for equality of means is $t = (\bar{x}_1 - \bar{x}_2)/((s_1^2/n_1) + (s_2^2/n_2))^{0,5}$ (unequal variances) where $n_1$ and $n_2$ are the sample sizes, $\overline{x}_1$ and $\overline{x}_2$ are the sample means, $S_1^2$ and $S_2^2$ are the samples variances, that follow the t-distribution with $\upsilon$ degrees of freedom where $\upsilon = ((S_1^2/n_1 + S_2^2/n_2)^2)/\left(\left((S_1^2/n_1)^2/(n_1-1)\right)+\left((S_2^2/n_2)^2/(n_2-1)\right)\right)$. The test for equality of variances is the two-tailed F test $F = S_1^2/S_2^2$, where the null of equal variances is rejected if $F < F_{1-(\alpha/2);n_1-1;n_2-1}$ or $F > F_{\alpha/2;n_1-1;n_2-1}$. *, **, *** denote values that are statistically significant at the 10, 5 and 1% levels, respectively.

**Table 2 Estimated Coefficients and Standard Errors on Trading and Non-Trading Period Returns of Day-of-the-Week, Holiday and Turn-of-the-Month Effect Model**

| *ETF* | *SPY* | | *DIA* | | *QQQ* | | *IWM* | |
|---|---|---|---|---|---|---|---|---|
| | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* |
| ($\beta_0$) Constant | 0.0091 | 0.0200 | 0.0083 | 0.0200 | 0.0122 | 0.0352 | -0.0035 | 0.0283 |
| ($\beta_1$) Monday-night | 0.0328 | 0.0316 | 0.0070 | 0.0320 | 0.0336 | 0.0485 | 0.0401 | 0.0410 |
| ($\alpha_1$) Monday-day | -0.0416 | 0.0431 | 0.0032 | 0.0438 | -0.0626 | 0.0725 | -0.0885 | 0.0601 |
| ($\beta_2$) Tuesday-night | 0.0498 | 0.0311 | 0.0360 | 0.0308 | 0.0466 | 0.0466 | **0.0830*** | 0.0444 |
| ($\alpha_2$) Tuesday-day | -0.0095 | 0.0427 | -0.0115 | 0.0428 | -0.1093 | 0.0723 | 0.0194 | 0.0631 |
| Wednesday-night | --------- | ---------- | --------- | --------- | --------- | --------- | --------- | --------- |
| ($\alpha_3$) Wednesday-day | -0.0113 | 0.0421 | -0.0005 | 0.0416 | -0.0114 | 0.0796 | 0.0403 | 0.0630 |
| ($\beta_3$) Thursday-night | 0.0055 | 0.0303 | -0.0093 | 0.0309 | 0.0568 | 0.0517 | 0.0379 | 0.0427 |
| ($\alpha_4$) Thursday-day | -0.0268 | 0.0400 | 0.0005 | 0.0422 | 0.0235 | 0.0676 | -0.0224 | 0.0597 |
| ($\beta_4$) Friday-night | 0.0023 | 0.0327 | -0.0345 | 0.0330 | 0.0195 | 0.0528 | 0.0143 | 0.0469 |
| ($\alpha_5$) Friday-day | -0.0401 | 0.0421 | -0.0355 | 0.0421 | **-0.1173*** | 0.0708 | -0.0002 | 0.0615 |
| ($\gamma_1$) PreH-night | 0.0575 | 0.0497 | 0.0046 | 0.0490 | -0.0212 | 0.0799 | -0.0161 | 0.0668 |
| ($\delta_1$) PreH-day | 0.0399 | 0.0771 | 0.1025 | 0.0745 | 0.1575 | 0.1304 | 0.0313 | 0.1128 |
| ($\gamma_2$) PostH-night | -0.0232 | 0.0697 | 0.0231 | 0.0662 | -0.0703 | 0.1022 | -0.0468 | 0.0974 |
| ($\delta_2$) PostH-day | 0.0594 | 0.0905 | 0.0315 | 0.0866 | 0.0514 | 0.1693 | 0.1141 | 0.1359 |
| ($\theta_1$) LastD-night | **0.1750**** | 0.0849 | **0.0678*** | 0.0410 | **0.1066*** | 0.0625 | 0.0737 | 0.0604 |
| ($\vartheta_1$) LastD-day | **-0.1272*** | 0.0735 | **-0.1998**** | 0.0778 | -0.1394 | 0.1086 | -0.1201 | 0.0949 |
| ($\theta_2$) FirstD-night | **0.1138**** | 0.0480 | **0.1419***** | 0.0462 | **0.2219***** | 0.0710 | **0.1682*** | 0.0716 |
| ($\vartheta_2$) FirstD-day | **0.1902**** | 0.0823 | 0.1142 | 0.0850 | 0.0326 | 0.1373 | -0.0013 | 0.1327 |
| ($\theta_3$) SeconD-night | -0.0150 | 0.0421 | -0.0218 | 0.0423 | -0.0170 | 0.0728 | -0.0123 | 0.0571 |
| ($\vartheta_3$) SeconD-day | 0.0104 | 0.0716 | 0.0327 | 0.0690 | 0.0617 | 0.1573 | 0.0347 | 0.1156 |
| $T$ | 9067 | | 8029 | | 7455 | | 6539 | |
| $R^2$ | 0.0031 | | 0.0033 | | 0.0031 | | 0.0025 | |
| Adjusted $R^2$ | 0.0010 | | 0.0009 | | 0.0006 | | -0.0003 | |
| $F$ - statistic | 1.4689* | | 1.394 | | 1.2517 | | 0.878 | |
| Breusch-Godfrey | 35.933*** | | 52.775*** | | 41.51*** | | 37.36*** | |
| White | 180.08*** | | 155.56*** | | 242.27*** | | 200.64*** | |
| A. Dickey-Fuller test | - 45.024*** | | -35.967*** | | -34.412*** | | -31.632*** | |
| Philips-Perron test | -97.433*** | | -89.418*** | | -85.858*** | | -81.469*** | |

