Intraday Correlation Patterns Between the S&P 500 and Sector Indices

Michael J. Bommarito II

Department of Financial Engineering, University of Michigan, Ann Arbor

Department of Political Science, University of Michigan, Ann Arbor

Center for the Study of Complex Systems, University of Michigan, Ann Arbor

Abstract

In this brief research paper, I explore patterns in intraday return and volume correlation between the S&P 500 and sector indices, as represented by minutely data from Aug. 23 to Oct. 1 for the SPDR exchange-traded funds. Notably, there is evidence of two previously unreported time-of-day effects. First, there is a "U-shaped" pattern in return correlation, characterized by higher correlation at open and close and lower correlation during mid-day hours. Second, volume correlation is marked by lower values in the morning and increasing values in the afternoon. In some cases, this trend even takes the infamous "hockey-stick" shape, exhibiting stable values in the morning but sharply increasing values in the late afternoon. To ensure that these patterns are not a function of the choice of correlation window size, I confirm that these patterns are qualitatively stable over correlation windows ranging from 10 minutes to 90 minutes. These findings indicate that non-time-stationary patterns exist not only for volume and volatility, as previously reported, but also for the correlation of return and volume between the market and sector indices. These results have possible implications for intraday market efficiency and for trading strategies that rely on intraday time-stationarity of return or volume correlation.

Keywords: return correlation, volume correlation, intraday correlation, sector correlation

1. Introduction

On September 10th, Kristina Peterson of the Wall Street Journal wrote an article asserting what many market participants already felt - market activity has become increasingly concentrated within the first and last moments of the day's trading session ([17]). Though this U-shaped volume phenomenon is not a new finding, Peterson's article claims that this U-shape has become more exaggerated recently. I do not attempt to confirm, deny, or explain this claim of an increasingly exaggerated U-shaped volume pattern. Instead, I would like to ask a related question: do intraday patterns exist in the *correlation* of return or volume within the market?

Previous research on time-of-day effects in volume and volatility have provided evidence of varying market efficiency in equity, fixed-income, and foreign exchange markets. For example, Wood, McInish, and Ord find that the magnitude and standard deviations of NYSE returns are substantially higher in the first and last 30 minutes of market sessions ([19]). Jain and Joh also demonstrate that the intraday NYSE volume follows a similar U-shape ([13]). A number of other authors have confirmed these and other results in equity and equity derivative markets ([3, 14, 18, 15]). In the very liquid index futures markets, Chan and Ekman independently confirm that return volatility exhibits a U-shape ([7, 9]). Intraday patterns have likewise been studied in foreign exchange and fixed-income markets by many authors ([4, 8, 2, 16]. Though there are differences in the type and degree of time-of-day effect, every market has exhibited some intraday pattern.

Email address: mjbommar@umich.edu (Michael J. Bommarito II)

Preprint submitted to SSRN *October 3, 2010*

Though empirical volume and volatility patterns have been thoroughly investigated, there has been comparatively little work on theory to explain these patterns. Brockman and Chung identify three existing theories that address at least intraday liquidity patterns ([6]). In the first theory, differences in the bid-ask spread are a function of time-varying information asymmetries. In the second theory, proposed by Brock and Kleidon, liquidity demand is more inelastic at open and close and liquidity suppliers exploit these differences ([5]). In the third theory, differences are a result of inventory management and the heterogeneity of market participants. A number of author have use these ideas to develop indirect tests of their theories ([1, 10]). However, a number of intraday and interday phenomena remain unexplained by any existing theory.

The motivation for this study is certainly empirical. The results, however, do have important possible theoretical implications. Just as U-shaped intraday patterns in volume and volatility suggest varying levels of market efficiency, the existence of intraday patterns in market correlation may also indicate varying levels of efficiency. If the dependence between two assets exhibits variations that cannot be explain by different information processes, then there may be strong evidence of market inefficiency. From a practical standpoint, there are many arbitrage strategies whose performance could be improved by incorporating intraday correlation patterns. Furthermore, any strong patterns may serve as a caution to theories or strategies that make assumptions about the dynamics of intraday correlation.

The remainder of this paper is organized as follows: in section 2, I describe the dataset constructed for this study and explain the approach for calculating return and volume correlation over varying time scales; in section 3, I present individual results and demonstrate that the claims in the abstract hold across nearly all time scales and assets; in section 4, I conclude with discussion of these results and provide direction for future work.

2. Data & Methods

In order to investigate intraday patterns, I construct a dataset of minutely closing prices and volumes from August 23rd, 2010 to October 1st, 2010 for the following symbols: SPY, XLB, XLE, XLF, XLI, XLP, XLU, XLV, XLY.[1][2] These exchange-traded funds represent some of the most heavily traded securities in the global equity market. Therefore, though they all have some tracking error with respect to their underlying indices, the combination of their liquidity and tradability makes them best suited for this analysis and the application thereof. SPY corresponds to the S&P 500. XLB, XLE, XLF, XLI, XLP, XLU, XLV, and XLY correspond to sector indices for basic materials, energy, financials, industrials, consumer staples, utilities, health care, and consumer discretionary respectively. As a whole, this dataset contains 28 market sessions, 10,938 minutely prices for each asset, and represents market action of over 8.6 billion traded shares and 622 billion dollars of transactions across all 9 assets.

Given this data, I write the price of an asset j in period i as $P_{i,j}$. The number of shares traded on asset j at period i is likewise written as $V_{i,j}$. For the price and volume matrices, period $i = 1$ corresponds to 9:30AM on August 23rd, 2010 and period $i = 10,938$ corresponds to 4:00PM on October 1st, 2010. From this data, I calculate the return matrix R by calculating $R_{i,j} = \log(P_{i+1,j}) - \log(P_{i,j})$ for $i = 1, \ldots, 10,937$.

Next, I calculate the trailing return correlation and volume correlation between a sector and the S&P 500 for a specified correlation window τ. Formally, I calculate this trailing correlation coefficient $\rho_{A,B}(i)$

[1]This data was obtained from the Power E*Trade Pro trading platform, a product of E*TRADE Securities LLC.

[2]XLK, the Select Sector SPDR ETF tracking the technology sector, is excluded due to a data issue caused by the author's mistake.

between assets A and B over a window of size τ at period i as

$$\mu_j(i) = \frac{1}{\tau} \sum_{t=i-\tau+1}^{i} R_{t,j}$$

$$\sigma_j^2(i) = \frac{1}{\tau - 1} \sum_{t=i-\tau+1}^{i} (R_{t,j} - \mu_j(i))^2$$

$$\rho_{A,B}(i) = \frac{1}{\tau - 1} \frac{\sum_{t=i-\tau+1}^{i} (R_{t,A} - \mu_A)(R_{t,B} - \mu_B)}{\sigma_A \sigma_B}$$

There are two important points to make with respect to these equations. The first is to note that ρ contains the covariance of the assets in the numerator and the volatility of the assets in its denominator. In other words, patterns in correlation can be driven both by changes in the actual covariance and/or by changes in the volatility of the assets. As discussed above, researchers previously uncovered U-shaped intraday patterns in volume and volatility. Given this fact, if a U-shaped pattern exists in intraday correlation, it could be the result of two possible cases. In the first case, the covariance of the S&P 500 and sector indices is constant intraday and any non-stationarity is driven by the known non-stationarity in the denominator. In the second case, the covariance between the S&P 500 and sector indices may not be constant and the pattern in covariance may interact directly with any patterns in volatility.

If a right-side-up U-shape is observed in correlation, this means that the covariance is not only U-shaped itself, but that covariance's U-shape is at least an order of magnitude greater than volatility's U-shape. This can be understood by noting the U-shape of volatility in the denominator of correlation would result either in an upside down U-shape ($\cap$) if covariance were constant or a relatively flat line ($-$) if covariance's U-shape is on the same order of magnitude. Therefore, the existence of a U-shape in correlation should indicate an entirely different phenomenon than U-shapes previously noted in volatility. This intraday U-shape in correlation is a much more damning indication of market inefficiency than U-shapes in liquidity, as it indicates that structural relationships within the market are not stable within the day.

The second point to emphasize in these equations is that τ is an important parameter. τ corresponds to the number of samples used in the correlation calculation and can be thought of as the "memory" of the estimate. Smaller values of τ are less strong statistically but more responsive to changes in correlation over time. Larger values of τ yield smoother, more reliable estimates but take more periods to show real changes in correlation if the underlying parameter varies over time. It is also important to ensure that correlation is only calculated within a day, not across sessions, as this would include jumps from the market close to the next market open. Accordingly, the first τ periods of each session are thrown out so as to not include data from the previous market session. The result is a matrix of return correlations over time, written C^R, and a matrix of volume correlations over time, written C^V. $C_j^R(i)$ corresponds to the correlation of returns between sector j and the S&P 500 at period i.[3] $C_j^V(i)$ likewise corresponds to the correlation of volumes between sector j and the S&P 500 at period i.

