NIGHT AND DAYTIME EFFECTS IN US EQUITY EXCHANGE-TRADED FUND RETURNS

João Dionísio Monteiro[1] *e José Ramos Pires Manso* [2]

Abstract

This paper examines average returns over periods when markets are open and when markets are closed using a sample of the four major US equity exchange-traded funds (ETFs) in the period January 1996 to January 2014. First, we examine common day and night effects across ETFs. Second, we examine common day and night effects decomposed by day of the week. In the analysis we use panel data regression models with estimators of the standard errors of the estimated parameters robust to various departures of the least square residuals from independent and identically distributed assumptions. In previous studies (Cliff et al., 2008), it was observed the surprising result that returns during the night period are strongly positive and significant and the night minus day return differences are also pervasively positive and significant. Our results show a marked decrease and the disappearance of the night and day effect from 2006. Results show that in this asset class night returns are no longer consistently higher than the day returns, overall and across days of the week. Another puzzling fact in light of the asset pricing models, already evidenced in previous studies, but which tends to remain, is that the volatility of day returns is significantly higher than the volatility of night returns.

KEYWORDS: night and daytime effects, market efficiency, US equity exchange-traded funds

1. INTRODUCTION

In financial markets, information flows continuously throughout twenty-four hours a day but price variations are not continuous due to periodic market closure. Sudden changes in daily transaction regimes, when markets open and close, have important implications for the dynamics of prices in the short term. Several researchers have presented evidence on the impact of periodic market closure in transaction volume, liquidity, volatility and pricing (Jones, Kaul and Lipson, 1994; George and Hwang, 2001). On the other hand, several theoretical papers have sought to model the implications of periodic market closure for equilibrium prices (Foster and Wiswanathan, 1990; Slezak, 1994; Hong and Wang, 2000). However, theoretical models suggest different predictions of the effects of periodic market closure at the first moment of returns.

On one hand, various models proposed in theoretical papers predict lower returns during non-trading periods than in trading periods (Slezak, 1994; Hong and Wang, 2000), a prediction consistent with the evidence documented in the early empirical studies of the weekend effect on returns. On the other hand, other theoretical papers predict higher returns during

[1] jdm@ubi.pt, NECE, University of Beira Interior
[2] pmanso@ubi.pt, NECE, University of Beira Interior

non-trading periods to compensate liquidity providers for bearing additional risk. In this line, the model of Longstaff (1995) predicts higher returns over periods of market closure arising from a liquidity related non-marketability effect.

The purpose of this paper is to add to the body of empirical papers on this theme an empirical analysis of the effects of the periodic market closure on the first moment of return distributions in US equity exchange-traded funds (ETFs). First, we examine whether there exist a common night and day effect across ETFs. Second, we extend the analysis and examine whether there exist a common night and day effect decomposed by day of the week. The analysis is carried out using panel-data regression models with standard errors of the estimated parameters corrected for several departures of the least square residuals from independent and identically distributed assumptions.

Our major findings on the US equity ETFs are: i) No significant overall common and robust night and daytime effects is exhibited in our sample of US ETFs; only a positive and marginally significant night average return is found while the day average return is not significantly different from zero; ii) Concerning common night and day effect, decomposed by day of the week, only the estimated coefficient (average) for Tuesday-night is positive and significantly different from zero; these findings contrast with evidence obtained by previous empirical studies on the same type of assets in US equity market. Cliff *et al.* (2008) found pervasive evidence of a significant difference, first, between overall night and day returns and, then, between night and day returns by day of the week; iii) our results confirm the evidence obtained in previous studies that the volatility is pervasive and significantly higher during the trading periods than during the market closing periods, contradicting the predictions of asset pricing models that the market price of the risk would be positive. These results continue to present a challenge to the asset pricing models in explaining why in diversified portfolios, as are the ETFs, zero average returns (daytime period) have higher volatility than that associated with slightly positive average returns (night period).

The remainder of this paper is organized as follows. Section 2 reviews the literature on the effects of the periodic closure of markets in returns when they are open and closed. Section 3 presents the data and the methodology used in the study. Section 4 present and discusses the results. In section 5 conclusions are presented.

2. LITERATURE REVIEW

In recent decades calendar anomalies have been widely surveyed in empirical papers but most studies used daily closing prices or volume weighted averages prices to identify these seasonality effects. Pearce (1996), for example, for US equity markets, used daily returns from 1972 to 1994 on a variety of portfolios to simultaneously test for weekend effects, post-holiday effects, January effects and serial correlations. Evidence was obtained that the calendar anomalies were more pronounced for smaller than larger firms. The returns for smaller firms were consistently lower after weekends and consistently higher preceding holidays. Thaler (1987) offered various hypotheses that can, partially, explain these abnormal returns, namely, the timing of the arrival of good and bad news. Thaler gives the example of bad news being postponed until after close on Friday. Schwert (2002) found that many market anomalies (size effect, value effect, weekend effect and dividend yield effect) have weakened or disappeared after research articles about them have been published.

A related topic less developed but equally important to the literature of market efficiency is the analysis of regularities in intraday period returns. Although empirical literature is not completely consensual about patterns of intraday returns within this topic of market efficiency, the most striking and surprising result is that various studies obtained evidence that overnight returns are strongly positive and the returns during the day are close to zero and sometimes negative. Wood, McInish and Ord (1985), using intraday data for an equally-weighted index of NYSE listed stocks, from September 1971 to February 1972 and for the calendar year 1982, show that close-to-open returns account for two-thirds of close-to-close returns in the 1971-1972 period. However, in the 1982 period, close-to-open returns account for a percentage that is not statistically different from zero. Branch and Ma (2006) find a very strong negative autocorrelation between the overnight and the intraday return. The study analyses stocks on the NYSE, AMEX and NASDAQ over two periods between 1994 and 2005 and divide the stocks into size categories. The authors find significant relationships for each sub-sample. The correlations do tend to be almost monotonically stronger as market capitalization decreases. They find a powerful negative correlation between the overnight and intraday returns and a significant relationship between overnight and the prior intraday and prior overnight returns. The signs of the correlations alternate, with adjacent periods having negative correlations and one step back being positively related. Using samples[3] of a diverse set of data in the period 1993-2006, Cliff et. al. (2008) also 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 are strongly positive and returns during the day are close to zero and sometimes negative. The authors show that this day and night effect is found on individual equities, equity indices, exchange traded funds and futures contracts on equity indices and is robust in the NYSE, NASDAQ, AMEX and Chicago Mercantile Exchanges.

These results constitute a reversal on the evidence reported by studies on the weekend effect in returns. French (1980), Agrawal and Ikenberry (1994), Wang *et al.*(1997) and Zainudin *et al.*(1997), among others, found that weekend returns were negative (from Friday close to Monday close) and that a significant part of this effect was generated from Friday close to Monday open. Evidence of positive and significant overnight returns contrasts with previous reported evidence under the day-of-the-week effect where, in general, returns in some transaction days are larger than in others. The most commonly reported anomaly in this regard is the significantly lower returns, if not negative, on Mondays and, usually, higher returns on Friday (Jaffe and Westerfield, 1985; Chang *et al.*, 1993, 1998). Several hypotheses have been proposed to explain the occurrence of the day-of-the-week effect: information release hypothesis, where companies delay disclosure of negative information until late in the week and the information processing hypothesis, linked with the asymmetry in information costs between small and large investors (Thaler, 1987).

The timing of information releases was also used to explain the potential day and night effect. Early papers on the timing of earnings announcements found that companies had a tendency to publicize bad news after the market close. Patell and Wolfson (1982) find that good news are more likely to be disclosed while the markets are open but bad news are more likely to be released after the market close. In turn, Bagnoly, Clement and Watts (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. Damadoram (1989) shows that despite the 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, these announcements of bad news can explain only a small part the weekend effect. Doyle and Magilke (2009) reexamine the conventional wisdom that managers are more likely to delay disclosure of unexpected negative information until after the close of the markets. They 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 or on Monday-to-Thursday period.