*Notes:* The dummy variable reference in regressions is the Wednesday night return. See data section for sample periods of ETFs. F-statistic is of the null hypothesis that all slope coefficients are zero. Breusch-Godfrey test is for serial correlation in least squares residuals; White test is for heteroscedasticity in least squares residuals; A. Dickey-Fuller and Philips-Perron tests is for unit root tests in returns; all standard errors and p-values are corrected for heteroscedasticity and autocorrelation of unknown form using the Newey-West covariance matrix. The Monday night return is calculated from Friday-close to Monday-open; Friday night return is calculated from Thursday-close to Friday-open; the night return of the last day of the month is calculated from close price of the penultimate day to open price of the last day of the month. The estimated model is $r_t = \beta_0 + \sum_{j=1}^{4}\beta_j x_{j,t} + \sum_{k=1}^{5}\alpha_k x_{k,t} + \sum_{l=1}^{2}\gamma_l h_{l,t} + \sum_{m=1}^{2}\delta_m h_{m,t} + \sum_{n=1}^{3}\theta_n y_{n,t} + \sum_{p=1}^{3}\vartheta_p y_{p,t} + \varepsilon_t$. *, **, *** denote values that are statistically significant at the 10, 5 and 1% levels, respectively.

## 4.2 Regression-Based Results

The estimated coefficients and standard errors of the specified parameters in equation (2) are presented, for each ETF, in Table 2. At the end of the table are presented the $R^2$, the adjusted $R^2$, an F-test of the null hypothesis that all coefficients are jointly zero, the Breusch-Godfrey, the White and the Augmented Dickey-Fuller and Philips–Perron' statistics. The Augmented Dickey-Fuller and Philips-Perron' tests reject the null hypothesis of a unit root in the return series, suggesting that all return series are stationary and suitable for a regression-based analysis.

The Breusch-Godfrey Lagrange multiplier test is used to test for high-order serial correlation in the least squares residuals. The null hypothesis of no serial correlation is rejected for all ETFs suggesting the presence of high-order serial correlation in the residuals. The White test is used to test for heteroscedasticity in the residuals and the null hypothesis of no heteroscedasticity is also rejected for all ETFs return series. Accordingly, the standard errors of parameters are estimated using the Newey-West procedure to correct for the effects of heteroscedasticity and autocorrelation in the residuals.

Consider the combined calendar effect regression models for the SPY and DIA. In no case are the estimated coefficients for the days of the week, by trading and non-trading periods, significant. For the QQQ (IWM), only the Friday daytime (Tuesday night) return is negative (positive), but marginally significant. These results are in line with those obtained in the tests for mean return differences in Table 1. For the pre- and post-holidays effects, and across ETFs, in no case are there significant estimated coefficients during the respective trading and non-trading periods. These results appear to be consistent with those obtained by Chong *et al*. (2005) in their analysis of the US stock market using daily returns, where the prevailing evidence points to the disappearance of this effect after 1997.

In terms of the turn-of-the-month days, and across ETFs, the estimated coefficients for trading and non-trading periods of the last and the first trading day of the following month are significant. For SPY, return coefficients for the non-trading period of the last and first trading days of the month are significantly positive at the 0.05 level. The daytime return of the first (last) day of the month is significantly positive (negative) at the 0.05 (0.10) level. At first glance, these results appear to be consistent with those obtained in previous studies using close-to-close daily returns on the US equity market (McConnell and Xu, 2008). These authors find that this effect is persistent in US equity market in the period 1926 to 2005. However, in the present study, significant coefficients do not extend to the second trading day of the month.

For the DIA´s model, the significant estimated coefficients are almost identical to those observed for the SPY's model, reflecting the strong correlation between these two ETFs. However, the estimated coefficient for the trading period of the first day of the month is no longer significant. As expected, these results reflect the strong common behavior pattern of these two ETFs.

For the QQQ's regression model, only the estimated coefficients for the non-trading period of the last and the first trading days of the month are significant and positive. For the regression model of the IWM, only the estimated coefficients on the

Tuesday night and the night return of the first trading day of the month are positive, though marginally significant.

Across ETFs' combined calendar effect models, only the estimated coefficients for the night and daytime returns of the last and first trading days of the month are significant, with the common pattern of the last and first days' night returns being positive and the first day's daytime return being negative, though only significant on SPY and DIA. Additionally, in all four models, the null hypothesis of joint insignificance is not rejected in three out of the four models, with joint insignificance rejected for the SPY model but at the 0.10 level.

Thus, results of the estimated coefficients of the combined calendar models for the four ETFs do not offer empirical support for the existence of day-of-the-week and pre- and post-holidays effects during trading and non-trading daily periods. For the turn-of-the-month days, some evidence is exhibited but not entirely consistent across the four ETFs calendar models. In interpreting summary statistics of the estimated models, it is relevant to note the low $R^2$ values. These results are expected in regressions that intend to model daily returns and even more so in regressions modeling night and daytime daily returns. As exhibited in Table 1, this seems to be due to the large variance of the dependent variable, i.e., a higher variation coefficient.

### 4.3 Multiple Structural Change Results

In the previous section, and at first glance, the significant coefficients of the combined calendar models would appear to offer some evidence for a single calendar effect in the turn-of-the-month days during the night and daytime periods. However, the low $R^2$ values of the regressions and the insignificant joint F-test could also indicate that the significant coefficients could be the result of occasional or time period-dependent momentary effects. On the other hand, some earlier studies showed evidence that the calendar effects over time may be subject to changes in trading procedures due to regulatory developments, changes in informational market efficiency and other exogenous phenomena to market activity. Thus, in order to examine the long run stability in the estimated models' parameters, we initially conducted cumulative sum of square residual (SSR) tests. Graphs of the cumulative SSR, however, suggest breaks on the return series.