Figure 1 plots the behavior of C_{XLB}^R for $\tau = 10$ on the left and $\tau = 60$ on the right over the entire time period of the dataset. Figure 2 likewise plots the behavior of C_{XLB}^V for $\tau = 10$ on the left and $\tau = 60$ on the right over the same interval. These figures illustrate the importance of τ as described above. Smaller values of τ result in correlation values that are much more dynamic but sometimes harder to interpret. Larger values of τ provide smoother, more easily understood pictures of correlation but may miss phenomena that occur on small time scales. It is important to understand that some statements may be true for some values of τ but not others.

In order to detect intraday patterns, these matrices must now be processed to correspond to hours of

[3] If i is one of the first τ periods of a market session, this value is coded as NaN and is not included in later statistical analysis. Note that this implies that the first hours of the market session may contain fewer samples than later hours. Values may also be coded as NaN in the event that the variance of either the S&P 500 or the sector index is 0 over the current correlation window.

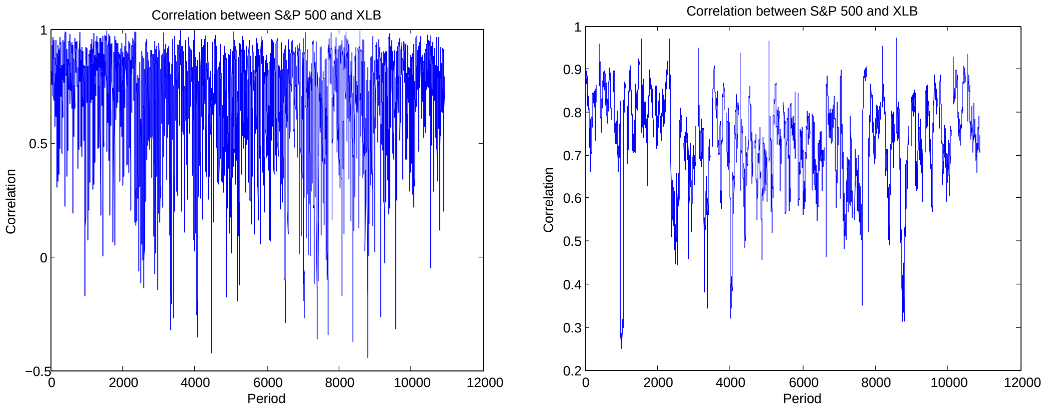

Figure 1: Return correlation between SPY and XLB (C^R_{XLB}) for $\tau = 10$ (left) and $\tau = 60$ (right).

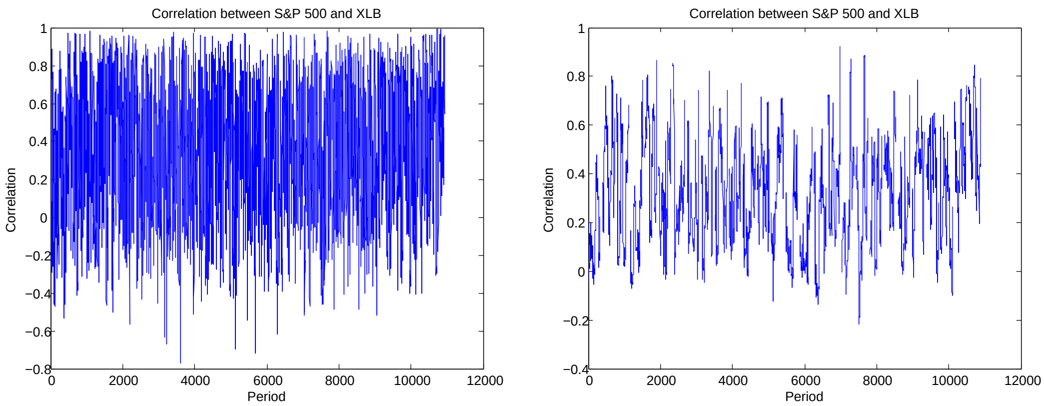

Figure 2: Volume correlation between SPY and XLB (C^V_{XLB}) for $\tau = 10$ (left) and $\tau = 60$ (right).

the day. To do so, I bin each value of C^R and C^V by the hour in which the period occurred.[4] This binning allows us to look at correlation dynamics that appear to be dependent on the hour of the day in which they occur. Each binned hour contains 60 minutes from 28 market sessions, yielding 1680 samples per bin.[5]

3. Results

Figure 3 shows the pattern of hourly return correlation ($\tau = 60$) over the day. Each point in the figure represents the mean correlation for that hour, averaged across all sector indices. The bars in the figure give mean standard deviation intervals above and below the mean correlation point. Though no hour has a statistically different mean from another hour, the U-shaped pattern is clearly visible for $\tau = 60$.

Figure 4 shows that this U-shape is also present for values of $\tau \in \{10, 30, 90\}$. In order to put numbers to this shape, I compare the correlation of the bottom of the U with the sides of the U by comparing the mean correlation around the lunch hour with the correlation away from the lunch hour. Here, I take the lunch period to be one hour on either side of noon, adjusted for the window size τ - in other words, $\lfloor (12 + \frac{\tau}{60} \rfloor \pm 1$.

[4]For example, values of C^R at 3:00PM on September 2nd and 3:59PM on September 9th are contained in the same 3PM (15) bin.

[5]The number of samples in the 9AM and 10AM bins depends on the value of τ. For example, if $\tau = 30$, then the 9AM bin contains 0 samples and the 10AM bin contains 1680 samples. If $\tau = 60$, the 9AM bin contains 0 samples and the 10AM bin now only contains $28 \cdot 30 = 840$ samples.

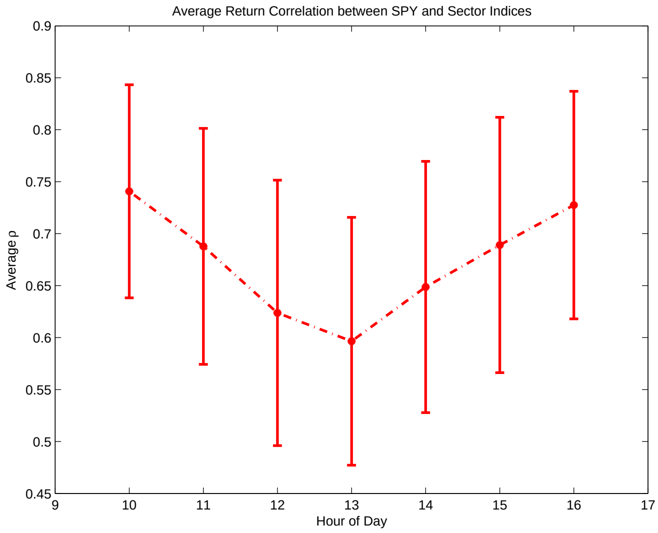

Figure 3: Return correlation ($\tau = 60$) to SPY, averaged over all sector indices. Error bars indicate the mean standard deviation.

The results are shown in Table 1 and clearly demonstrate that the mean of the mid-day correlation is lower than the mean of the opening and closing hours' correlation across all sector indices.

A two-sample Kolmogorov-Smirnov test rejects the possibility that the distributions are identical at the $\alpha = 0.001$ level for every sector for $\tau = 10, 30, 60, 90$. Furthermore, the Wilcoxon rank sum test rejects the hypothesis that the distributions have identical mean at the $\alpha = 0.001$ level for all sectors for $\tau = 30, 60, 90$ ([12]). For $\tau = 10$, the hypothesis is rejected for five of eight sectors. Though not all of these differences are statistically significant, the trend as a whole is present across all sectors and all time scales.[6]

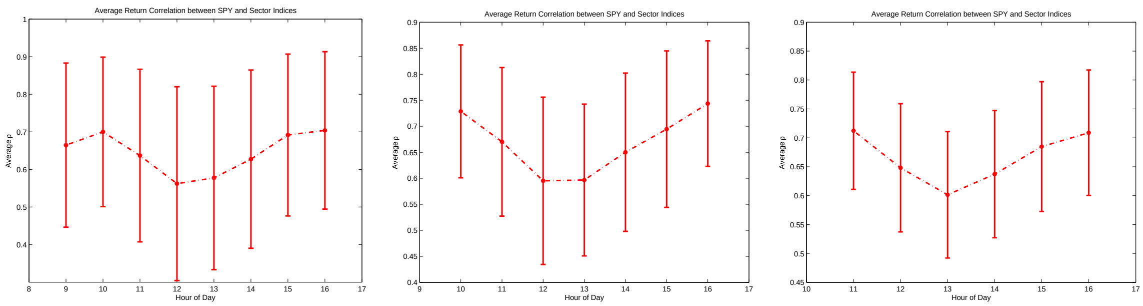

Figure 4: Return correlation to SPY, averaged over all sector indices. Error bars indicate the mean standard deviation. Left: $\tau = 10$, Center: $\tau = 30$, Right: $\tau = 90$.