[3] Individual stocks included in the S&P 500 index, individual stocks of technology companies included in the AMEX interactive week internet index, a sample of 14 of the largest exchange traded funds and the intraday behavior of the S&P 500 E-mini futures contract.

Cliff *et. al.* (2008) also test the possible explanation of the timing of disclosure. Using a sample of earnings announcements, where the time of disclosure and the unexpected earning signals (positive, negative, neutral) are identified, results show that although there was a trend in the period of analysis (2000-2005) for managers to disclose positive unexpected earnings after the market closes, this tendency does not explain the significant day and night effect on returns.

Another proposed argument to explain the night and daytime effect are the effects of asset liquidity. Amihud (2002) documents a negative relationship between various measures of liquidity and future stock returns: increased (lower) risk or transactions costs of low (high) liquidity stocks would predict more (less) night minus daytime return spread. Using various measures to proxy liquidity, Cliff *et al.* (2008) find that positive night minus day return spread is not mainly due to the high risk or to the high transaction costs of illiquid stocks.

Given evidence that the positive and significant night minus day return spread constitutes an anomaly in US equity markets, which appears to have a substantial magnitude and that goes against the predictions suggested by models based on the risk of assets, this pattern is an intriguing fact and a challenge for the asset pricing literature to explain it. In the following section we present the data and methods used to examine hypotheses of the day and night effects, decomposed by day of the week and pre- and post-holiday.

3. DATA AND METHODOLOGY

The data employed in this study are opening (first recorded trade) and closing (last recorded trade) daily prices from a group of ETFs based on the main US equity market indices. 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. In our analysis we use broad-based index ETFs. The ETFs used are the DIA (representing the Dow Jones Industrial Average 30), the IWM (representing the Russel 2000 index - a small-cap US companies index), the QQQQ (representing the NASDAQ 100 index) and the SPY (SPYDERs - representing the S&P 500 index).

The process to create an ETF begins when an ETF manager (the sponsor) submits a plan with the U.S. Securities and Exchange Commission to create an ETF. Once the plan is approved, the sponsor forms an agreement with an authorized participant, generally a market maker, specialist or large institutional investor, who is empowered to create or redeem ETF shares. The authorized participant borrows stock shares, places those shares in a trust and uses them to form ETF creation units. These are bundles of stocks varying from 10,000 to 600,000 shares, but 50,000 shares is what's commonly designated as one creation unit of a given ETF. Then, the trust provides shares of the ETF, which are legal claims on the shares held in the trust to the authorized participant. Because this transaction is an in-kind trade - that is, securities are traded for securities - there are no tax implications. Once the authorized participant receives the ETF shares, they are sold to the public on the open market just like stock shares (SEC, 2012).

When investors want to sell their ETF holdings, they can do so by one of two methods. The first is to sell the shares on the open market. This is generally the option chosen by most individual investors. The second option is to gather enough shares of the ETF to form a creation unit, and then exchange the creation unit for the underlying securities. This option is generally only available to institutional investors due to the large number of shares required to form a creation unit. When these investors redeem their shares, the creation unit is destroyed and the securities are turned over to the redeemer. This option has no tax implications for the portfolio (SEC, 2012).

The return series of ETFs were obtained from www.finance.yahoo.com. The return series span from the period 3rd January 1994 to 3rd January 2014. However, according to Kelly and

Clark (2011), the liquidity of the ETFs was poor during the first half of the 90's and has vastly improved during the second half of this decade. To determine from where to start the analysis, we follow the criteria used by Kelly and Clark (2011). Kelly and Clark computed the 5th percentile of sorted opening and closing times. 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.

Based upon their criteria, DIA data are used from 1998, IWM data are used from 2001, QQQQ data are used from 1999 and SPY data are used from 1996. Until the middle of the first decade of this century, while the AMEX exchange was the primary exchange for most of the ETFs, they also actively traded on other exchanges. From the second half of this decade the QQQQ ETF is primarily traded on NASDAQ exchange and the other ETFs are primarily traded in NYSE ARCA, a subsidiary of NYSE Euronext, the second largest electronic communication network in terms of shares traded in US markets.

We calculate the returns in the ETFs during the two daily time sub-periods: night (close--to-open prices) and daytime (open-to-close prices) returns. As is common in the analysis of daily and intraday financial data, we work with log returns: $r^C_{i,t}=ln[P_{i,t}/P_{i,t-1}].100$ where $P_{i,t}$ is the ETF level i at the end of day t and the C superscript stands for daily close-to-close return. We decompose the continuously compound close-to-close return on day t for the ETF i as $r^C_{i,t}=r^N_{i,t}+r^D_{i,t}$ where $r^N_{i,t}$ stands for the night and $r^D_{i,t}$ for the daytime return. The reported average returns are geometric averages returns and therefore its sign indicates whether the ETF gained or lost value during this intraday range over the sample period.

Two approaches are used to examine the hypotheses of day and night effects on ETFs. The first involves a descriptive analysis of the returns and tests of equality of means returns using parametric tests.

The second is a panel data regression-based approach. The following panel data regression model is specified to capture night and daytime effects common in all ETFs:

$$r_{i,t}=\beta_0+\beta_1 x_{i,t\_night}+\varepsilon_{i,t} \qquad (1.1)$$

where the depend variable $r_{i,t}$ is a scalar accounting for $r^N_{i,t}$ or $r^D_{i,t}$, x_{i,t_night} is a dummy variable taking the value of one for the night period t in the i th ETF and zero otherwise, β are parameters to be estimated and $\varepsilon_{i,t}$ is the error term. Then, the following panel data regression model is specified to capture common night and daytime effects, by day-of-the-week:

$$r_{i,t}=\beta_0+\sum^4_{k=1}\beta_k x_{k,i,t\_daytime}+\sum^5_{k=1}\theta_k x_{k,i,t\_night}+\varepsilon_{i,t} \qquad (1.2)$$

where $x_{k,i,t\_daytime}$ and x_{k,i,t_night} are dummy variables for the daytime and night periods, in the k day of the week, in the i th ETF, respectively. β_k and θ_k are parameters to be estimated and $\varepsilon_{i,t}$ is the error term. The night and daytime average returns, by day of the week, are the coefficients (given the reference category) from a regression of the panel of returns (stacked for all ETFs, dates, and daily sub-periods) on dummy variables for the daily time periods, by day of the week.

To evaluate the statistical significance of the parameters and test for the robustness of the day and night effects in the US equity ETFs, we use various robust variance-covariance matrix estimators for the panel data models to account for departures of the residuals from homoscedasticity, temporal, contemporaneous and cross-temporal independence. As is common in the analysis of times-series cross-section (TSCS) data in finance, the disturbances $\varepsilon_{i,t}$ exhibits evidence to be heteroskedastic within and between groups, autocorrelated up to some lag, contemporaneously and possibly cross-temporal correlated between groups in different time periods (Petersen, 2008). The errors may be heteroskedastic between groups and autocorrelated within panels but must have zero conditional mean, $E[\varepsilon_{i,t}\,|\mathbf{X}'_{it}]=0$, where the vector of independent variables is $\mathbf{X}'_{it}$. We make the following assumptions about correlations between errors:

ETF effects: The errors may exhibit ETF effects, meaning that errors may have arbitrary correlation across time for a particular fund: $E[\varepsilon_{i,t}\ \varepsilon_{i,t-k}|\mathbf{X}'_{i,t}, \mathbf{X}'_{i,t-k}] \neq 0$ for $k\neq 0$.

Time effects: The errors may exhibit time effects, meaning that errors may have arbitrary correlation across ETFs at a moment in time: $E[\varepsilon_{i,t}\ \varepsilon_{j-k}|\mathbf{X}'_{i,t}, \mathbf{X}'_{j,t}] \neq 0$ for $i\neq j$.