To assess the stability of the significant coefficients, we run multiple structural change tests using Bai and Perron's (1998, 2003) method. This procedure is applied to the equation model (2) admitting a pure structural change model where all parameters are allowed to vary across multiple regimes. In this method, the break dates and the model coefficients across the various regimes are the parameters of interest to be estimated. Table 3 present results of the stability tests on parameters over time. Based on this procedure, the modeling strategy of the return series over calendar effects is similar to that performed in the previous section, i.e., we apply the Bai and Perron (1998)' estimation and test procedures to the model (2). The aim is to test for the existence of multiple structural changes in the mean level of the return series across the various calendar effects. In the procedure specification, we set the trimming $\varepsilon = 0.05$, the maximum permitted number of breaks is set at $M = 5$ and the various tests uses a 5% significance level.

**Table 3 Multiple Structural Change Test Results Across ETFs**

| ETF | ***SPY*** | ***DIA*** | ***QQQ*** | ***IWM*** |
|---|---|---|---|---|
| Specification | $q = all\ dummy\ variables$ | | $M = 5$ | $\varepsilon = 0.05$ |
| $h$ (observations) | 469 | 401 | 373 | 327 |
| $Sup\ F_T(1)$ | 42.248** | 49.458** | 47.347** | 24.877 |
| $Sup\ F_T(2)$ | 28.874** | 31.719** | 56.960 ** | 37,295 ** |
| $Sup\ F_T(3)$ | 33.513 ** | 35.770** | 67.755** | 47.297** |
| $Sup\ F_T(4)$ | 34.726** | 39.138 ** | 59.126 ** | 50.875** |
| $Sup\ F_T(5)$ | 37.190 ** | 38.382** | 54.816 ** | 53.016** |
| $UDmax\ F_T$ | 42.24 ** | 49.458** | 67.755** | 53.016** |
| $WDmax\ F_T$ | 46.656** | 49.458** | 78.501 ** | 66.510** |
| $Sup\ F_T(1\|0)$ | 42.248 ** | 49.966 ** | 47.347 ** | 27. 877 |
| $Sup\ F_T(2\|1)$ | 45.787** | 44.872 ** | 26.414 | --- |
| $Sup\ F_T(3\|2)$ | 37.854** | 50.728** | --- | --- |
| $Sup\ F_T(4\|3)$ | 43.433** | 22.518 | --- | --- |
| $Sup\ F_T(5\|4)$ | 31.679 | --- | --- | --- |
| Selected breaks | 4 | 3 | 1 | 0 |

*Notes:* This table presents statistics' values in the equations (6), (7) and (8) for the pure structural change model in equation (2). $\varepsilon = 0.05$ stands for the trimming coefficient, $M = 5$ stands for the maximum number of allowed break dates in the model; $q$ stands for the number of regressors with allowed changing coefficients and $h$ stands for the minimum number of observations in each regime, ** denotes values that are statistically significant at the 5 level.

Since in the previous section tests results on residuals showed that return series exhibited serial correlation and heteroscedasticity, in estimating the models we allow for different distributions of the errors across the various regimes and the estimated standard errors and p-values are corrected for serial correlation and heteroscedasticity using the Newey-West covariance matrix.

For the SPY, the $Sup\ F_T(k)$ tests are all significant for k between 1 and 5. The $UDmax\ F_T$ and the $WDmax\ F_T$ tests are also significant, suggesting that at least one break is present. To determine the number of breaks, the sequential procedure is applied using a 5% level which selects four breaks (the $Sup\ F_T(5|4)$ test is not significant). The four break dates are estimated at 12/12/1996, at 05/02/1998, at 01/27/1999 and at 04/16/2001. For the DIA, the $Sup\ F_T(k)$ tests are all significant for k between 1 and 5. The $UDmax\ F_T$ and $WDmax\ F_T$ tests are also significant, suggesting that at least one break is present. The selection procedure selects three breaks, with the break dates estimated at 01/11/1999, at 01/28/2000 and at 03/12/2001. For the QQQ, the $Sup\ F_T(k)$ tests are all significant for k between 1 and 5. The $UDmax\ F_T$ and the $WDmax\ F_T$ tests are also significant suggesting that at least one break is present. The selection procedure selects one break, with the break date estimated at 09/20/2001. For the IWM, the $Sup\ F_T(k)$ tests are all significant for k between 2 and 5. The $UDmax\ F_T$ and the $WDmax\ F_T$ tests are also significant but the selection procedure do not select any break, suggesting no structural change in the mean level of returns over the entire sample period in this ETF.

For the SPY, estimated calendar models across the various identified regimes are presented in Table 4. Across the first four regimes, from 1/02/1996 to 4/16/2001, there is a high variability in the significant coefficients. Each one of the first three identified sub-periods comprise about one year and the fourth about two years of daily trading and non-trading period returns. In the first four regimes, the null hypothesis of joint insignificance of parameters is rejected at the 5% level in the first and second and at the 1% level in the third and fourth. In the fifth regime, the longest, the null hypothesis of joint insignificance cannot be rejected. These results suggest that the significant calendar effects, evidenced during the first four sub-time periods, would be motivated by momentary effects generated by specific market conditions during these sub periods. For SPY, these results suggest that persistent and significant day-of-the-week, pre- and post-holidays and turn-of-the-month effects do not exist. Thus, these results would not support any profitable trading strategy based on these calendar effects. The estimated results for the fifth and last regime, from 04/17/2001 to 01/03/2014, despite the estimated coefficient for the night return of the first trading day of the month being marginally positive, suggest the disappearance of calendar effects observed in the previous sub–periods.

For the DIA, results of the estimated models for the various regimes are shown in Table 5. As in the SPY case, there is also in the ETF a strong variability in the estimated coefficients in the first three regimes. Each one of the first three regimes comprises approximately 1 year of intraday returns. The fourth and last regime coincides with that observed for SPY. In this last regime, the estimated coefficient for the non-trading period of the first trading day of the month is significantly positive. In the first three regimes, the null hypothesis of joint insignificance is rejected. This suggest that the estimated calendar effects have some explanatory power in returns but that these effects are also specific and restricted to these short sub-periods, not allowing delineating profitable trading strategies. Similarly, in the fourth regime, we observe the decline and the disappearance of calendar effects in this ETF, despite the estimated coefficient for the night return of the first trading day of the month being significantly positive. Similarly, the null hypothesis of joint insignificance for this regime cannot be rejected.