The results above clearly indicate that a U-shaped pattern exists in the correlation of *returns* between between the S&P 500 and sector indices. One might wonder whether intraday patterns also exist in the correlation of *volumes* between the S&P 500 and sector indices. This analysis proceeds just as above, except instead of calculating the equations above on the return matrix R, the mean, standard deviation, and correlation of the volume matrix V are calculated. Figure 2 above shows sample behavior of volume correlation.

Figure 5 shows the behavior of the volume correlation with a trailing hourly ($\tau = 60$) window over the course of a day, just as Figure 3 demonstrated for return correlation above. These calculations are also

[6]See Figures 7 through 11 in the appendix for all figures, including figures for all values of τ and separate for all sector indices.

τ	**Mid-Day Correlation**	**Open/Close Correlation**
5	0.559	0.641
10	0.592	0.677
30	0.621	0.704
60	0.623	0.711
90	0.629	0.702

Table 1: Return correlation comparison between Mid-Day and open/close.

averaged over all sector indices and error bars show mean standard intervals. Unlike Figure 3, however, this volume correlation plot indicates that the later afternoon hours have higher correlation than earlier session hours. Figure 6 shows that this trend is also quite clear for values of $\tau \in \{10, 30, 90\}$.

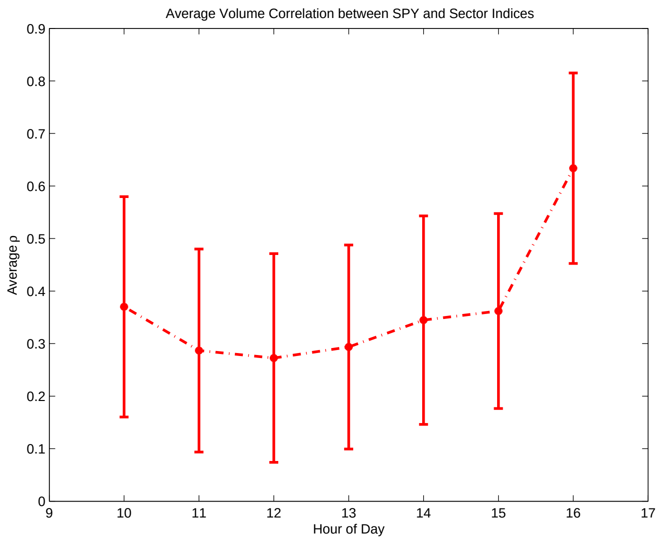

Figure 5: Volume correlation ($\tau = 60$) to SPY, averaged over all sector indices. Error bars indicate the mean standard deviation.

This trend, marked by stable correlation values in the first half of the market session and sharply rising values in the second half of the market session, is often described as a "hockey stick." The hockey stick is incredibly clear for lower values of τ as seen in Figure 6. In order to more rigorously describe this trend, I put numbers to the phenomenon by calculating the means for both the morning and afternoon correlations. The results are shown in Table 2.

A two-sample Kolmogorov-Smirnov test rejects the possibility that the morning and afternoon distributions are identical at the $\alpha = 0.001$ level for all but one sector for $\tau = 10, 30, 60, 90$. XLU, the utilities sector ETF, was only significant at the $\alpha = 0.005$ level for $\tau = 10$ and 60. Furthermore, the Wilcoxon rank sum test rejects the hypothesis that the morning and afternoon distributions have equal median at the $\alpha = 0.001$ level for nearly all sectors for $\tau = 10, 30, 60, 90$. Again, XLU is only significant at the $\alpha = 0.01$ level for $\tau = 10$ and 90. The test cannot reject the hypothesis for any reasonable value of α for $\tau = 30$ and 60. The test also cannot reject the hypothesis for XLY, the consumer discretionary sector ETF, for $\tau = 60$ and 90. Though the not every difference between morning and afternoon sessions is significantly different, the trend is convincing when considering that it holds qualitatively true across all sectors and values of τ. There are a number of possible explanations for this trend in correlation including re-balancing arbitrage and concentrated market interest.

τ	**Morning Correlation**	**Afternoon Correlation**
5	0.314	0.467
10	0.316	0.464
30	0.312	0.425
60	0.310	0.409
90	0.294	0.396

Table 2: Volume correlation comparison between morning and afternoon.

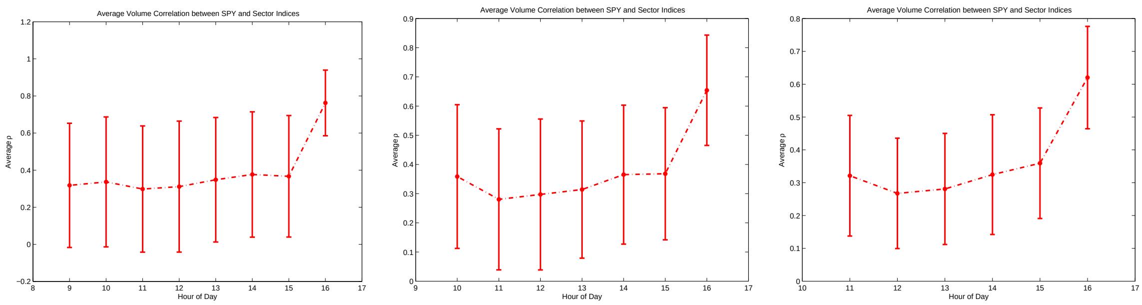

Figure 6: Volume correlation to SPY, averaged over all sector indices. Error bars indicate the mean standard deviation. Left: $\tau = 10$, Center: $\tau = 30$, Right: $\tau = 90$.

4. Discussion

The results contained above provide evidence for the existence of two striking and previously unreported patterns. First, the correlation between the S&P 500 and sector indices is not time-stationary within a day. This is the result of an extreme U-shape in covariance and is a phenomena separate from the previously observed U-shape in volatility. Furthermore, this U-shape is possibly evidence of market inefficiency. Whereas market liquidity or volatility may vary with demand elasticity, information asymmetries, and inventory management, the structural relationship between the S&P 500 and market sectors should not exhibit intraday patterns. The second phenomenon observed above is that the correlation of volume between the S&P 500 and sectors indices exhibits lower correlation in the morning and rapidly increasing correlation in the afternoon. This trend even takes the infamous “hockey stick” shape for some sector indices and values of τ. Both of these results have implications for strategies that are based on the assumption that correlation between sector indices and the market are time-stationary within a day.

This brief note has a number of shortcomings. Foremost among them is that the data is taken from a small period of time in the market and includes only 28 market sessions. Therefore, the results reported herein may represent dynamics that are local to this period of time. This does, however, imply that the results are immediately relevant and do not include what might be described as “less relevant” data. The rapid growth of high-frequency trading has likely altered market microstructure over the past five years. Second, the dataset is based on exchange-traded funds. Exchange-traded funds exhibit a number of peculiarities that lead to possible tracking error. However, exchange-traded funds are by far the best tradable, highly liquid proxies to the S&P 500 and market sectors that are available. Third, the volume correlation results are likely affected by the previously observed U-shape in volume. I hope to disentangle this relationship in future work by detrending the data for previously observed intraday patterns. In future work, I hope to expand upon this dataset, provide a more full description and model of the return and volume correlation processes, and include other asset class indices such as fixed-income or commodities.

Acknowledgements

I would like to thank the Center for the Study of Complex Systems (CSCS) at the University of Michigan for a fruitful research environment. This work was partially supported by an NSF-IGERT fellowship through

the Center for the Study of Complex Systems (CSCS) at the University of Michigan, Ann Arbor.

References

[1] A. Admati, P. Pfleiderer. *A theory of intraday patterns: Volume and price variability.* Review of Financial Studies 1:3-40, 1988.

[2] T. Andersen, T. Bollerslev. *Deutsche MarkDollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies.* Journal of FInance, 53:219-265, 2002.