Persistent common shocks: The errors may exhibit persistent common shocks, meaning that we allow for some correlation between different ETFs in different time periods, but these shocks fade away over time, and may be ignored after L periods. So $E[\varepsilon_{i,t}\ \varepsilon_{j,t-k}|\ \mathbf{X}'_{i,t}, \mathbf{X}'_{i,t-k}]=0$ for $i\neq j$ and $|k|>L$.

To understand the differences between time effects, ETF effects, and persistent common shocks, consider the following data generating process (Thompson, 2011):

$$\varepsilon_{i,t}= \theta_i' f_t+\omega_{it}+u_{it}, \tag{2}$$

$$\omega_{it}=\varphi\omega_{i,t-1}+\delta_{it}, \qquad \omega_{i,o}=0,$$

where f is a vector of random factors common to all ETFs, and θ is a vector of factor loadings specific to ETF i. u_{it} and δ_{it} are random shocks, uncorrelated across both ETF and time. The ω_{it} term generates ETF effects - shocks specific to ETF i. $\theta_i' f_t$ generates both time effects and persistent common shocks. When f is uncorrelated across time, we have time effects but no persistent common shocks - ETFs are correlated with one another at a moment in time, but different ETFs in different time periods are uncorrelated. When f is persistent, we have both time effects and persistent common shocks. We assume that the autocorrelations for f_t disappear after L periods[4] . In the following sub-sections we present the various variance--covariance matrix estimators robust to the various behavior assumptions of the errors.

3.1. Beck-Katz method

The Parks-Kmenta method, originally proposed by Parks (1967) and then improved by Kmenta (1986), seeks to take into account heteroscedasticity between panels as well as temporal and spatial dependence in the residuals of TSCS models. This method uses an application of the Generalized Least Squares (GLS) estimation. This estimation is based on the assumption that the variance-covariance matrix of the errors, $\mathbf{\Omega}$, is known. However, since the $\mathbf{\Omega}$ matrix is unknown, this method uses the feasible GLS (FGLS) using a consistent estimator of this matrix to obtain a consistent estimator of the coefficient vector, β. This method combines assumptions about serial correlation (within panels), contemporaneous correlation (between panels) and heteroscedasticity of disturbances across panels:

$$E[\varepsilon^2_{it}|\mathbf{X}_{it}]=\sigma_{ii} \tag{3.1}$$

$$E[\varepsilon_{it}\ \varepsilon_{jt}|\ \mathbf{X}_{it}, \mathbf{X}_{jt}]=\sigma_{ij} \text{ if } i\neq j, \tag{3.2}$$

$$\varepsilon_{it}=\rho_i\ \varepsilon_{i,t-1}+\upsilon_{it}, \tag{3.3}$$

where ρ_i is the first-order autoregressive coefficient, allowing the value of this parameter to vary from panel to panel, being necessary to find consistent estimators of the σ_{ii}, σ_{ij} and ρ_i

---

[4] More generally, shocks to $\mathbf{f}_t$ could decay slowly but not completely disappear after L periods. For example, $\mathbf{f}_t$ could follow a first-order autoregressive process. While this would violate the assumption, we assume that after some time the correlation between shocks is small enough that it can be ignored. Autoregressive processes could be handled by allowing the lag length L to grow with the sample size (see for example Newey and West, 1987).

(i.e., matrix elements of $\mathbf{\Omega}$). The estimated GLS results of the Parks-Kmenta method are given by

$$\hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{y}), \quad (4)$$

$$\hat{\mathbf{V}}\mathbf{ar}(\hat{\boldsymbol{\beta}}_{GLS}) = (\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{X})^{-1}. \quad (5)$$

The variance-covariance matrix of the disturbances can be written in terms of the Kronecker product

$$E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \mathbf{\Omega} = \mathbf{\Sigma}_{\mathbf{m\,x\,m}} \otimes \mathbf{I}_{\mathbf{T_i\,x\,T_i}}, \quad (6)$$

and the $\hat{\mathbf{\Sigma}}_{\mathbf{m\,x\,m}}$ matrix is estimated as

$$\hat{\mathbf{\Sigma}}_{i,j} = (\hat{\boldsymbol{\varepsilon}}_i'\hat{\boldsymbol{\varepsilon}}_j)/T_{i,j}, \quad (7)$$

where $\hat{\boldsymbol{\varepsilon}}_i$ and $\hat{\boldsymbol{\varepsilon}}_j$ are the residuals for panels *i* and and *j* is the number of matched time period observations between panel *i* and *j*.

The Parks-Kmenta method consists of three sequential FGLS transformations and is estimated as follows. In the first step OLS pooled regressions (1.3) or (1.4) are estimated. Then residuals from the first step are used to estimate panel *i* - specific serial correlation. In the second step, ρ_i are used to transform the model into a model with serially independent errors. Residuals of the second step are then used to estimate contemporaneous correlation and heteroscedasticity of the errors across panels. In the third step the data is again transformed with estimates of the correlations and variances to allow the parameter estimation of equations (1.3) or (1.4) by OLS or the coefficient estimates in equation (4), with errors now corrected of the three previous assumptions, and the standard error estimates of the parameters in equation (5).

According to Beck and Katz (1995), although the GLS method has optimal properties for handling TSCS data, the Parks-Kmenta FGLS method does not hold the same properties. This is because, although the FGLS uses an estimate of the error process, the FGLS´s formula for standard errors assumes that the variance-covariance matrix is known, not estimated. According to Beck and Katz this is a problem in TSCS models because the errors' process has a large number of parameters. In contexts where the number of time- periods is not much higher than the number of cross-sectional units, this oversight would lead standard error estimates of the estimated coefficients to understate their true variability. In the dataset of this study, however, as the number of time-periods points is immeasurably greater than the number of cross-sectional units, overconfidence in the standard errors should not occur because the $Ti|m$ ratio is high.

In the Beck and Katz (1995)' method disturbances are assumed to be heteroskedastic, contemporaneously correlated (across panels), autocorrelated (within panel) and the first-order autocorrelation parameter can be constant or different for each panel. This method produces OLS estimates of $\boldsymbol{\beta}$ when autocorrelation is not specified or uses Prais-Winsten method to produce parameters' estimates when autocorrelation is specified. By specifying autocorrelation, parameters' estimates are conditional to estimates of the autocorrelation parameters. To estimate the parameters' variance-covariance matrix, which is asymptotically efficient under the assumed covariance structure, this method also uses the variance-covariance matrix of the disturbances and the estimation method of its elements is similar to that used by the Parks--Kmenta method, which are described in equations (6) and (7), respectively. Parks-Kmenta and Beck-Katz estimators are both consistent as long as the conditional average $(\mathbf{x}_{it}'\boldsymbol{\beta})$ is correctly specified. However, if the assumed covariance structure of the errors is correct, parameter estimates generated by Parks-Kmenta method are more efficient.

The estimated variance-covariance matrix of parameters of OLS or Prais-Winsten is given by

$$\hat{\mathbf{V}}\mathbf{ar}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})^{-1} \tag{8}$$

where $\hat{\boldsymbol{\Omega}}$ is the full estimated covariance matrix of disturbances. When errors are serially correlated, this method assumes a first-order autoregressive process of errors and uses the GLS to estimate the parameters whose estimates are conditional on the estimated value of .