**Table 4 Estimated Coefficients and Standard Errors on the Trading and Non-Trading Period Returns on Day-of-the-week, Pre- and Post-Holiday and Turn-of-the-month Effect Model for Multiple Structural Changes for SPY ETF**

| ***SPY ETF*** | ***1/02/1996 – 12/12/1996*** | | ***12/12/1996-5/02/1998*** | | ***2/06/1998-1/27/1999*** | | ***1/28/1999-4/16/2001*** | | ***4/17/2001-1/03/2014*** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* |
| ($\beta_0$) Constant | 0.0500 | 0.0702 | 0.0609 | 0.0558 | 0.1219 | 0.1018 | 0.0013 | 0.0563 | -0.0068 | 0.0244 |
| ($\beta_1$) Monday-night | 0.0129 | 0.0918 | 0.1406 | 0.0889 | -0.0929 | 0.1634 | 0.1042 | 0.0783 | 0.0190 | 0.0392 |
| ($\alpha_1$) Monday-day | -0.0106 | 0.1528 | -0.1500 | 0.1901 | -0.3127 | 0.2130 | 0.1093 | 0.1264 | -0.0363 | 0.0510 |
| ($\beta_2$) Tuesday-night | 0.1100 | 0.0870 | 0.0098 | 0.1068 | 0.1398 | 0.1357 | 0.0696 | 0.0798 | 0.0392 | 0.0384 |
| ($\alpha_2$) Tuesday-day | **-0.2021*** | 0.1152 | **0.3739*** | 0.2099 | -0.0450 | 0.1845 | **-0.2805**** | 0.1299 | 0.0210 | 0.0501 |
| Wednesd-night | --------- | --------- | --------- | --------- | --------- | --------- | --------- | --------- | --------- | --------- |
| ($\alpha_3$) Wednesd-day | -0.0149 | 0.1497 | -0.2049 | 0.1341 | -0.0528 | 0.1923 | -0.1781 | 0.1327 | 0.0398 | 0.0494 |
| ($\beta_3$) Thursday-night | -0.0358 | 0.0913 | 0.0103 | 0.0882 | **-0.449***** | 0.151 | 0.1394 | 0.0859 | 0.0177 | 0.0367 |
| ($\alpha_4$) Thursday-day | -0.0189 | 0.1197 | **-0.3683***** | 0.1262 | -0.0429 | 0.2390 | 0.0572 | 0.1196 | -0.0070 | 0.0475 |
| ($\beta_4$) Friday-night | 0.0049 | 0.1274 | 0.0011 | 0.1198 | 0.0915 | 0.1331 | **0.1725*** | 0.0940 | -0.0320 | 0.0391 |
| ($\alpha_5$) Friday-day | 0.0438 | 0.1296 | -0.0928 | 0.1746 | 0.1314 | 0.1928 | **-0.3556**** | 0.1440 | -0.0026 | 0.0475 |
| ($\gamma_1$) PreH-night | -0.0516 | 0.0802 | **0.4098**** | 0.1634 | **0.2721*** | 0.159 | 0.1329 | 0.1904 | 0.0114 | 0.0561 |
| ($\delta_1$) PreH-day | **-0.3163*** | 0.1614 | -0.3750 | 0.3377 | -0.0449 | 0.3774 | 0.2207 | 0.2871 | 0.0766 | 0.0859 |
| ($\gamma_2$) PostH-night | **-1.0121**** | 0.3989 | 0.1256 | 0.1253 | **0.6070**** | 0.286 | 0.0050 | 0.1827 | -0.0266 | 0.0805 |
| ($\delta_2$) PostH-day | 0.1296 | 0.3499 | -0.0209 | 0.3051 | -0.1611 | 0.4137 | 0.0758 | 0.2764 | 0.0878 | 0.1072 |
| ($\theta_1$) LastD-night | 0.1999 | 0.2561 | -0.0634 | 0.4272 | **1.5071**** | 0.673 | -0.0182 | 0.2710 | 0.1281 | 0.0849 |
| ($\vartheta_1$) LastD-day | -0.0923 | 0.2057 | -0.0977 | 0.3066 | **-1.5087**** | 0.661 | 0.1482 | 0.2517 | -0.0713 | 0.0682 |
| ($\theta_2$) FirstD-night | 0.1473 | 0.1423 | **0.2901*** | 0.1524 | -0.0756 | 0.1737 | 0.1272 | 0.0918 | **0.1125*** | 0.0612 |
| ($\vartheta_2$) FirstD-day | **0.5934***** | 0.2025 | **0.5215**** | 0.2382 | 0.6164 | 0.4276 | 0.0413 | 0.1969 | 0.1122 | 0.1014 |
| ($\theta_3$) SeconD-night | 0.0797 | 0.1421 | -0.0687 | 0.0907 | -0.1001 | 0.1239 | -0.1057 | 0.1614 | 0.0063 | 0.0502 |
| ($\vartheta_3$) SeconD-day | -0.0601 | 0.1927 | 0.1097 | 0.2574 | -0.1804 | 0.5051 | -0.1311 | 0.3017 | 0.0398 | 0.0731 |
| $R^2$ | 0.0716 | | 0.0607 | | 0.0899 | | 0.0358 | | 0.0019 | |
| Adjusted $R^2$ | 0.0335 | | 0.0288 | | 0.0532 | | 0.0191 | | -0.0009 | |
| $F$ - statistic | 1.8796** | | 1.9033** | | 2.4463*** | | 2.1605*** | | 0.6720 | |
| $T$ | 483 | | 580 | | 490 | | 1118 | | 6398 | |

*Notes:* The regression constant term (reference category) is the Wednesday night mean return. F-statistic is of the null hypothesis that all slope coefficients are zero. All standard errors and p-values are corrected for heteroscedasticity and autocorrelation of unknown form using the Newey-West covariance matrix. $T$ stands for sample size, either of night and day time returns, in each regime. The Monday night return is calculated as from Friday-close to Monday-open prices; the Friday night return is calculated as from Thursday-close to Friday-open prices; the night return of the last trading day of the month is calculated as from the close price of the penultimate trading day to open price of the last trading day of the month and the others calendar returns are calculated accordingly. The estimated model is $r_t = \beta_0 + \sum_{j=1}^{4}\beta_j x_{j,t} + \sum_{k=1}^{5}\alpha_k x_{k,t} + \sum_{l=1}^{2}\gamma_l h_{l,t} + \sum_{m=1}^{2}\delta_m h_{m,t} + \sum_{n=1}^{3}\theta_n y_{n,t} + \sum_{p=1}^{3}\vartheta_p y_{p,t} + \varepsilon_t$ where each regression coefficient captures the corresponding mean return difference relative to the constant term. *, **, *** denote values that are statistically significant at the 10, 5 and 1% levels, respectively.