[3] A. Atkins, S. Basu. *The effect of after-hours announcements on the intraday U-shaped volume pattern.* Journal of Business Finance and Accounting, 22:789-809, 1995.

[4] R. Baillie, T. Bollerslev. *Intra-day and inter-market volatility in foreign exchange rates.* Review of Economic Studies, 58:565-585, 1990.

[5] W. Brock, A. W. Kleidon. *Periodic market closure and trading volume: a model of intraday bids and asks.* Journal of Econ. Dyn. Control, 16:451489, 1992.

[6] P. Brockman, D. Chung. *Inter- and intra-day liquidity patterns on the Stock Exchange of Hong Kong.* Journal of International Financial Markets, Institutions, and Money, 8:277-298, 1998.

[7] K. Chan, K.C. Chan, G. A. Karolyi. *Intraday volatility in the stock index and stock index futures markets.* The Review of Financial Studies, 4:657-684, 1991.

[8] M. Cornett, T. Schwarz, A. Szakmary. *Seasonalities and intraday return patterns in the foreign currency futures market.* Journal of Banking and Finance, 19:843-869, 2000.

[9] P. Ekman. *Intraday patterns in the S&P500 index futures market.* Journal of Futures Markets, 12:365-81, 1992.

[10] F. D. Foster, S. Viswanathan, *A theory of interday variations in volumes, variances and trading costs in securities markets.* Review of Financial Studies, 3:593624, 1990.

[11] L. Harris. *A transaction data survey of weekly and intraday patterns in stock patterns.* Journal of Financial Economics 16:99117, 1986.

[12] M. Hollander, D. A. Wolfe. *Nonparametric Statistical Methods.* Hoboken, NJ: John Wiley and Sons, Inc., 1999.

[13] P. Jain, G. Joh. *The dependence between hourly prices and trading volume.* Journal of Financial and Quantitative Analysis, 23:269-84, 1988.

[14] L. Lockwood, S. Linn. *An examination of stock market return volatility during overnight and intraday periods, 1964-1989.* Journal of Finance, 45:591-601, 1990.

[15] T. McInish, R. Wood. *An analysis of intraday patterns in bid/ask spreads for NYSE Stocks.* Journal of Finance, 47:753-64, 1992.

[16] P. Pasquariello. *The Microstructure of Currency Markets: An Empirical Model of Intra-day Return and Bid-Ask Spread Behavior.* Working Paper, 2001.

[17] K. Peterson. *The Traders Who Skip Most of the Day*, Wall Street Journal. Sep. 10, 2010. `http://online.wsj.com/article/SB10001424052748704392104575475781704072278.html`.

[18] A. Sheikh, E. Ronn. *A Characterization of the Daily and Intraday Behavior of Returns on Options.* Journal of Finance, 49:557-579, 1994.

[19] R. Wood, T. McInish, J. Ord. *An investigation of intraday data for NYSE stocks.* Journal of Finance 40:72374, 1985.

Appendix

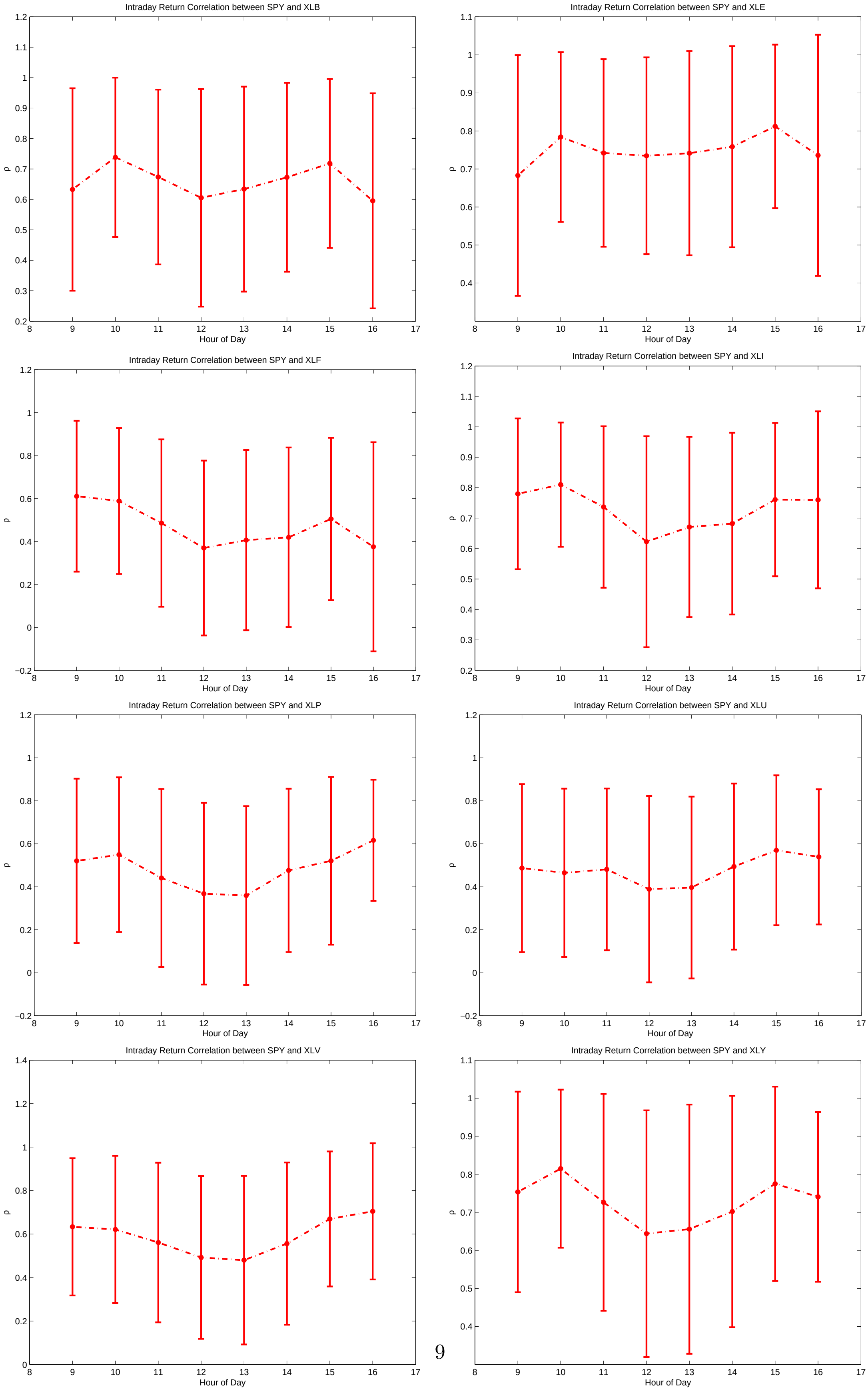

Figure 7: Panel of intraday return correlation between S&P 500 and sector indices for window size $\tau = 5$.Error bars indicate one standard deviation.