The Prais-Winsten method consists on the following. The 0-th iteration is obtained by estimating equations (1.3) or (1.4). An estimate of the serial correlation in residuals series is obtained. By default the following regression is used $\varepsilon_{i,t}=\rho_i\varepsilon_{i,t}+\delta_{it}$. Then the following transformation is applied and estimated:

$$(r_{it}-\rho_i\, r_{i,t-1})=\left[\mathbf{x}'_{it}-\rho_i\,\mathbf{x}'_{it-1}\right]\boldsymbol{\beta}+v_{it}. \tag{9}$$

By default, this estimation process is performed using an iterative mechanism until convergence is achieved; this process is repeated until the change in the ρ_i estimate is within a specified tolerance. Thus, the new parameter estimates are used to produce estimates of $\hat{r}_{it}=\mathbf{x}'_{it}\hat{\boldsymbol{\beta}}$ and then ρ_i is re-estimated using, by default, the regression defined by

$$(r_{it}-\hat{r}_{it})=\rho_i\left[r_{i,t-1}-\hat{r}_{i,t-1}\right]+\varepsilon_{it}. \tag{10}$$

Then equation (9) is re-estimated using the new estimate and continue to iterate between (9) and (10) until the estimated converges. The residuals, removed from the serial correlation, are then used to estimate the $E\left[\boldsymbol{\varepsilon\varepsilon}'\right]=\boldsymbol{\Omega}$ matrix of the residuals and the $\mathbf{Var}(\boldsymbol{\beta})$ matrix of parameter's standard errors.

3.2. Double clustered standard errors by time and stock indices effects

In finance panel data analysis, it is common that the error term includes a specific effect in the unit (index) as well as common shocks that affect all units (indexes), i.e., that the error term is correlated temporally (within indexes) and contemporaneously (between indexes). When using the usual standard errors that do not fit to the correlation between observations for the two dimensions, Cameron *et al.*(2006), Petersen (2009) and Thompson (2011) show that this choice of methodology can lead to standard errors that are very small. Small standard errors lead to higher than expected t statistics and F statistics, thus showing significance even when it does not exist. When we cluster along a single dimension (within or between) this can lead to understated standard errors. Thompson (2011) considers that when clustering is made by time allowing only units to be correlated with one another at a given time period, this procedure will ignore specific persistent effects to a unit. One way to simultaneously handle unit and time effects would be to use unit and time dummies (for example, we could cluster standard errors by time and include units fixed effects including in the regression dummy variables specific to the units). However, Thompson considers that this procedure does not take into account many relevant forms of correlated errors because inclusion of time periods and units dummies does not correctly model the structure of the temporal correlation (time dummies) and the autoregressive process (unit dummies) of the errors.

Cameron *et al.*(2006) and Thompson (2011) propose the double clustered standard error estimator robust to temporal (within units) and sectional (between units) correlation. In this regard, this study also uses the double clustered standard error estimator, also used by Cliff, Cooper and Gulen (2008). Double clustered standard errors are formulated with the following covariance matrix estimator of $\boldsymbol{\beta}$

$$\hat{\mathbf{V}}\mathbf{ar}(\hat{\boldsymbol{\beta}}) = \hat{\mathbf{V}}_{\mathbf{t}}+\hat{\mathbf{V}}_{\mathbf{i}}-\hat{\mathbf{V}}_{\mathbf{w}} \tag{11}$$

where the three matrices are the variances-covariances matrices time-clustered, index-clustered and the White (1980) consistent estimator (no clustering), respectively. Each matrix is a heteroscedasticity autocorrelated consistent (HAC) "sandwich" estimator which takes the form $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{S}(\mathbf{X}'\mathbf{X})^{-1}$, where each differ in the construction of $\mathbf{S}$, the spectral density matrix of $\hat{\mathbf{u}}_{it} = \mathbf{x}_{it}\hat{\varepsilon}_{it}$, where $\hat{\varepsilon}_{it}$ is the residual in t -th time period and i -th stock index and x_{it} represents the vector of the regressors` values in *t -th* time period and *t -th* stock index. The $\hat{\mathbf{V}}_{\mathbf{t}}$ term assumes that a (common) shock at any given time affects multiple indexes and so account for cross-index (contemporaneous) correlation. This term is computed as

$$\mathbf{S}_t = \sum_{t=1}^{T} \mathbf{\Omega}_t, \tag{12.1}$$

$$\mathbf{\Omega}_t = \mathbf{\Gamma}_0 + \sum_{n=1}^{N-1} \left[\mathbf{\Gamma}_n + \mathbf{\Gamma}'_n\right], \tag{12.2}$$

$$\mathbf{\Gamma}_n = \sum_{i=1}^{N-n} \mathbf{u}'_{it}\mathbf{u}_{i+n,t}. \tag{12.3}$$

The $\mathbf{\Gamma}_n$ matrix measures correlation between indices that are n - positions apart. Because, unlike the time, does not exist a natural order of the indexes $\mathbf{\Omega}_t$ captures all possible interactions between i and j indices at the t-time period and $\mathbf{S}_t$ captures all the possible time periods in the analysis. The clustered index term, $\hat{\mathbf{V}}_{\mathbf{i}}$, has a similar structure. This term is computed as

$$\mathbf{S}_i = \sum_{i=1}^{N} \mathbf{\Omega}_i, \tag{13.1}$$

$$\mathbf{\Omega}_i = \mathbf{\Gamma}_0 + \sum_{\tau=1}^{T-1} \left[\mathbf{\Gamma}_\tau + \mathbf{\Gamma}'_\tau\right], \tag{13.2}$$

$$\mathbf{\Gamma}_\tau = \sum_{t=1}^{T-\tau} \mathbf{u}'_{it}\mathbf{u}_{i,t+\tau}. \tag{13.3}$$

The $\mathbf{\Gamma}_\tau$ matrix measures correlation between residuals at the time t and $t+\tau$ for a given i -th index. Thus, $\mathbf{\Omega}$ is the HAC estimator for a single index. The above general version suggests that all possible lags from 0 to T-1 should be included. However, when an excessive number of lags are included, these estimators have poor finite sample properties. Standard implementations, such as the Newey-West method, modify the $\mathbf{\Omega}$ computation by truncating the summation at a much lower lag length k[5]. We follow this intuition and set K = 4 daily time sub-periods, taking into account the order of the significant autocorrelation coefficients in each index. Finally, the $\hat{\mathbf{V}}_w$ term is the White´s conventional estimator, ignoring any clustering, and is computed as

$$\hat{\mathbf{V}}_w = \sum_{i=1}^{N} \sum_{t}^{T_i} \mathbf{u}'_{it}\mathbf{u}_{it}. \tag{14}$$

This estimator is subtracted to avoid double counting of the common-index, common--time terms ($\mathbf{\Gamma}_0$) included in $\hat{\mathbf{V}}_t$ and $\hat{\mathbf{V}}_i$ estimators. All three $\hat{\mathbf{V}}$ estimators are multiplied by a finite sample adjustment of *(T-1)/(T-p)*, where p is the dimension of $\boldsymbol{\beta}$. The estimators $\hat{\mathbf{V}}_t$ and $\hat{\mathbf{V}}_i$ are also multiplied by *T/(T-1)* and *N/(N-1)*, respectively, to adjust for the number of clusters.

Thompson (2011) warns that it can sometimes be a disadvantage to double-clustering and that it is not always appropriate to use the most robust standard error formula. Thompson considers that the more robust standard error formula tend to have less bias, but more variance. The lower bias improves the performance of the statistical test. But the increase in variance often leads to find statistical significance even when it does not exist.

---

[5] The Newey-West procedure also weights Γ_τ by $w_\tau = 1 - \tau/(k+1)$ to ensure that the resulting matrix is positive definite.

