Overnight Returns & Intraday Returns. Are they Correlated?

Faisal Nazir Zargar¹, Suhail Qasim Mir² and Ishaq Ahmad Bhat³.

¹,²,³MBA Student, Centre for Management Studies, Jamia Millia Islamia, New Delhi.

Keywords: Efficient Market Hypothesis (EMH), Overnight Returns, Intraday Returns, Close to Close Returns, Augmented Dickey–Fuller test (ADF), Correlation, Regression.

Revolutions often spawn counterrevolutions and the efficient market hypothesis in finance is no exception. The intellectual dominance of the efficient-market revolution has more been challenged by economists who stress psychological and behavioral elements of stock-price determination and by econometricians who argue that stock returns are, to a considerable extent, predictable. An efficient market (weak form) will contain no significant price pattern, a view supported by numerous empirical studies. This study, however, reveals a very strong negative autocorrelation between overnight and intraday returns, and a strong negative cross correlation between overnight and intraday returns.

A generation ago, the efficient market hypothesis was widely accepted by academic financial economists. It was generally believed that securities markets were extremely efficient in reflecting information about individual stocks and about the stock market as a whole. The accepted view was that when information arises, the news spreads very quickly and is incorporated into the prices of securities without delay. Thus, neither technical analysis, which is the study of past stock prices in an attempt to predict future prices, nor even fundamental analysis, which is the analysis of financial information such as company earnings, asset values, etc., to help investors select “undervalued” stocks, would enable an investor to achieve returns greater than those that could be obtained by holding a randomly selected portfolio of individual stocks with comparable risk.

The efficient market hypothesis is associated with the idea of a “random walk,” which is a term loosely used in the finance literature to characterize a price series where all subsequent price changes represent random departures from previous prices. The
logic of the random walk idea is that if the flow of information is unimpeded and information is immediately reflected in stock prices, then tomorrow’s price change will reflect only tomorrow’s news and will be independent of the price changes today. But news is by definition unpredictable and, thus, resulting price changes must be unpredictable and random. As a result, prices fully reflect all known information, and even uninformed investors buying a diversified portfolio at the tableau of prices given by the market will obtain a rate of return as generous as that achieved by the experts. By the start of the twenty-first century, the intellectual dominance of the efficient market hypothesis had become far less universal. Many financial economists and statisticians began to believe that stock prices are at least partially predictable. A new breed of economists emphasized psychological and behavioral elements of stock-price determination, and came to believe that future stock prices are somewhat predictable on the basis of past stock price patterns as well as certain “fundamental” valuation metrics. Moreover, many of these economists were even making the far more controversial claim that these predictable patterns enable investors to earn excess risk-adjusted rates of return. Our study of the daily returns for stocks listed on the NSE reveals a potentially anomalistic relationship that may have significant market implications. According to the weak form of the efficient market hypothesis, which is supported by numerous studies, a time series of returns should either be uncorrelated or the magnitude of any possible correlation should be too small to be financially meaningful. Our finding, that overnight stock returns are strongly auto correlated with subsequent intraday returns, presents evidence that is inconsistent with the weak form of the efficient market hypothesis.

Literature Review
A number of issues related to the relationship between overnight and intraday returns are considered in the market microstructure literature. For example, Fuertes, Kalotychov & Natasha (2011) suggest that the information in Overnight & Intraday returns, unlike trading volume, improves the statistical accuracy of daily volatility forecasts. Fengzhong, Shich, Havlin & Stanley (2009) investigate that the two components of the Total daily return (close to close), the Overnight return (close to open), and the Intraday return (open to close) tend to be auto-correlated. Moreover they found that the cross correlations between the three different returns (total, overnight and intraday) are quite stable over the entire 20 year period for the stocks of NYSE.

Hong and Wang (2000) cite the following: "a, intraday mean return and volatility are U-shaped; b, intraday trading volume is U-shaped; c, open-to-open returns are more volatile than close-to-close returns; d, weekend returns are lower than weekday returns; and e, returns over trading periods are more volatile than returns over non-trading periods."

Stoll and Whaley (1990) report that open-to-open returns are more volatile than are close-to-close returns. They note that this greater volatility implies that the correlation
between the daytime and the following overnight returns exceeds that of the overnight and the following daytime returns, which in turn indicates that open-to-open returns are more likely to reverse than are close-to-close returns.

Pursuing another line of research, *Harris (1989)* reports that "prices systematically rise near the end of the day and most obviously on the last trade. The price rise is shown to be pervasive through time and across firms and strongest when the last transaction is near the close of trading."