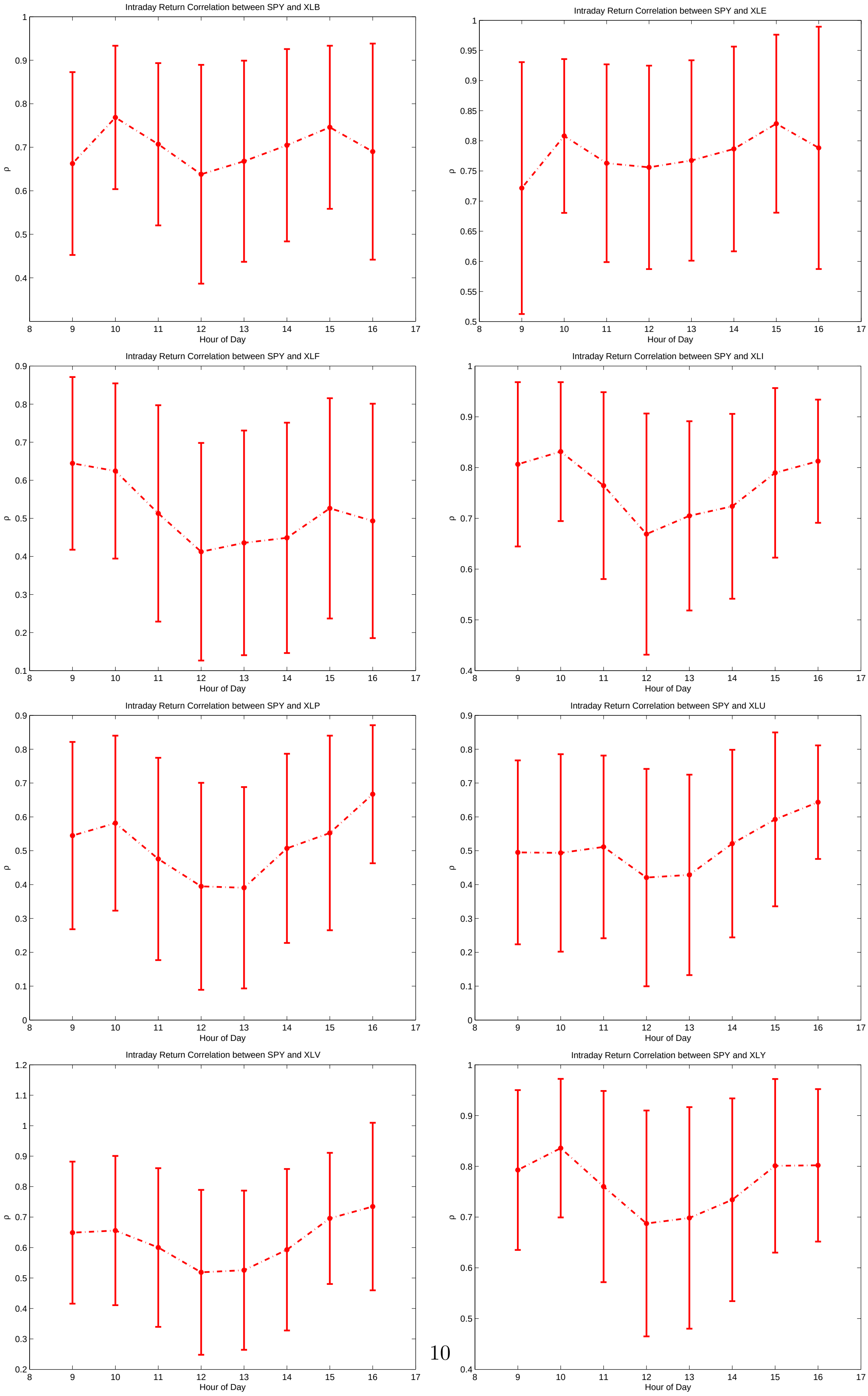

Figure 8: Panel of intraday return correlation between S&P 500 and sector indices for window size $\tau = 10$.Error bars indicate one standard deviation.

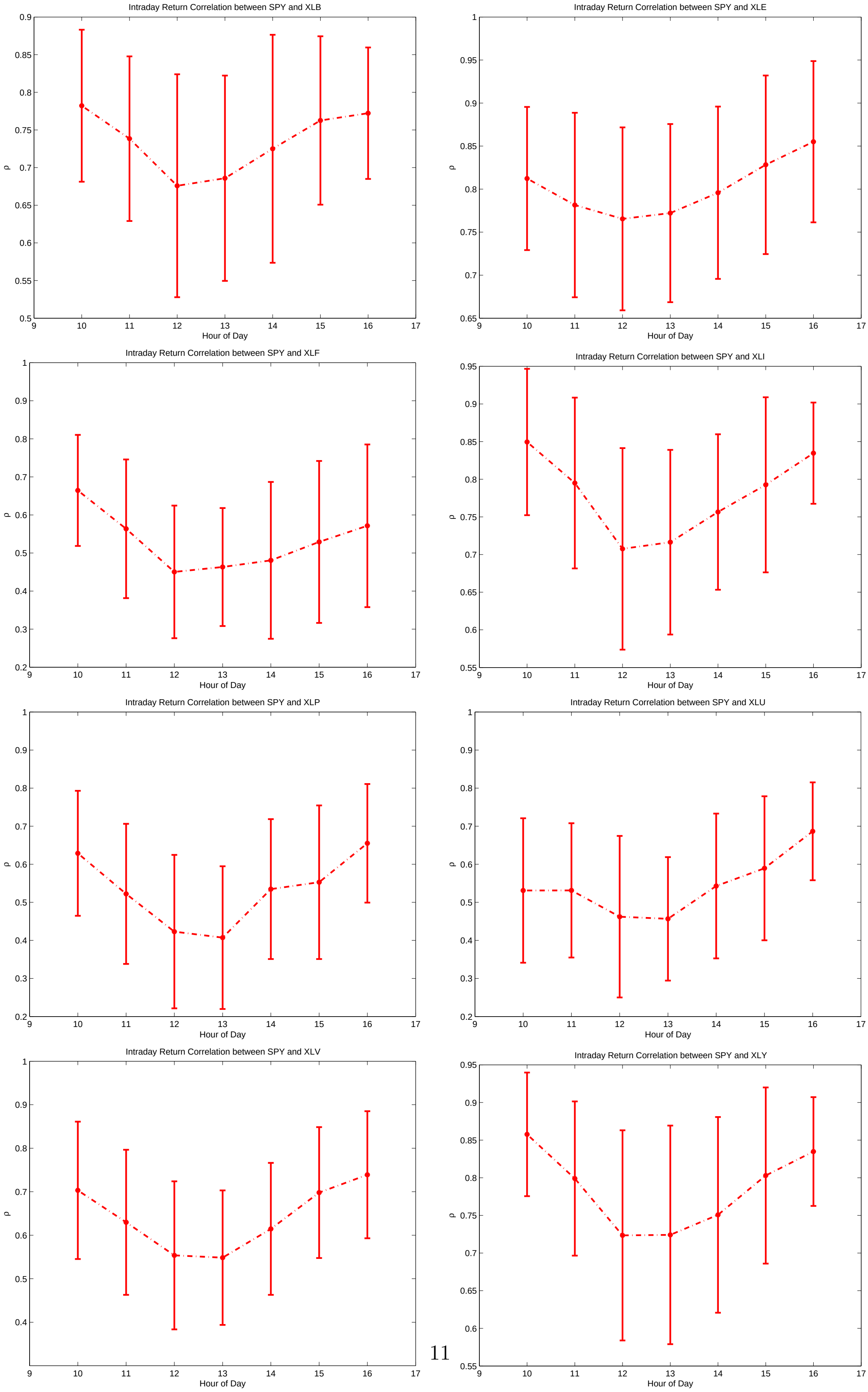

Figure 9: Panel of intraday return correlation between S&P 500 and sector indices for window size $\tau = 30$.Error bars indicate one standard deviation.