3.3. Robust standard errors to general forms of cross-sectional, temporal and cross-temporal dependence

It is common in economic and financial processes that the data behavior be heteroscedastic, correlated over time, correlated between subjects in a given time period and correlated between different subjects in different time periods. To estimate standard errors of the parameters that are robust to these features of the errors on the underlying panel data model, Driscoll and Kraay (1998) demonstrate that the standard nonparametric time-series covariance matrix estimator (Newey-West) can be modified such that it is robust to general forms of cross-sectional as well as temporal dependence. Driscoll and Kraay approach loosely applies a Newey-West - type correction to the sequence of cross-sectional averages of the moment conditions. Driscoll and Kraay´s standard errors for the coefficient estimates are then obtained as the squares roots of the diagonal elements of the asymptotic (robust) covariance matrix,

$$\hat{\mathbf{V}}\mathbf{ar}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{S}_T(\mathbf{X}'\mathbf{X})^{-1}, \tag{15}$$

where $\boldsymbol{S}_T$ is defined as in Newey-West (1987)

$$\mathbf{S}_T = \boldsymbol{\Omega}_0 + \sum_{j=1}^{m(T)} w(j,m)\left[\boldsymbol{\Omega}_j + \boldsymbol{\Omega}'_j\right] \tag{16}$$

where *m(T)* denotes the lag length up to which the residuals may be autocorrelated and the modified Bartlett weights *w(j,m)*=1-j/*(m(T)*+1) ensures positive semi-definiteness of $\boldsymbol{S}_T$ and smooth the sample autocovariance function such that higher-order lags receive less weight. The $\boldsymbol{\Omega}_j$ matrix is defined as

$$\boldsymbol{\Omega}_j = \sum_{t=j+1}^{T} \mathbf{u}_t \mathbf{u}'_{t-j}, \text{ with } \mathbf{u}_t = \sum_{i=1}^{N(t)} \mathbf{u}_{i,t} \tag{17}$$

In (17), the sum of the individual time t moment conditions $\mathbf{u}_{i,t}$ runs from 1 to *N(t)*, where N is allowed to vary with t. This is a small adjustment to Driscoll and Kraay (1998) original estimator introduced by Hoechle (2007) in his *xtscc* Stata program. This adjustment makes their estimator ready for use with unbalanced panels. For pooled OLS estimation, the individual orthogonality conditions $\mathbf{u}_{i,t}$ in (17) are the (kx1) dimensional moment conditions of the regression model, i.e., $\mathbf{u}_{i,t} = \mathbf{x}_{it}\varepsilon_{it}$. From (16) and 17) it follows that Driscoll and Kraay (1998)´s covariance matrix estimator equals the heteroskedasticity and autocorrelation consistent covariance matrix estimator of Newey-West (1987) applied to the time series of cross-sectional averages of the $\mathbf{u}_{i,t} = \mathbf{x}_{it}\varepsilon_{it}$. Hoechle (2007) considers that estimating the covariance matrix with this approach yields standard errors that are robust to general forms of cross-sectional and temporal dependence. In the following section, to estimate pooled OLS long panel data and fixed effects regression models with double clustered standard errors and standard errors robust to general forms of cross-sectional, temporal and cross-temporal dependence in residuals, we use Stata´s ivreg2 and xtivreg2 routines from Baum, Schaffer and Stillman, (2010) and Schaffer (2010), respectively.

4. RESULTS AND DISCUSSION

4.1. Summary statistics of night and day returns

Summary statistics for day and night returns series of ETFs during the entire sample period are reported in table 1. These returns are categorized according to the hypothesized day of the week effect for time period of the day. Results show that there is no pattern for the higher night average returns by day of the week across ETFs. In turn, lower day average returns are recorded on Friday in 3 out of 4 ETFs. Regarding the volatility of day returns (as measured by standard deviation) results show no pattern by day of the week on higher volatility. Concerning lower volatilities, these occurs at Friday.

The night average returns are higher on Tuesday in 3 out of 4 ETFs. In fact, every Tuesday night's returns are positive and significantly greater than zero in all 4 ETFs. Regarding lower night returns there is no pattern by day of the week across the ETFs. Regarding overnight volatility, the lowest value is exhibited on Wednesday in 3 out of 4 and the highest value is exhibited on Friday in 2 out of 4 ETFs (SPY and SPDR). Overall, for all ETFs, volatility during the day is about two times greater than that exhibited during the night period for all days of the week.

The distributional properties of the return series, by time period of day and day of the week, for all ETFs, are not normal. Since the sampling distribution of the skewness parameter of a normal distribution is normal with zero mean and standard deviation $\sqrt{6/T}$, where T is the sample size, virtually all returns by time period of the day and day of the week are positive or negative and significantly biased, but there is no common pattern as to the sign and magnitude of the bias. By time period of the day and day of the week, the SPY and SPDR ETFs have a sign and significance approximately identical. As to skewness, for day time returns, QQQQ and IWM ETFs also have a similar pattern by day of the week. It is observed that in 3 out of 4 ETFs (SPY, SPDR and IWM) night returns on Monday (weekend return) are positive and significantly skewed, suggesting a greater likelihood of the returns having higher values than under the normal. Conversely, day returns on Monday are negative and significantly skewed and are more likely to have lower extreme values than under the normal.

The degree of excess kurtosis across ETFs, for returns categorized by time period of the day and day of the week, is also high, suggesting leptokurtic (i.e., peaked) distributions with higher chances of generating extreme returns (i.e., fat tails) compared to a normal distribution. Since the sampling distribution of kurtosis parameter is normal under a normal distribution, with mean zero and standard deviation $\sqrt{24/T}$, where is the sample size, almost all kurtosis estimates for night and day returns categorized by day of the week are statistically significant at 1% level, there being no pattern in the magnitude estimate between day and night returns.

In table 1, the average return column also presents returns significantly different from zero. Only night returns are positive and significantly different from zero. Tuesday-night (all ETFs), Monday-night (SPY and IWM), Thursday-night (QQQQ) and pooled night returns (SPY, DIA and QQQQ) are positive and significantly different from zero. In no ETF day returns are significantly different from zero. These results suggest the possibility that the day and night effect might occur in some ETF, where in some days of the week night returns are positive and significantly higher than the corresponding day return. In general, day returns are not significantly different from zero. The significant higher volatility on day than night returns constitute a puzzle to the asset pricing literature in face of the significant lower average day than night returns.

In the last two columns of Table 1 are presented results of Student t-tests, with and without equal variances, for the equality of means between day and night returns, by day of week and overall return and for each ETF. For the SPY and SPDR ETFs it is observed that in no day of the week there is a significant difference between night and day average return.

For the SPY ETF, however, the overall night is significantly higher than the overall day average return at the 0, 05 level. To quadruple Q ETF, the Tuesday, Friday night and pooled returns are significantly higher than the corresponding day returns. To the quadruple Q ETF, Tuesday, Friday and overall night returns are significantly higher than the corresponding day returns. For the IWM ETF, only the night (weekend return) is significantly higher than the day Monday average return.