**Table 5 Estimated Coefficients and Standard Errors on Trading and Non-Trading Period Returns of Day-of-the-Week, Holiday and Turn-of-the-Month Effect Model for Multiple Structural Changes for DIA ETF**

| *DIA ETF* | *1/21/1998-1/11/1999* | | *1/12/1999-1/28/2000* | | *1/28/2000-3/12/2001* | | *3/12/2001-1/03/2014* | |
|---|---|---|---|---|---|---|---|---|
| | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* |
| ($\beta_0$) Constant | **0.1721**** | 0.086 | 0.0857 | 0.0831 | -0.0899 | 0.0738 | -0.0019 | 0.022 |
| ($\beta_1$) Monday-night | -0.1115 | 0.1582 | 0.0731 | 0.1098 | 0.1075 | 0.1212 | -0.0001 | 0.0356 |
| ($\alpha_1$) Monday-day | -0.3277 | 0.2193 | -0.1197 | 0.1850 | **0.6655***** | 0.174 | -0.0153 | 0.0467 |
| ($\beta_2$) Tuesday-night | 0.0894 | 0.1312 | -0.1430 | 0.1106 | **0.1753*** | 0.099 | 0.0348 | 0.0353 |
| ($\alpha_2$) Tuesday-day | -0.1071 | 0.1914 | **-0.428***** | 0.142 | 0.0850 | 0.1651 | 0.0223 | 0.0473 |
| Wedn.-night | --------- | --------- | --------- | --------- | --------- | --------- | --------- | --------- |
| ($\alpha_3$) Wedn.-day | -0.1795 | 0.1429 | -0.0771 | 0.1617 | -0.2075 | 0.1661 | 0.0346 | 0.0472 |
| ($\beta_3$) Thursday-night | **-0.5797***** | 0.153 | 0.0866 | 0.1199 | 0.0923 | 0.1152 | 0.0164 | 0.0343 |
| ($\alpha_4$) Thursday-day | -0.0624 | 0.2203 | -0.0984 | 0.1565 | 0.2005 | 0.1841 | -0.0030 | 0.0448 |
| ($\beta_4$) Friday-night | 0.0247 | 0.1274 | 0.1022 | 0.1376 | 0.0032 | 0.1302 | -0.0547 | 0.0362 |
| ($\alpha_5$) Friday-day | 0.1015 | 0.1956 | -0.1558 | 0.1558 | **-0.4767**** | 0.213 | 0.0033 | 0.0451 |
| ($\gamma_1$) PreH-night | 0.2477 | 0.1748 | 0.3792 | 0.2856 | -0.0414 | 0.1948 | -0.0235 | 0.0506 |
| ($\delta_1$) PreH-day | -0.1701 | 0.2559 | **0.534**** | 0.250 | 0.1249 | 0.3985 | 0.0827 | 0.0784 |
| ($\gamma_2$) PostH-night | **0.6553**** | 0.295 | 0.1947 | 0.2098 | -0.0097 | 0.2682 | -0.0245 | 0.0712 |
| ($\delta_2$) PostH-day | -0.2649 | 0.4366 | -0.1555 | 0.2511 | 0.2817 | 0.2893 | 0.0482 | 0.0957 |
| ($\theta_1$) LastD-night | 0.0440 | 0.2091 | **0.218*** | 0.115 | 0.2110 | 0.1528 | 0.0499 | 0.0458 |
| ($\vartheta_1$) LastD-day | **-1.5499**** | 0.635 | -0.3191 | 0.3013 | -0.1675 | 0.2922 | -0.0886 | 0.0661 |
| ($\theta_2$) FirstD-night | 0.1996 | 0.2463 | 0.0436 | 0.1534 | **0.2199**** | 0.100 | **0.1375***** | 0.051 |
| ($\vartheta_2$) FirstD-day | 0.5295 | 0.3322 | 0.1715 | 0.3403 | 0.2635 | 0.2362 | 0.0629 | 0.0969 |
| ($\theta_3$) SeconD-night | -0.1381 | 0.1455 | -0.0519 | 0.2118 | -0.0869 | 0.2070 | -0.0020 | 0.0454 |
| ($\vartheta_3$) SeconD-day | -0.0054 | 0.4718 | -0.0293 | 0.2357 | 0.3720 | 0.3477 | 0.0186 | 0.0669 |
| $R^2$ | 0.1013 | | 0.0615 | | 0.0912 | | 0.0020 | |
| Adjusted $R^2$ | 0.0652 | | 0.0264 | | 0.0594 | | -0.0008 | |
| $F$ - statistic | 2.8027*** | | 1.7555** | | 2.865*** | | 0.703 | |
| $T$ | 492 | | 529 | | 562 | | 6446 | |

*Notes*: The regression constant term (reference category) is the Wednesday night mean return. F-statistic is of the null hypothesis that all slope coefficients are zero. All standard errors and p-values are corrected for heteroscedasticity and autocorrelation of unknown form using the Newey-West covariance matrix. $T$ stands for sample size, either of night and day time returns, in each regime. The Monday night return is calculated as from Friday-close to Monday-open prices; the Friday night return is calculated as from Thursday-close to Friday-open prices; the night return of the last trading day of the month is calculated as from the close price of the penultimate trading day to open price of the last trading day of the month and the others calendar returns are calculated accordingly. The estimated model is $r_t = \beta_0 + \sum_{j=1}^{4} \beta_j x_{j,t} + \sum_{k=1}^{5} \alpha_k x_{k,t} + \sum_{l=1}^{2} \gamma_l h_{l,t} + \sum_{m=1}^{2} \delta_m h_{m,t} + \sum_{n=1}^{3} \theta_n y_{n,t} + \sum_{p=1}^{3} \vartheta_p y_{p,t} + \varepsilon_t$ where each regression coefficient captures the corresponding mean return difference relative to the constant term. *, **, *** denote values that are statistically significant at the 10, 5 and 1% levels, respectively.