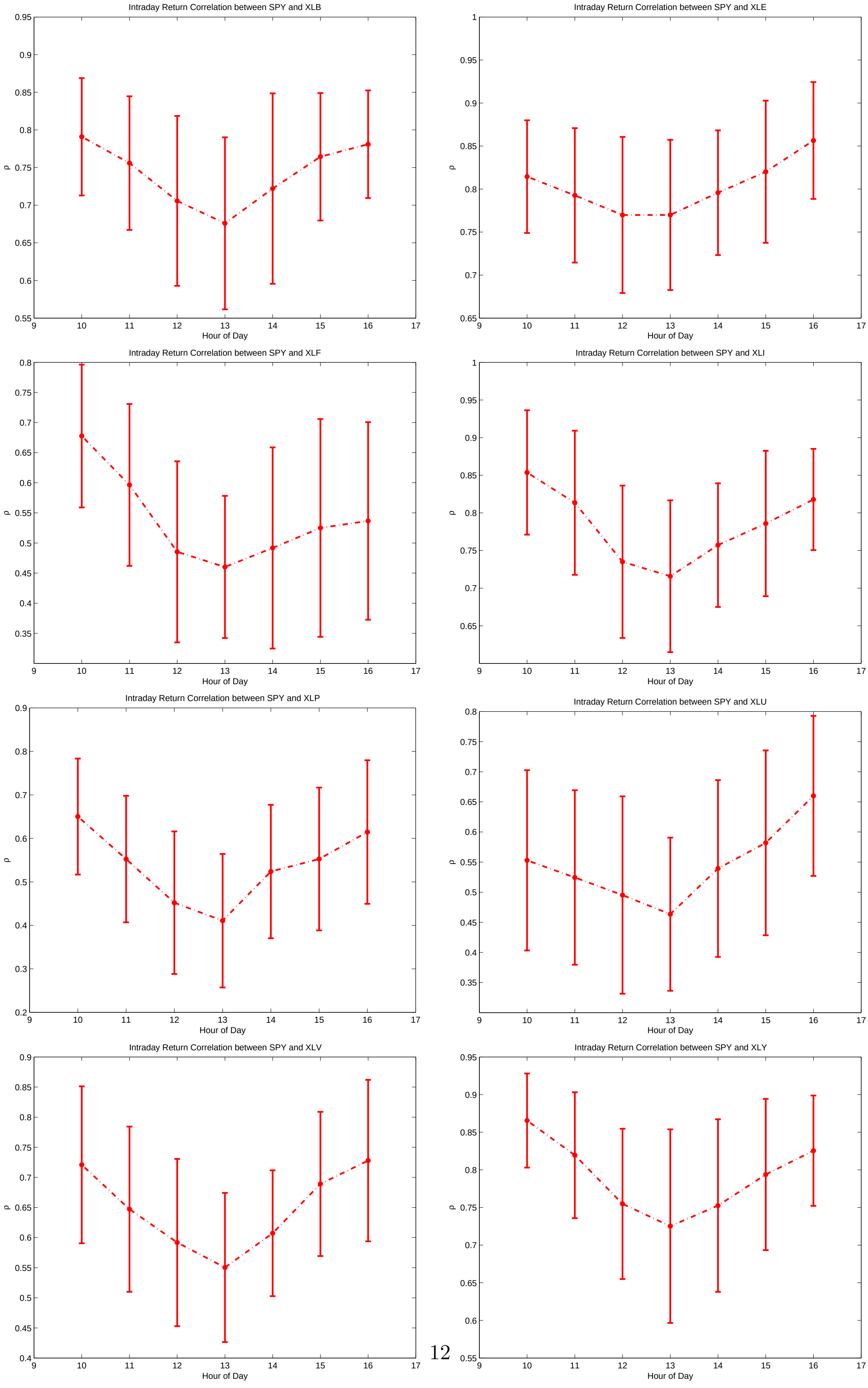

Figure 10: Panel of intraday return correlation between S&P 500 and sector indices for window size $\tau = 60$.Error bars indicate one standard deviation.

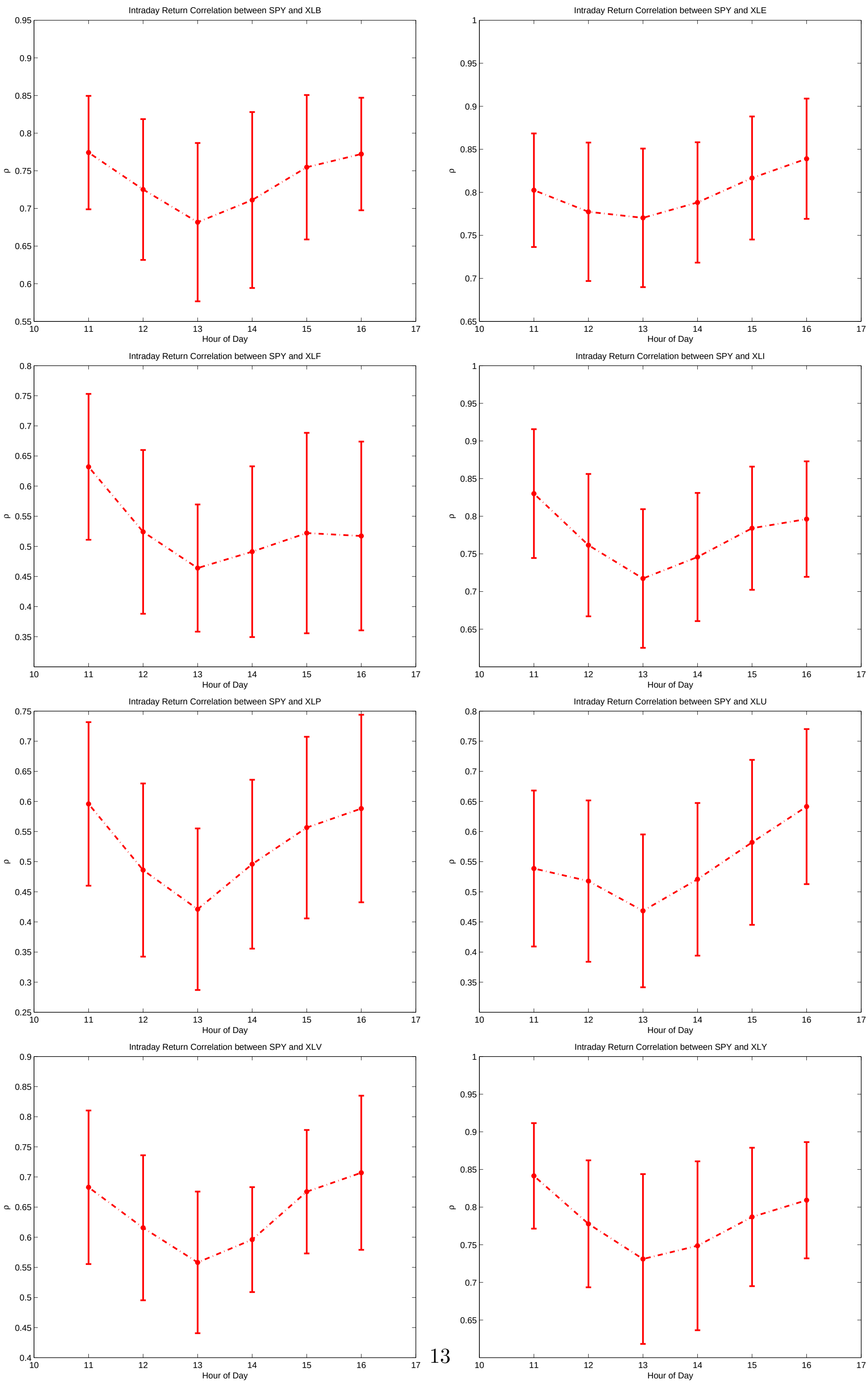

Figure 11: Panel of intraday return correlation between S&P 500 and sector indices for window size $\tau = 90$.Error bars indicate one standard deviation.

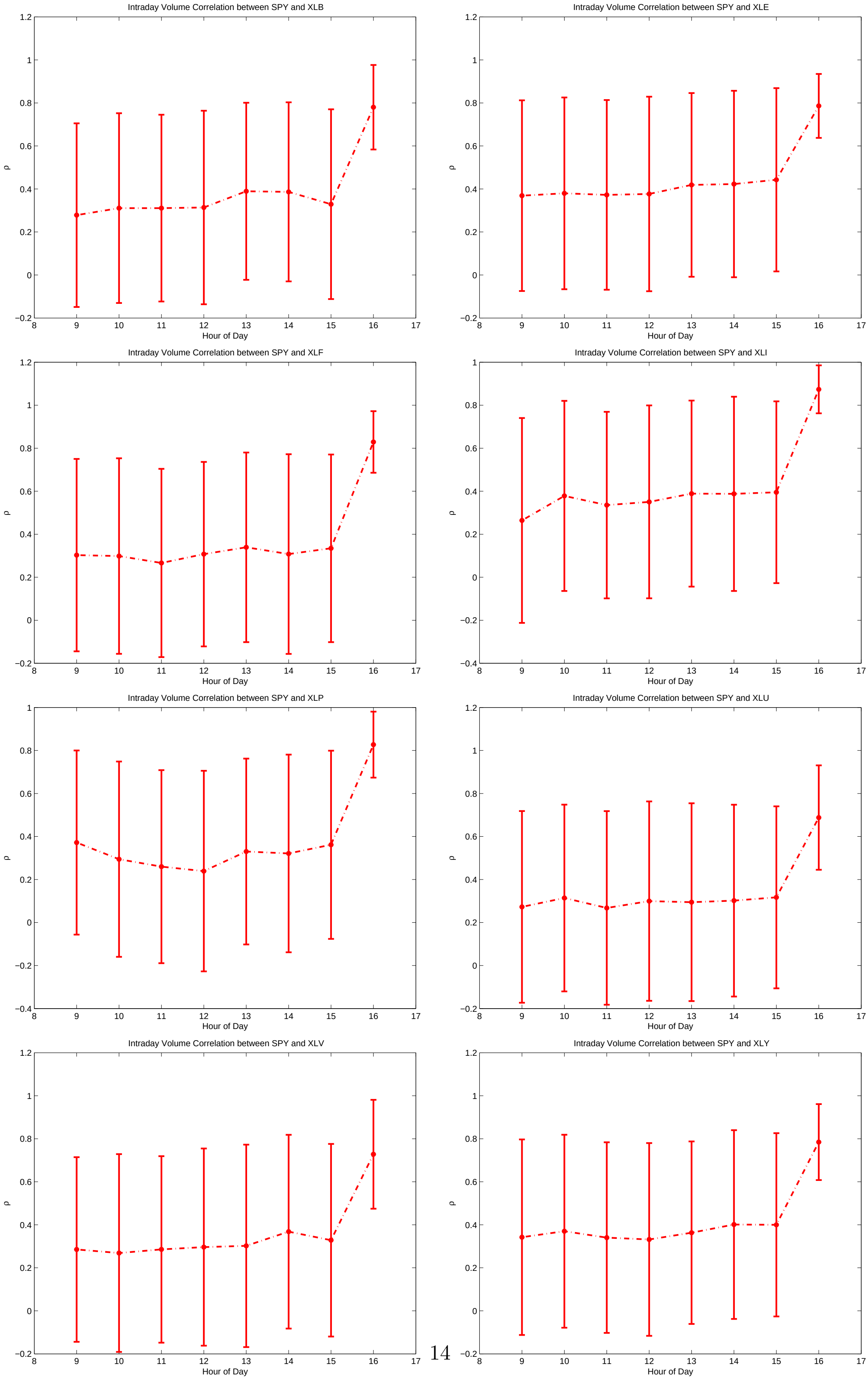

Figure 12: Panel of intraday volume correlation between S&P 500 and sector indices for window size $\tau = 5$.Error bars indicate one standard deviation.