Table 1: Descriptive analysis of night and day returns by day of the week

Exchange Traded Fund	Number	Mean (%)	Std. dev. (%)	Skewness	Kurtosis	t- test of equality of means (equal variances)	t- test of equality of means (unequal variances)
SPY							
Daytime							
Monday	855	-0.0155	1.1276	-0.7419	12.456	-1.589	-1.602
Tuesday	929	0.0104	1.1432	0.3870	10.652	-1.244	-1.276
Wednesday	930	0.0020	1.0894	0.1799	11.439	-0.266	-0.267
Thursday	912	-0.0131	1.0829	-0.7335	12.007	-0.817	-0.819
Friday	908	-0.0316	1.0050	0.1023	5.340	-1.263	-1.268
total	4534	-0.0093	1.0901	-0.1521	10.788	-2.306**	-2.325**
Overnight							
Monday	823	0.0578***	0.7082	0.6724	13.006	---	---
Tuesday	846	0.0657***	0.6332	0.6447	11.271	---	---
Wednesday	920	0.0129	0.5997	-0.6918	7.599	---	---
Thursday	903	0.0208	0.6332	-0.2906	6.878	---	---
Friday	881	0.0219	0.7710	-1.8800	26.618	---	---
total	4373	0.0350***	0.6710	-0.4736	16.341	---	---
DIA							
Daytime							
Monday	755	0.0194	1.0660	-0.2820	13.3011	-0.270	-0.273
Tuesday	821	0.0018	1.0383	0.2270	10.3669	-1.165	-1.193
Wednesday	825	0.0096	1.0472	0.0623	12.6368	-0.061	-0.061
Thursday	810	0.0107	1.0732	-0.6773	12.8931	0.099	0.100
Friday	804	-0.0322	0.9701	0.1826	5.6626	-0.313	-0.314
total	4015	0.0017	1.0390	-0.1150	11.354	-0.740	-0.746
Overnight							
Monday	726	0.0320	0.6888	0.4618	10.936	---	---
Tuesday	746	0.0522***	0.5919	1.1312	11.95	---	---
Wednesday	815	0.0121	0.5697	-0.5804	6.611	---	---
Thursday	803	0.0063	0.6198	-0.1226	7.594	---	---
Friday	782	-0.0186	0.7340	-2.3344	35.416	---	---
total	3872	0.0161*	0.6430	-0.5431	18.894	---	---
QQQQ							
Daytime							
Monday	701	-0.0504	1.6341	0.4818	12.198	-0.167	-0.165
Tuesday	762	-0.0886	1.8076	-0.6387	6.8034	-2.173**	-2.245**
Wednesday	765	0.0048	1.8635	0.9034	17.470	-0.186	-0.187
Thursday	752	0.0396	1.6841	-0.0991	6.1786	-0.566	-0.568
Friday	749	-0.0994	1.5386	0.1086	7.8503	-2.026**	-2.036**
total	3729	-0.0385	1.7115	0.1652	10.887	-2.017**	-2.022**

Overnight							
Monday	676	-0.0298	2.8160	-21.882	538.48	---	---
Tuesday	692	0.0731***	0.7913	0.4330	7.327	---	---
Wednesday	755	0.0190	0.9479	0.1910	10.019	---	---
Thursday	745	0.0794***	0.9251	0.6980	7.7974	---	---
Friday	728	0.0399	1.0541	-0.3125	12.539	---	---
total	3596	0.0369*	1.4836	-28.125	1317.11	---	---
IWM							
Daytime							
Monday	615	-0.0906	1.3703	-0.8618	7.821	-2.195**	-2.213**
Tuesday	669	0.0260	1.4458	-0.2212	6.891	-0.986	-1.014
Wednesday	671	0.0365	1.4429	0.0288	8.193	0.603	0.606
Thursday	659	-0.0269	1.4330	-0.4729	9.795	0.306	0.305
Friday	657	-0.0083	1.3171	0.5877	9.160	-0.476	-0.478
total	3271	-0.0113	1.4034	-0.1941	8.397	-0.823	-0.822
Overnight							
Monday	593	0.0528*	0.8259	0.2368	8.503	---	---
Tuesday	607	0.0907***	0.7562	1.8274	23.101	---	---
Wednesday	662	-0.0012	0.7196	-0.2882	7.160	---	---
Thursday	652	-0.0647	2.8199	-22.830	561.88	---	---
Friday	638	0.0218	0.9272	-0.1987	32.434	---	---
total	3152	0.0181	1.4718	-33.233	1571.4	---	---

Sample period is 20th January 1998 to 3rd January 2014 for SPDR, 1st January 1996 to 3rd January 2014 for SPY, 10thMarch 1999 to 3rd January 2014 for QQQQ and 30th May 2000 to 3rd January 2014 for IWM.

4.2. Panel data regression-based analysis of night and day time effects

This section presents the estimated panel data regression models, first, by day time period and, then, simultaneously, by day time period and day of the week. In these long panel data models we use standard error estimators of the coefficients robust to departures from independent and identically distributed (i.i.d.) residuals over time and across sectional dimensions. The panels comprise $N=4$ ETFs and T varies by ETF, i.e., we use an unbalanced panel.

With the first regression model specified in Eq. (1.1) we intend to examine whether ETFs show significant common high-low tendencies in average returns during the night and day time periods. In the second regression model, specified in Eq. (1.2), we examine whether there occurs significant common high-low regularities in mean returns during the night and day time periods by day of the week. The occurrence of such abnormal mean returns across ETFs by time period of the day and day of the week could be profitably used for timing the deals: when buying / selling at the open and selling / buying at the close of the market.

The estimated coefficients and standard errors of the parameters specified in Eq. (1.1) to testing for the day and night common effects across ETF are presented in Table 2. The table also includes the R^2, the adjustedR^2 and an F-test (Wald χ^2- test) of the null hypothesis that the slope coefficients are jointly zero. The column of the pooled OLS regression also includes the White's general test, the modified Wald test and the Breusch- Pagan LM test. The White´s general test is used to test for heteroscedasticity in the residuals, where the assumption of normally distributed residuals is relaxed. The modified Wald test is used to test for groupwise heteroscedasticity in the residuals of the pooled OLS regression and the Breusch-Pagan LM test is used to test for contemporaneous correlation of residuals across ETFs. We also conducted tests for cross-temporal pairwise correlations between residuals from ETFs which proved significant at several lags.

Thus, the null hypotheses of no within heteroscedasticity, no groupwise heteroscedasticity, no within serial correlation, no contemporaneous and no cross-temporal correlations in residuals across ETFs from pooled OLS regression were rejected. Following this, regression models were estimated with standard errors corrected for these departures of the residuals from i.i.d assumptions using the panel corrected standard errors method (Beck-Katz method), the pooled OLS method with standard errors clustered by time and ETF, the pooled OLS and the OLS fixed effect methods with standard errors corrected for heteroscedasticity, serial correlation, contemporaneous and cross-temporal correlations in residuals.

Table 2: Estimated coefficients and standard errors of night and daytime effects for U.S equity exchange-traded funds panel data regressions

Model		**Pooled OLS (1)**	**PCSE (3)**	**Double-clustered (4)**	**Driscoll-Kraay (5)**	**Fixed - Driscoll-Kraay (6)**
Constant (daytime)		-0.0139	-0.0141	-0.0139	-0.0139	------
	(s.e.)	(0.0097)	(0.0153)	(0.0169)	(0.0178)	------
Overnight		0.0410***	0.0412*	0.0410*	0.0410*	0.0410*
	(s.e.)	(0.0139)	(0.0217)	(0.0211)	(0.0216)	(0.0216)
NxT		30542	30542	30542	30542	30542
R^2		0.0003	0.0003	0.0003	0.0003	0.0003
Adjusted R^2		0.0003	--------	-------	--------	-------
F- statistic (Wald chi2)		8.65***	(3.59*)	2.84	3.59*	3.59*
White Modified		1.35	---------	---------	---------	---------
Wald-stat		997.83***	---------	---------	---------	---------
Breusch-Pagan LM stat		30150***	---------	---------	---------	---------

Sample period is 20th January 1998 to 3rd January 2014 for SPDR, 1st January 1996 to 3rd January 2014 for SPY, 10thMarch 1999 to 3rd January 2014 for QQQQ and 30th May 2000 to 3rd January 2014 for IWM. F-test and Wald chi2-test are for the null hypothesis that all slope coefficients are zero. White is the White (1980) heteroskedasticity general test for pooled OLS regression model. Modified Wald stat. test for groupwise heteroscedasticity across stock indices in the context of an OLS regression fit to pooled cross-section time series data, following page 598 of Green (2000). Breusch-Pagan (1980) LM statistic test for contemporaneous correlation of residuals across stock indices. (1) stands for OLS regression fit to pooled cross-section time series data, (2) stands for Beck-Katz method, (3) stands for OLS regression fit to pooled cross-section time series data with standard errors clustered by time and stock index effects, (4) stands for OLS regression fit to pooled cross-section time series data with standard errors estimated following Driscoll and Kraay (1998)´s covariance matrix estimator and (5) stands for fixed effects regression model with standard errors estimated following Driscoll and Kraay (1998)´s covariance matrix estimator. Asterisks stands for significance at the *0.10, **0.05, *** 0.01 levels.