For the QQQ, estimated models are presented in Table 6. For this ETF, Bai and Perron (1998)'s method only identifies two regimes over the entire sample period. The first regime encompasses about two and a half years of night and daytime returns and the significance and sign of the estimated coefficients in this regime are identical to those exhibited in Table 2 for this ETF over the entire sample period, except in the coefficient for the Tuesday daytime return. However, the null hypothesis of joint insignificance of the estimated coefficients is not rejected. The second regime spans from mid-September 2001 until the end of the entire sample period. In this regime, and similarly to that observed in the full sample period, the estimated coefficient for the night return of the first trading day of the month is significantly positive. However, this significant effect, along with the significantly negative effect of the daytime return of the last trading day of the month, reveal no significant power in explaining the trading and non-trading period return variance since the null of joint insignificance of the coefficients fails to be rejected.

**Table 6 Estimated Coefficients and Standard Errors on Trading and Non-Trading Period Returns of Day-of-the-Week, Holiday and Turn-of-the-Month Effect Model for Multiple Structural Changes for the QQQ ETF**

| ***ETF QQQ*** | ***3/11/1999 - 9/20/2001*** | | ***9/20/2001-1/03/2014*** | |
|---|---|---|---|---|
| | *Coeff.* | *Std.error* | *Coeff.* | *Std.error* |
| ($\beta_0$) Constant | -0.0688 | 0.1342 | 0.0277 | 0.0316 |
| ($\beta_1$) Monday-night | 0.1760 | 0.1743 | 0.0070 | 0.0454 |
| ($\alpha_1$) Monday-day | -0.0145 | 0.2781 | -0.0694 | 0.0658 |
| ($\beta_2$) Tuesday-night | 0.2362 | 0.1682 | 0.0075 | 0.0446 |
| ($\alpha_2$) Tuesday-day | **-0.5534*** | 0.2875 | -0.0157 | 0.0651 |
| Wedn.-night | --------- | --------- | --------- | --------- |
| ($\alpha_3$) Wedn.-day | -0.1273 | 0.3146 | 0.0129 | 0.0696 |
| ($\beta_3$) Thursday-night | 0.2999 | 0.2046 | 0.0074 | 0.0454 |
| ($\alpha_4$) Thursday-day | 0.1564 | 0.2681 | -0.0025 | 0.0608 |
| ($\beta_4$) Friday-night | 0.2770 | 0.1965 | -0.0331 | 0.0490 |
| ($\alpha_5$) Friday-day | -0.2390 | 0.2878 | -0.0947 | 0.0627 |
| ($\gamma_1$) PreH-night | 0.0556 | 0.3705 | -0.0366 | 0.0632 |
| ($\delta_1$) PreH-day | 0.3489 | 0.6276 | 0.1273 | 0.1080 |
| ($\gamma_2$) PostH-night | -0.3756 | 0.3686 | -0.0012 | 0.0990 |
| ($\delta_2$) PostH-day | -0.6815 | 0.7184 | 0.2024 | 0.1429 |
| ($\theta_1$) LastD-night | **0.5814***** | 0.1945 | 0.0091 | 0.0590 |
| ($\vartheta_1$) LastD-day | 0.1514 | 0.4759 | **-0.2001**** | 0.0873 |
| ($\theta_2$) FirstD-night | **0.4862*** | 0.2587 | **0.1648**** | 0.0658 |
| ($\vartheta_2$) FirstD-day | -0.1353 | 0.4638 | 0.0580 | 0.1327 |
| ($\theta_3$) SeconD-night | 0.0060 | 0.3182 | -0.0157 | 0.0594 |
| ($\vartheta_3$) SeconD-day | -0.0027 | 0.7904 | 0.0782 | 0.1010 |
| $R^2$ | 0.0206 | | 0.0031 | |
| Adjusted $R^2$ | 0.0057 | | 0.0001 | |
| $F$ - statistic | 1.3846 | | 1.0336 | |
| $T$ | 1268 | | 6188 | |

*Notes:* The regression constant term (reference category) is the Wednesday night mean return. F-statistic is of the null hypothesis that all slope coefficients are zero. All standard errors and p-values are corrected for heteroscedasticity and autocorrelation of unknown form using the Newey-West covariance matrix. $T$ stands for sample size, either of night and day time returns, in each regime. The Monday night return is calculated as from Friday-close to Monday-open prices; the Friday night return is calculated as from Thursday-close to Friday-open prices; the night return of the last trading day of the month is calculated as from the close price of the penultimate trading day to open price of the last trading day of the month and the others calendar returns are calculated accordingly. The estimated model is $r_t = \beta_0 + \sum_{j=1}^{4}\beta_j x_{j,t} + \sum_{k=1}^{5}\alpha_k x_{k,t} + \sum_{l=1}^{2}\gamma_l h_{l,t} + \sum_{m=1}^{2}\delta_m h_{m,t} + \sum_{n=1}^{3}\theta_n y_{n,t} + \sum_{p=1}^{3}\vartheta_p y_{p,t} + \varepsilon_t$ where each regression coefficient captures the corresponding mean return difference relative to the constant term. *, **, *** denote values that are statistically significant at the 10, 5 and 1% levels, respectively.