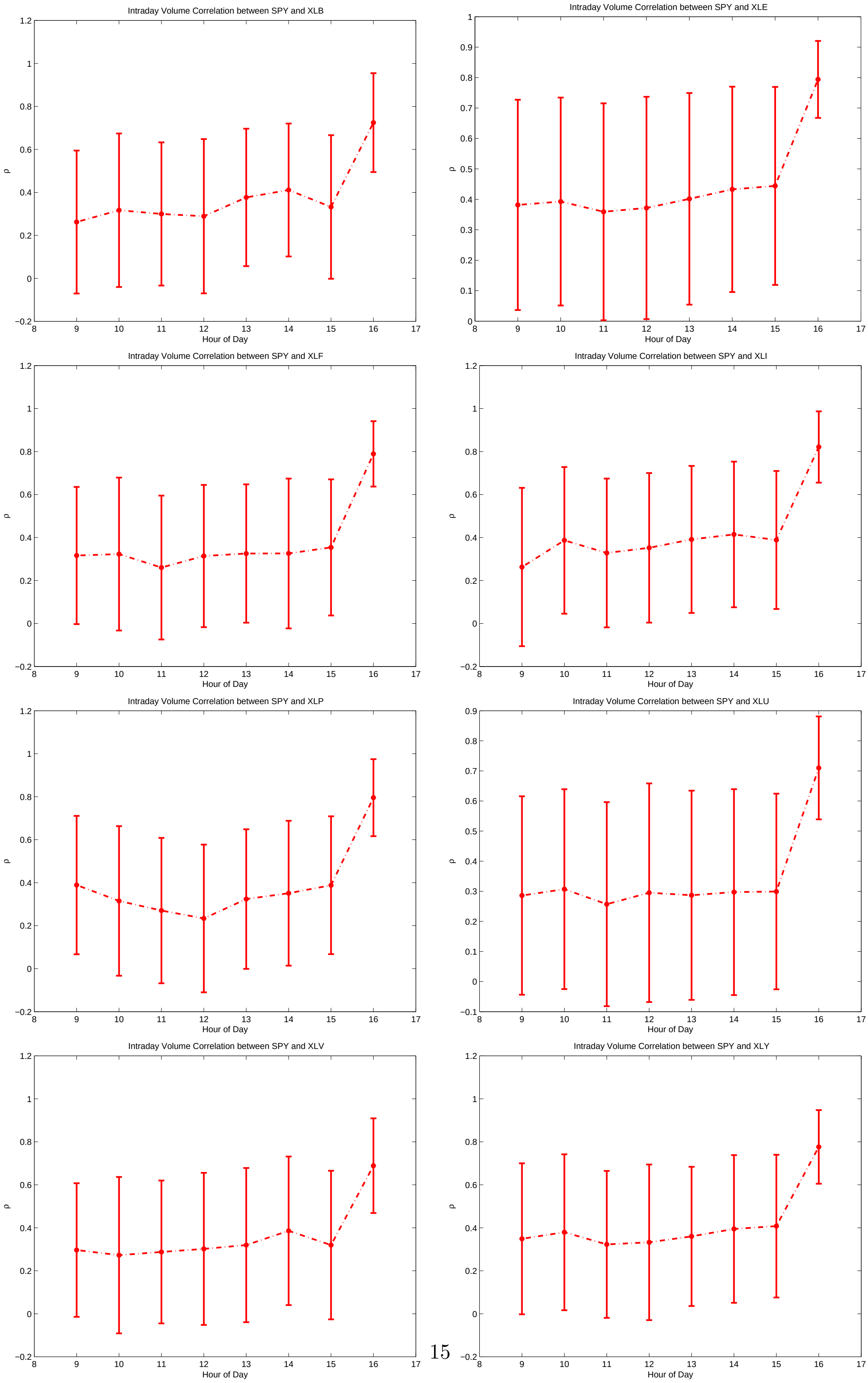

Figure 13: Panel of intraday volume correlation between S&P 500 and sector indices for window size $\tau = 10$.Error bars indicate one standard deviation.

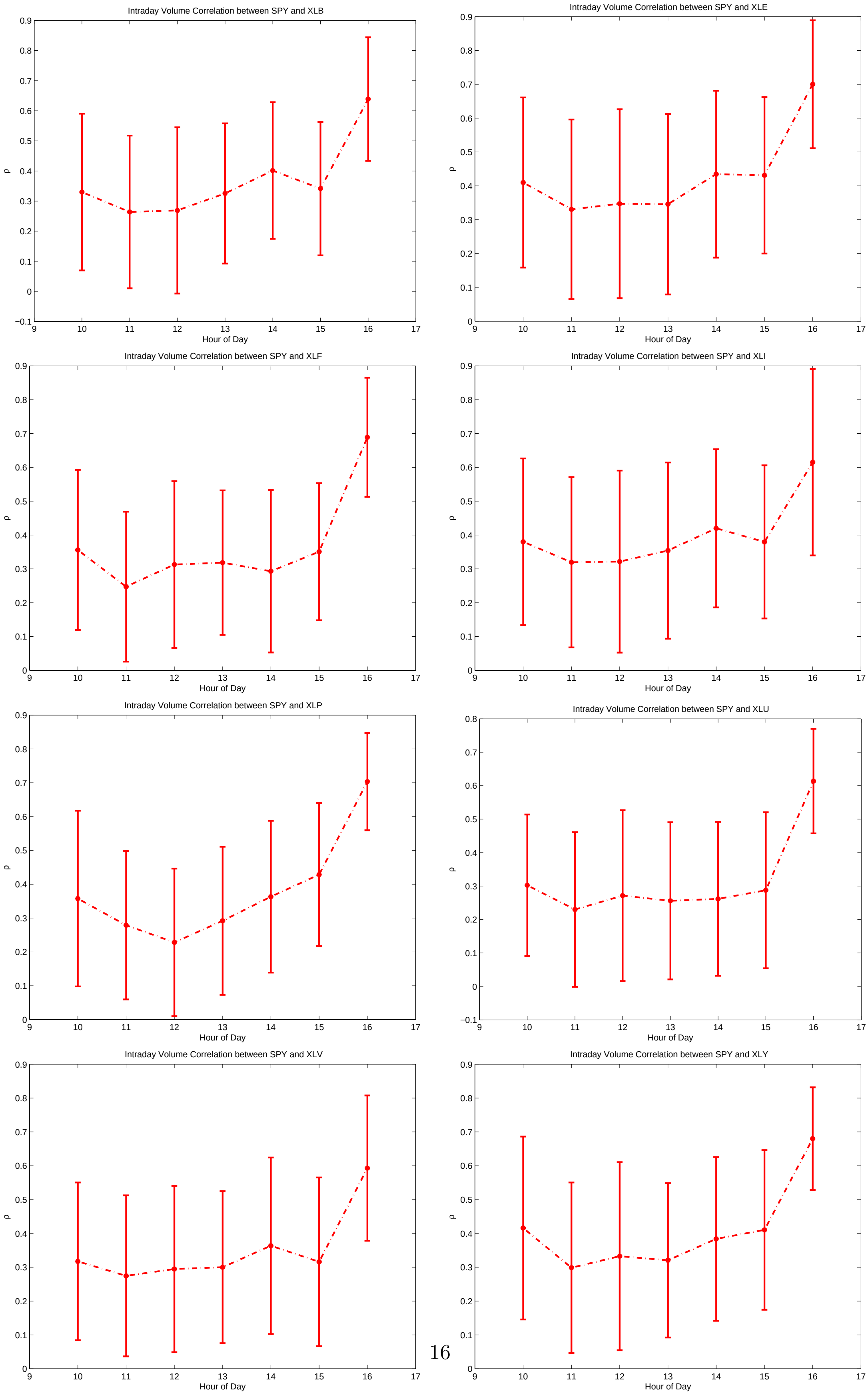

Figure 14: Panel of intraday volume correlation between S&P 500 and sector indices for window size $\tau = 30$.Error bars indicate one standard deviation.