In the pooled OLS regression, the estimated coefficient for night return is significantly positive while that for day return is not significantly different from zero. When standard errors of the parameters are adjusted for departures of the residuals from i.i.d. assumptions, the significance of the estimated coefficient for the night return decreases, being now marginally significant. These results suggest that the equity ETFs are characterized by a common night positive albeit marginally significant effect. These results also suggest a sharp decrease in the intensity of the night effect in view of the results reported in previous studies on the same type of assets (Cliff et al., 2008), providing, albeit marginally, some support to theories that predict higher returns during non-trading periods to compensate suppliers investors of liquidity for the risk endured during the period of non-marketability (Longstaf, 1994). However, the lower values of the volatility on night relatively to the day returns constitute a puzzle on the risk-return space to the asset pricing literature.

Table 3: Estimated coefficients and standard errors of night and daytime effects by day of the week for U.S equity exchange-traded funds panel data regressions

Model	**Pooled OLS (1)**	**PCSE (3)**	**Double-clustered (4)**	**Driscoll-Kraay (5)**	**Fixed - Driscoll-Kraay (6)**
Constant-daytime	-.0306	-.0304	-.0306	-.0306	-------
(s.e.)	(.0225)	(.0354)	(.0395)	(.0421)	-------
Tuesday-daytime	.0184	.0181	.0184	.0184	.0184
(s.e.)	(.0312)	(.0491)	(.0558)	(.0619)	(.0619)
Wednesday-daytime	.0426	.0418	.0426	.0426	.0426
(s.e.)	(.0312)	(.0490)	(.0549)	(.0602)	(.0602)
Thursday-daytime	.0334	.0327	.0334	.0334	.0334
(s.e.)	(.0313)	(.0492)	(.0525)	(.0590)	(.0589)
Friday-daytime	-.0125	-.0128	-.0125	-.0125	-.0125
(s.e.)	(.0313)	(.0493)	(.0524)	(.0568)	(.0568)
Monday-night.0597*	.0594	.0597	.0597	.0597	
(s.e.)	(.0321)	(.0504)	(.0478)	(.0532)	(.0532)
Tuesday-night.0999***	.0998**	.0999**	.0999**	.0999**	
(s.e.)	(.0319)	(.0500)	(.0471)	(.0486)	(.0486)
Wednesday-night	.0418	.0420	.0418	.0418	.0418
(s.e.)	(.0312)	(.0492)	(.0433)	(.0478)	(.0478)
Thursday-night	.0438	.0436	.0438	.0438	.0438
(s.e.)	(.0314)	(.0494)	(.0474)	(.0533)	(.0533)
Friday-night.0464	.0463	.0464	.0464	.0464	
(s.e.)	(.0316)	(.0497)	(.0506)	(.0511)	(.0511)
NxT	30542	30542	30542	30542	30542
R^2	0.0006	0.0006	0.0006	0.0006	0.0006
Adjusted R^2	0.0003	--------	-------	--------	-------
F- statistic					
(Wald chi2)	1.97**	(7.22)	1.00	1.11	1.11
White Modified	7.84	---------	---------	---------	---------
Wald-stat	997.93***	---------	---------	---------	---------
Breusch-Pagan					
LM stat	30139***	---------	---------	---------	---------

Sample period is 20th January 1998 to 3rd January 2014 for SPDR, 1st January 1996 to 3rd

January 2014 for SPY, 10th March 1999 to 3rd January 2014 for QQQQ and 30th May 2000 to 3rd January 2014 for IWM. F-test and Wald chi2-test are for the null hypothesis that all slope coefficients are zero. White is the White (1980) heteroskedasticity general test for pooled OLS regression model. Modified Wald stat. test for groupwise heteroscedasticity across stock indices in the context of an OLS regression fit to pooled cross-section time series data, following page 598 of Green (2000). Breusch-Pagan (1980) LM statistic test for contemporaneous correlation of residuals across stock indices. (1) stands for OLS regression fit to pooled cross-section time series data, (2) stands for Beck-Katz method, (3) stands for OLS regression fit to pooled cross-section time series data with standard errors clustered by time and stock index effects, (4) stands for OLS regression fit to pooled cross-section time series data with standard errors estimated following Driscoll and Kraay (1998)´s covariance matrix estimator and (5) stands for fixed effects regression model with standard errors estimated following Driscoll and Kraay (1998)´s covariance matrix estimator. Asterisks stands for significance at the *0.10, **0.05, *** 0.01 levels.

Table 3 presents the estimated coefficients and standard errors of the parameters specified in Eq. (1.2) for common day and night effects decomposed by day of the week. In the methods using robust variance-covariance matrix estimators of the parameters, only the Tuesday-night coefficient is positive and significant.

In the pooled OLS method, in addition to the Tuesday-night also the Monday-night coefficient (i.e., weekend return) is positive and marginally significant. Another interesting result, although not statistically significant, is that negative mean returns are only observed during the Monday and Friday day time period.

The results shown in Table 3, concerning the entire sample period, are not consistent with those obtained by Cliff *et al.* (2008) in a sample of 13 US equity ETFs, and individually for the SPY ETF, in the period May 1995 through December 2006. Cliff *et al.* (2008) found that there are significant pervasive differences between day and night returns by day of the week for the sample and for the individually SPY. Our results only exhibit the occurrence of a positive and significant difference between Monday daytime and Tuesday night returns.

Considering that the sampling period in this study goes beyond 2006, until early 2014, the comparison of these results suggests that the US equity ETF market will have become more (weak form) efficient, leading to a decrease or disappearance of the day and night effects previously evidenced by Cliff *et al.* (2008) in the sample period 1995-2006.

In addition to the longer sampling period used in this study, which may have reflected the increased (weak form) efficiency in the 2nd half of the sample period, differences with the results of Cliff *et al.*(2008) may also be due to both institutional and procedural transformations that have taken place from 2006. In August 2006, the Securities Exchange Commission-SEC, mandated by the Congress of United States, adopted the National Market System (NMS) Regulation which introduced regulation designed to protect the rights of the investing public, improve the market´s informational efficiency, reduce transaction costs and, ultimately, provide access to untapped financial resources belonging to investors and companies, while brightening the long-term investment prospects in the market. The principal rules and initiatives of the NMS Regulation are aimed to reduce "trade-throughs" (order protection rule), to allow the efficient access to quotes by market participants and rules that provide for the efficient dissemination of market data to the investing public. The procedure "trade-throughs" represent the execution of trades at a price that is inferior to the price of a protected quotation, often representing an investor limit order, displayed by another trading center (SEC, 2005). This type of transaction, occurring at a rate of 1 in 11 actively traded Nasdaq stocks, prevents transparency in the price discovery process by precluding investors access to more favorable quotes, namely counterparty limit orders (SEC, 2005, pg.20). To qualify for the application of this rule, the transaction center should be able of displaying and executing trades in under two seconds, requiring that the trading center automate its system in order to ensure that the speed requirement is met. Plausibly, the implementation of this rule (Rule 611) severely hampered the role of the specialist on the NYSE where an automatic order execution system exist, decreasing its ability to influence, among others, opening and closing prices.

Also, the implementation of the access rule (rule 610) in this regulation, which oversees the fair and efficient access to quotes by market participants, allows the disclosure of protected quotes (bid and ask prices of limit orders), allowing an improvement in the informational efficiency and contributing to the narrowing of the capabilities of market makers affecting in own benefit the price discovery, in particular, opening and closing prices. Likely, these procedural and institutional changes will have contributed to the decline and disappearance of the hypothesized night and day return effect in the second part of the sample period.