Finally, for the IWM, the Bai and Perron (1998)'s method does not identify different regimes in the calendar effects on the trading and non-trading period returns over their entire sample period. The absence of different regimes of calendar effects in this ETF is not surprising given that their sample period begins in 06/02/2001, approximately coinciding with the beginning of the last identified regime in the previous three ETFs and where the calendar effects on trading and non-trading period returns tend to disappear. In this latter sub-period and across ETFs, the exception is the positive and significant effect on the night return of the first trading day of the month.

Results suggest that the calendar effects on trading and non-trading daily period returns in US stock markets, using as a proxy returns in the US equity ETFs of the major US stock indices, exhibit a strong variability in the significant coefficients in the short sub-samples until the beginning of 2001. After this date, significant calendar effects tend to disappear, only remaining a significant positive effect on the night return of the first trading day of the month. Also, across ETFs and for this last sub-period, the null of joint insignificance of the parameters in the model fails to be rejected. Thus, in the US equity ETF market, results do not support the existence of calendar (day-of-the-week, pre- and post-holiday and turn-of-the-month) effects on trading and non-trading daily returns, in particular from 2001 onwards.

## 5. Conclusion

This study examines the presence of calendar effects on trading and non-trading daily period returns in the US equity market, using return series of the four major Exchange Traded Funds (ETFs) actively traded and that track major stock market indices in US markets: the SPY (S&P 500 index), the DIA (Dow Jones Industrial Average index), the QQQ (NASDAQ 100 index) and the IWM (Russell 2000 index). The regression model simultaneously specifies three calendar effects in night and daytime returns: by day-of-the-week, by pre- and post-holidays and by turn-of-the-month days' effects. The specified model allows us to examine the relative strength of each effect.

Obtained results during the full sample period are not consistent with those observed in previous studies using these same ETFs, although using other methods and shorter and earlier sample periods. Specifically, almost all mean return differences between trading and non-trading periods, by days-of-the-week and across the set of ETFs, are not statistically significant. On the other hand, by days-of-the-week, but consistent with previous studies, we continue to observe the pervasive fact that the volatility of the trading is significantly higher than the volatility of the non-trading daily period return.

Over the entire sample period, regression results for the combined calendar effect model only exhibit significant coefficients on the last and the first trading day of the month for the SPY and DIA and in a smaller number in the QQQ and IWM ETFs. The night return of the first trading day of the month is the only significant and positive coefficient across all ETFs, though marginally significant in the IWM.

To examine over time the persistence or the fixedness of the calendar effects, we employ Bai and Perron (1998, 2003)'s method to identify multiple structural changes. Across ETFs, results show the existence of several regimes in calendar effects, except for the IWM where only a single one regime is observed over the entire sample period. After identification of the regime break dates over the entire sample period, regression models were re-estimated. For each of these three ETFs, the significant estimated coefficients vary considerably across regimes and with a short duration, from one to two years. A common pattern across ETFs is observed in the last regime, beginning in 2001, and the single one for the IWM. In this last and considerably longer regime, significant effects observed in previous regimes disappear, except for the significant positive effect on the night return of the first trading day of the month which tends to persist.

In summary, before 2001, results exhibit a high instability in significant coefficients. From 2001 onwards, and across ETFs, there is a decrease or even the disappearance of calendar effects in trading and non-trading daily period returns. Results obtained in the present study contrast with those obtained in previous studies by Cliff *et al*. (2008) and Kelly and Clark (2011) using this same ETF group, although using earlier and shorter sample periods, ending at 2006. For the exhibited instability by the significant coefficients across ETFs over the sub-sample periods prior to 2001, we cannot exclude the possibility that these significant and unstable observed effects could have been momentary and motivated by market-specific conditions in such short time periods. On the other hand, and in line with the absence of these effects from 2001 onwards, results are consistent with the nature of this asset class, i.e., these ETFs are broadly diversified portfolios with diversification of private information, with higher liquidity and lower transaction costs (bid-ask spreads); these characteristics mitigate adverse selection, induces uninformed investors to trade these securities (Hasbrouck, 2003) and improves information impounding.

Other factors could have contributed to the absence of calendar effects, namely the growth in futures market, the increased trading by institutional managers in this asset class and the regulatory changes introduced by SEC (2005) in the US market trading microstructure. Hasbrouck (2003) found that for the S&P 500 and NASDAQ 100 indexes, price discovery is dominated by futures, specifically by E-minis futures contracts, and not by ETFs trading. The regulatory changes introduced by SEC (2005) could have contributed to improvements in information impounding at the open and close of the markets and reduced trading activity leeway by specialists in NYSE and market makers in NASDAQ in open and close price discovery.

Our results suggest that, from 2001 onwards, open and close price discovery mechanisms may have become more efficient. Given our inconsistent results with previous studies using this same ETF group, it would be adequate to carry out further studies to examine the robustness of our results using this and other asset groups, with other methods and with an extended sample period.

REFERENCES


Abhyankar A, Ghosh D, Levin E, Limmack R (1997): Bid-Ask Spreads, Trading Volume and Volatility: Intraday Evidence from the London Stock Exchange. *Journal of Business Finance and Accounting*, 24(3&4):343-362.

Abraham A, Ikenberry D (1994): The Individual Investor and the Weekend Effect. *Journal of Financial and Quantitative Analysis*, 29(2):263-277.

Amihud Y (2002): Illiquidity and Stock Returns: Cross-Section and Time Series Effects. *Journal of Financial Markets*, 5:31-56.

Ariel R (1990): High Stock Returns before Holidays: Existence and Evidence on Possible Causes. *Journal of Finance*, 45(5):1611-1626.

Bagnoli M, Clement M, Watts S, (2005): Around-the-Clock Media Coverage and the Timing of Earnings Announcements. Available at SSRN: http://ssrn.com/abstract=570247.

Bai J, Perron P (1998): Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66(1):47-78.

Bai J, Perron P (2003): Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics*, 18:1-22.

Barclay M, Hendershott T (2003): Price Discovery and Trading after Hours. *Review of Financial Studies*, 16(4):1041-1073.

Berkman H, Koch P, Tuttle L, Zhang Y (2012): Paying Attention: Overnight Returns and the Hidden Cost of Buying at the Open. *Journal of Financial and Quantitative Analysis*, 47(4):715-741.