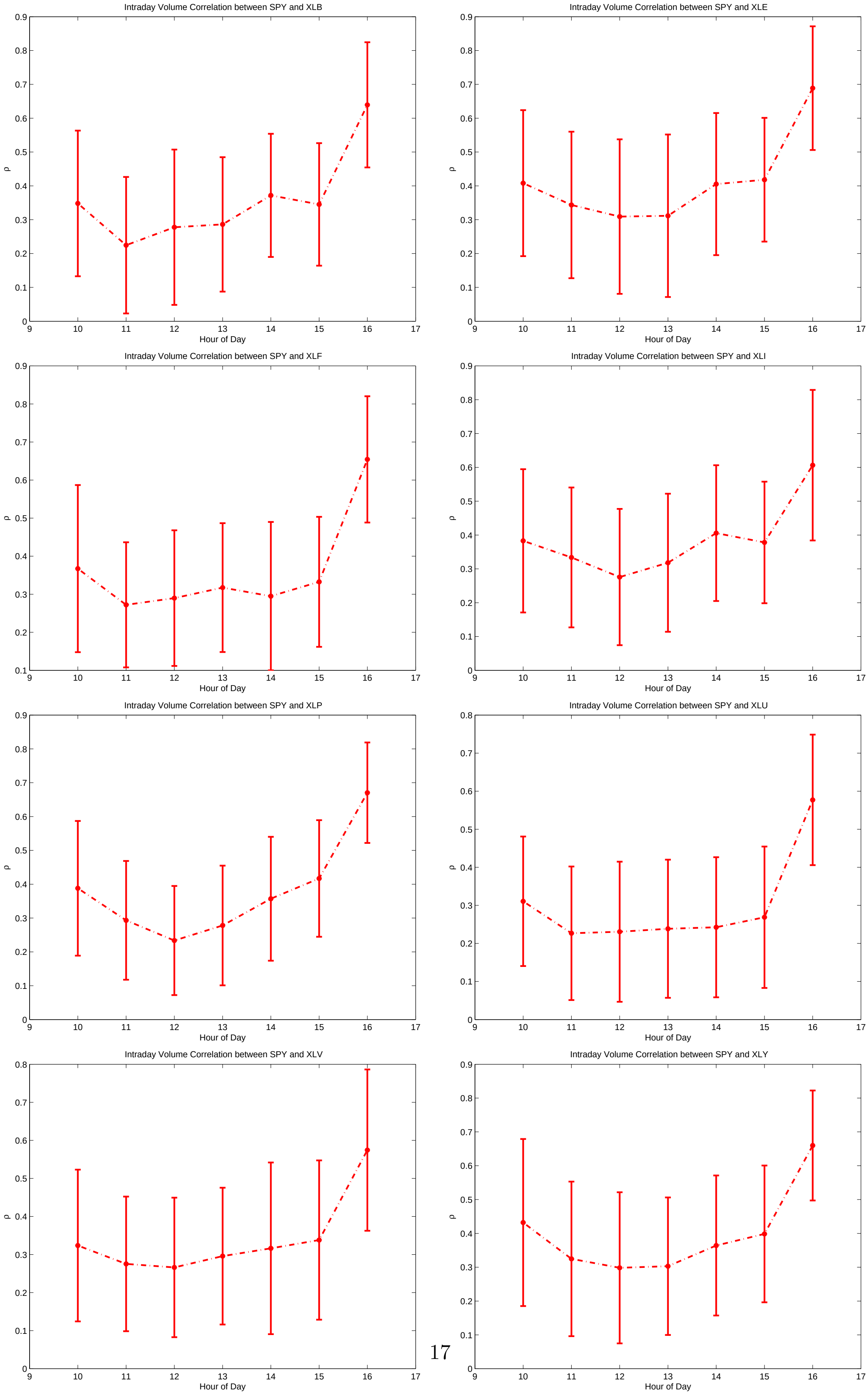

Figure 15: Panel of intraday volume correlation between S&P 500 and sector indices for window size $\tau = 60$.Error bars indicate one standard deviation.

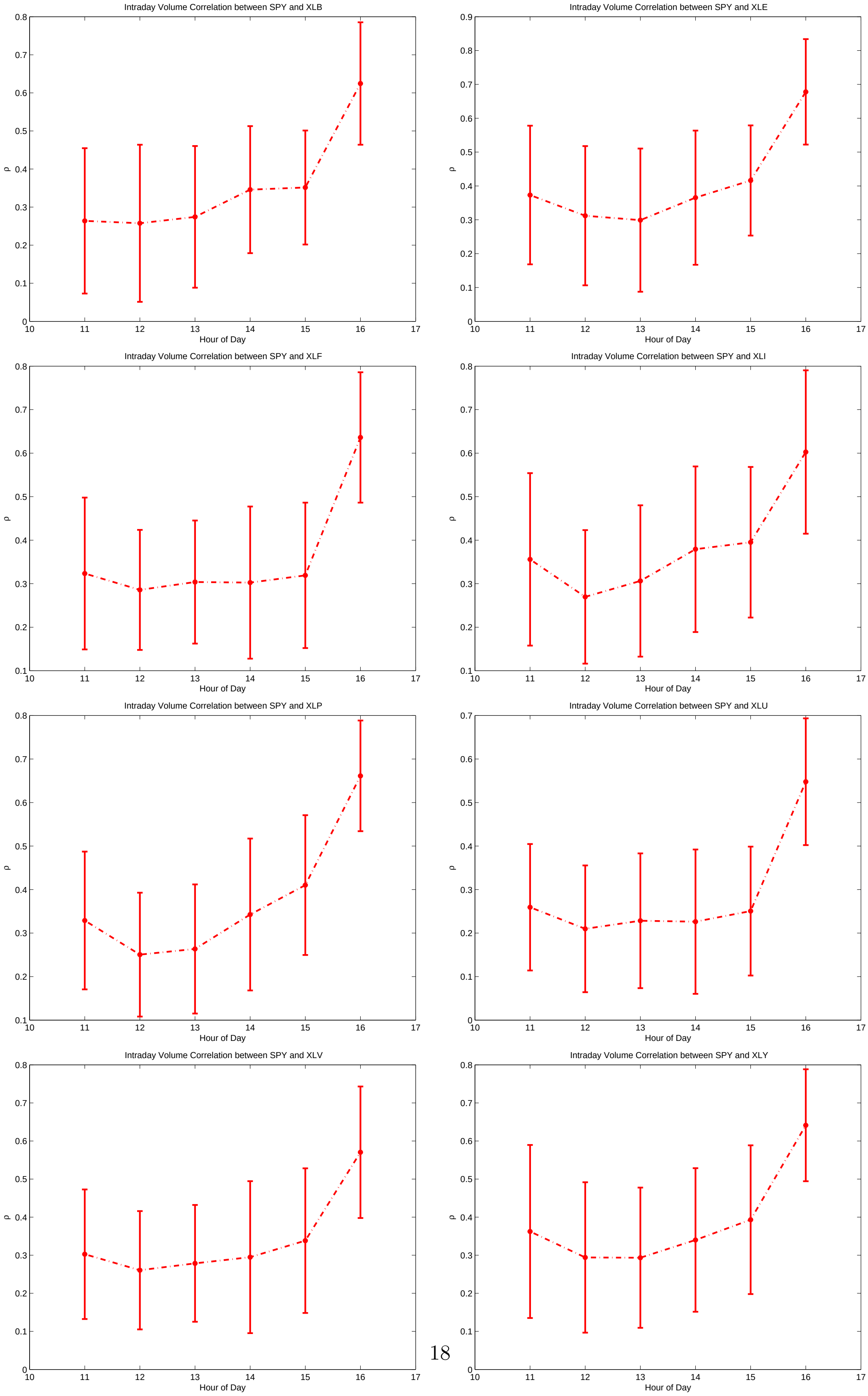

Figure 16: Panel of intraday volume correlation between S&P 500 and sector indices for window size $\tau = 90$.Error bars indicate one standard deviation.