5. CONCLUSION

This study examines the presence of night and daytime common effects in a set of US equity Exchange-Traded Funds returns series from January 3rd, 1994 to January 3rd, 2014. Two hypothesized manifestations of the effects are examined: the pooled night and daytime effect and the night and daytime effect by day of the week. To examine these effects, classical parametric tests are applied and panel data regressions models are estimated using robust variance-covariance matrix estimators of the models' parameters. Panel data regression models are estimated with standards errors corrected for the presence of within and groupwise heteroscedasticity and serial, contemporaneous and cross-temporal correlation in the least squares residuals of the US equity ETFs. To date, other studies have empirically evaluated the night and daytime effects in the US equity ETFs returns. The most prominent one, the Cliff *et al.* (2008)´s study, use several asset types, including a sample of large ETFs, which replicate the behavior of the principal US stock indices. Their study cover the period from 1995 to 2006 while our goes further and covers the period of 1995 to the beginning of 2014, covering the post implementation period of the NMS Regulation, enacted by the Securities Exchange Commission in 2006, which will have introduced significant changes in the procedures of transactions in the US equity markets.

Regarding the common overall night and daytime effect in the ETFs return series, using panel data regressions with robust variance-covariance matrix estimators, a positive and marginally significant night average return is found while the day average return is not significantly different from zero. Ours findings on the night and daytime effects, by and large, are not consistent with those evidenced in the Cliff *et al.* (2008)'s study.

As regards the hypothesized common night and daytime effect decomposed by day of the week, over the entire sample period and for regressions using robust variance-covariance matrix estimators, only the estimated coefficient for Tuesday-night is significantly positive. For the pooled OLS regression, with standard errors not corrected from departures of the residuals from i.i.d. assumptions, the estimated coefficient for Monday-night is also positive but marginally significant.

These results contrast with those obtained in previous studies using a sample of US equity ETF but with a shorter sampling period. The study carried out by Cliff *et al.* (2008) in a sample of 13 US Equity ETFs, and individually in the SPY ETF, showed a pervasive evidence for the day and night effect overall and decomposed by day of the week in this type of assets, i.e., a significantly positive night and a not so significantly different from zero day average return. Ours findings show a marked decrease or even the disappearance of the day and night effects previously evidenced by Cliff *et al.* (2008).

The disappearance of the effects may imply that US equity markets became gradually more (weak form) efficient as from 2006. Several contributing factors are possible, including the wide supremacy of US markets on a global scale in terms of transaction volume, liquidity and capital admitted to public trading, which will have attracted a growing number of domestic and foreign investors, accompanied by a decrease in transaction costs, especially those related to brokerage and information procurement. We also conjecture that the observed discrepancy between previous findings and ours can be attributed to the legal and institutional changes introduced with the NMS Regulation passed by SEC (2005), which reduced the intervention of specialists in the NYSE trading platforms and the narrowing of the capabilities of market makers affecting in own benefit the price discovery in NASDAQ trading platforms.

6. REFERENCES

Agrawal, A. and Ikenberry D.L. (1994), The individual investor and the weekend effect, *Journal of Finance and Quantitative Analysis*, 29: 263-277.

Amihud, Y. (2002), Illiquidity and stock returns: cross-section and time series effects, J*ournal of Financial Markets*, 5: 31-56.

Baum, C.F., Schaffer, M.E. and Stillman, S. (2010), ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. (http://ideas.repec.org/c/boc/bocode/s425401.html).

Beck, N. and Katz, J.N. (1995), What to do (and not to do) with time-series cross-section data, *American Political Science Review*, 89: 634-647.

Branch, B.S. and Ma, A. (2008), The overnight return, one more anomaly, Available at SSRN.

Breusch, T. and Pagan, A. (1980), The LM test and its application to model specification in econometrics, *Review of Economic Studie*s, 47: 239–254.

Brown, M. B. and Forsythe, A.B. (1974), Robust Tests for the Equality of Variances, *Journal of the American Statistical Association*, 69: 364-367.

Cameron, C., Gelbach, J. and Miller, D. (2006), Robust inference with multi-way clustering, NBER Technical Working Paper, no. 327.

Chang, E.C., Pinegar, J.M. and Ravichandra, R. (1993), International evidence on the robustness of the day of the week effect, *Journal of Financial and Economic Analysis*, 28: 497-513.

Chang, E.C., Pinegar, J.M. and Ravichandra, R. (1998), U.S. day-of-the-week effect and asymmetric responses to macroeconomic news, *Journal of Banking and Finance*, 22: 513-534.

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

Damodaran, A.(1989), The weekend effect in information releases: a study of earnings and dividend announcements, *Review of Financial Studies*, 2: 607-623.

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

Driscoll, J. and Kraay, A. (1998), Consistent covariance matrix estimation with spatially dependent panel data, *Review of Economics & Statistics*, 80: 549-560.

Foster, F.D. and Viswanathan, S. (1990), A theory of interday variations in volume, variance, and trading costs in securities markets, *Review of Financial Studies*, 4: 593-624.

French, K. (1980), Stock returns and the weekend effect, *Journal of Financial Economics*, 8: 55-69.

George, T.J. and Hwang C. (2001), Information flow and pricing errors: a unified approach to estimation and testing, *Review of Financial Studies*, 14: 979-1020.

Greene, W. (2000), *Econometric Analysis*. Upper Saddle River, NJ: Prentice–Hall.

Hoechle, D. (2007), Robust standard errors for panel regressions with cross-sectional dependence, *The Stata Journal,* 7: 281-312.

Hong, H. and Wang, J. (2000), Trading and returns under periodic market closures, J*ournal of Finance*, 55: 297-354.

Jaffe, J.F., Westerfield, R., (1985), "The weekend effect in common stock returns: the international evidence", *Journal of Finance,* 40, 433-454.

Jones, C.M., Kaul, G. and Lipson, M. (1994), Information, Trading and Volatility, *Journal of Financial Economics*, 36: 127-154.

Kelly, M. A. and Clark, S. P. (2011), Returns in trading versus non-trading hours: The difference is day and night, J*ournal of Asset Management*, 12, 132-145.

Kmenta, J. (1986), *Elements of Econometrics*. 2d ed. New York, Macmillan.

Longstaff, F. (1995), How much can marketability affect security value?, *Journal of Finance,* 50: 1767-1774.

Nelson, C.R. and Schwert, G.W. (1982), Tests for predictive relationships between time series variables: a Monte Carlo investigation, *Journal of the American Statistical Association*, 77: 11-18.

Parks, R. (1967), Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated, J*ournal of the American Statistical Association*, 62: 500-509.

Pattel, J. and Wolfson, M. (1982), Good news, bad news, and the intraday timing of corporate disclosures, *The Accounting Review*, 47: 509-527.

Pearce, D. (1996), The robustness of calendar anomalies in daily stock returns, *Journal of Economics and Finance*, 20, 69-80.

Petersen, M. A. (2008), Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies*, 22: 435-480.

Schaffer, M. E. (2010), xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. (http://ideas.repec.org/c/boc/bocode/s456501.html)

Schwert,W. (2002), Anomalies and market efficiency, NBER working article 9277.

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)

Slezak, S. (1994), A theory of the dynamics of security returns around market closures, *Journal of Finance*, 49: 1163-1212.

Thaler, R. (1987), Seasonal movements in security prices II: weekend, holiday, turn of the month and intraday effects, *Journal of Economic Perspectives*, 1, 169-177.

Thompson, S.B. (2011), Simple formulas for standard errors that cluster by both firm and time, *Journal of Financial Economics*, 99: 1-10.

Wang, K., Li, Y. and Erickson, J. (1997), A new look at the Monday effect", *Journal of Finance,* 52: 2171-2186.

Wood,R.A., McInish, T.H. and Ord, J.K. (1985), An investigation of transaction data for NYSE stocks, *Journal of Finance*, 40: 723-739.

Zainudin, A. and Coutts, J.A. (1997), A securities price anomalies in the London International Stock Exchange: a 60 year perspective, *Applied Financial Economics*, 7: 455-464.