Branch B, Ma A (2006): The Overnight Return, One More Anomaly. SSRN, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=937997.

Brockman P, Michayluk D, (1998): The Persistent Holiday Effect: Additional Evidence. *Applied Economic Letters*, 5(4):205-209.

Cadsby C, Ratner M (1992): Turn-of-the-Month and Pre-Holiday Effects on Stock Returns: Some International Evidence. *Journal of Banking and Finance*, 16(3):497-509.

Chang E, Pinegar J, Ravichandran R (1993): International Evidence on the Robustness of the Day-of-the-Week Effect. *Journal of Financial and Quantitative Analysis*, 28(4):497-513.

Chang E, Pinegar J, Ravichandran R (1998): U.S. Day-of-the-Week Effect and Asymmetric Responses to Macroeconomic News, *Journal of Banking and Finance*, 22(5):513-534.

Chong R, Hudson R, Keasey K, Littler K (2005): Pre-Holiday Effects: International Evidence on the Decline and Reversal of a Stock Market Anomaly. *Journal of International Money and Finance*, 24(8): 1226-1236.

Chow E, Lee Y, Liu Y (2004): Intraday Information, Trading Volume, and Return Volatility: Evidence from the Order Flows on the Taiwan Stock Exchange. *Academia Economic Papers*, 32(1):107-148.

Cliff, MT, Cooper M, Gulen H (2008): Return differences between trading and non-trading hours: like night and day. Available at SSRN 1004081.

Damodaran A (1989): The Weekend Effect in Information Releases: a Study of Earnings and Dividend Announcements. *Review of Financial Studies*, 2(4):607-623.

Doyle J, Magilke M (2009): The Timing of Earnings Announcements: An Examination of the Strategic Disclosure Hypothesis. *The Accounting Review*, 84(1):157-182.

Fama E (1970): Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25(2):383-417.

Foster F, Viswanathan S (1993): Variations in Trading Volume, Return Volatility and Trading Costs: Evidence on Recent Price Formation Models. *Journal of Finance*, 48(1):187-211.

French K (1980): Stock Returns and the Weekend Effect. *Journal of Financial Economics*, 8(1):55-69.

French K, Roll R (1986): Stock Return Variances: the Arrival of Information and the Reaction of Traders. *Journal of Financial Economics*, 17(1):5-26.

George T, Hwang C (2001): Information Flow and Pricing Errors: a Unified Approach to Estimation and Testing. *Review of Financial Studies*, 14(4):979-1020.

Harris L (1989): A Day-End Transaction Price Anomaly. *Journal of Financial and Quantitative Analysis*, 24(1):29-45.

Hasbrouck J (2003): Intraday Price Formation in US Equity Index Markets. *Journal of Finance*, 58(6): 2375-2400.

Hong H, Wang J (2000): Trading and Returns under Periodic Market Closures. *Journal of Finance*, 55(1)297-354.

Jiang C, Likitapiwat T, McInish T (2012): Information Content of Earnings Announcements: Evidence from After-Hours Trading. *Journal of Financial and Quantitative Analysis*, 47(6):1303–1330.

Kelly M, Clark S (2011): Returns in Trading versus Non-Trading Hours: The Difference is Day and Night. *Journal of Asset Management*, 12(2):132-145.

Lachance M (2015): Night Trading: Lower Risk but Higher Returns? San Diego State University, Available at SSRN: https://ssrn.com/abstract=2633476

Lakonishok J, Smidt S (1988): Are Seasonal Anomalies Real? A Ninety-Year Perspective. *Review of Financial Studies*, 1(4):403-425.

Lockwood L, Linn S (1990): An Examination of Stock Market Return Volatility during Overnight and Intraday Periods, 1964-1989. *Journal of Finance*, 45(2):591-601.

Longstaff F (1995): How Much Can Marketability Affect Security Values? *Journal of Finance*, 50(5): 1767-1774.

Lou D, Polk C, Skouras S (2015): A Tug of War: Overnight versus Intraday Expected Returns, London School of Economics, *Working Paper*.

McConnell J, Xu W (2008): Equity Returns at the Turn-of-the-Month. *Financial Analysts Journal*, 64(2):49-64.

McInish H, Wood R (1992): An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks. *Journal of Finance*, 47(2)753-764.

Patell J, Wolfson M (1982): Good News, Bad News, and the Intraday Timing of Corporate Disclosures. *Accounting Review*, 57(3):509-527.

Pearce D (1996): The Robustness of Calendar Anomalies in Daily Stock Returns. *Journal of Economics and Finance*, 20(3):69-80.

Pettengill G (2003): A Survey of the Monday Effect Literature. *Quarterly Journal of Business and Economics,* 42(3/4):3-28.

Qiu, M, Cai T (2013): On Overnight Return Premiums of International Stock Markets, http://www.nzfc.ac.nz/archives/2013/programme/.

Rogalsky R (1984): New Findings Regarding Day-of-the-Week Returns over Trading and Non-Trading periods: a note. *Journal of Finance*, 39(5):1603-1614.

SEC (2005): Securities and Exchange Commission, Regulation NMS: Final Rules and Amendments to Joint Industry Plans. Washington, D.C., 2005. (http://www.sec.gov/rules/final/34-51808.pdf).

Stoll H, Whaley R (1990): Stock Market Structure and Volatility. *Review of Financial Studies*, 3(1): 37-71.

Thaler R (1987): Seasonal Movements in Security Prices II: Weekend, Holiday, Turn-of-the-Month and Intraday Effects. *Journal of Economic Perspectives*, 1(1):169-177.

Tompkins R, Wiener Z (2008): Bad Days and Good Nights: A Re-Examination of Non-Traded and Traded Period Returns. *Working Paper Series*. http://www.ssrn.com.

Wood R, McInish T, Ord J (1985): An Investigation of Transaction Data for NYSE Stocks. *Journal of Finance*, 40(3):723-739.

Zainudin A, Coutts J (1997): A Securities Price Anomalies in the London International Stock Exchange: a 60-Year Perspective. *Applied Financial Economics*, 7(5):455-464.