Our study seeks to add to this growing body of knowledge regarding non-randomness in returns during the course of the day which do not assume any particular form of distribution for stock returns.

**Sample and Method**

Our study covers the time frame of January 1 2007 to December 31, 2013. We used the (National Stock Exchange) database for daily stock prices. We categorize daily returns into three different groups - overnight returns, intraday returns, and close-to-close returns. The three groups of returns are calculated as follows.

\[
\text{overnight return}_t = \left( \frac{\text{open}_t - \text{close}_{t-1}}{\text{close}_{t-1}} \right) \times 100\% \tag{1}
\]

\[
\text{intraday return}_t = \left( \frac{\text{close}_t - \text{open}_t}{\text{open}_t} \right) \times 100\% \tag{2}
\]

\[
\text{close-to-close return}_t = \left( \frac{\text{close}_t - \text{close}_{t-1}}{\text{close}_{t-1}} \right) \times 100\%. \tag{3}
\]

In stage I, Augmented Dickey–Fuller test (ADF) is used to determine whether there exist any autocorrelation within the time series data of our variables (ON, ID, lagID, lagON & CC). In Stage II, regression analysis is used to determine the strength of relationship between our variables (ON, ID, lagID, lagON & CC). We use the basic model for our regression analysis:

\[ ID = \beta_0 + \beta_1 \times ON + \beta_2 \times lagID + \beta_3 \times lagON + e, \]

Where the betas are various coefficients, *ID* and *ON* are the intraday and overnight returns of day *t*, and *lagID* and *lagON* are the intraday and overnight returns of day *t-1*. In stage III, we compute the correlation between our variables (ON, ID, lagID, lagON & CC).
Results and Discussion

Table I: Results of Augmented Dickey–Fuller test (ADF).

<table>
<thead>
<tr>
<th>Variable</th>
<th>t-Statistic</th>
<th>Critical value@5%</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>-64.539</td>
<td>-2.861</td>
<td>.0001</td>
</tr>
<tr>
<td>ON</td>
<td>-66.913</td>
<td>-2.861</td>
<td>.0001</td>
</tr>
<tr>
<td>lagID</td>
<td>-64.539</td>
<td>-2.861</td>
<td>.0001</td>
</tr>
<tr>
<td>lagON</td>
<td>-66.913</td>
<td>-2.861</td>
<td>.0001</td>
</tr>
<tr>
<td>CC</td>
<td>-63.317</td>
<td>-2.861</td>
<td>.0001</td>
</tr>
</tbody>
</table>

Table II: Results of Regression Analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std.Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ON</td>
<td>0.31139</td>
<td>0.05873</td>
<td>5.30172</td>
<td>0.000</td>
</tr>
<tr>
<td>LAGID</td>
<td>0.99503</td>
<td>0.00920</td>
<td>108.090</td>
<td>0.000</td>
</tr>
<tr>
<td>LAGON</td>
<td>-0.28326</td>
<td>0.05875</td>
<td>-4.82123</td>
<td>0.000</td>
</tr>
<tr>
<td>C</td>
<td>-1.06E-05</td>
<td>7.96E05</td>
<td>-0.1333</td>
<td>0.894</td>
</tr>
</tbody>
</table>

Table III: Results of Correlation Analysis.

<table>
<thead>
<tr>
<th></th>
<th>ID</th>
<th>ON</th>
<th>LAGID</th>
<th>LAGON</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ON</td>
<td>-0.16</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAGID</td>
<td>0.042</td>
<td>0.094</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAGON</td>
<td>-0.016</td>
<td>0.060</td>
<td>-0.17</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>0.969</td>
<td>0.264</td>
<td>0.064</td>
<td>-0.001</td>
<td>1.000</td>
</tr>
</tbody>
</table>

A quick scan of the results of the ADF test as presented in Table I clearly reveal that the ADF t-statistic for all the variables (ON, ID, lagID, lagON & CC) is highly negative when compared to the respective critical value at 5%. This in other words means that the time series data of none of the variables shows the characteristics of random walk and as such the assumption of the data being efficient in weak form stands refuted.

The results of regression analysis as presented in Table II reveal that the hypothesis of the data/variables being unrelated can be questioned since the probability of the three variables (ON, lagID, lagON) tends to be zero which is less than the level of significance (5%).
The correlation results as put forth in Table III reveal that both statistically and economically significant relationships exist between the overnight returns and the subsequent intraday returns (-.16). A similar relationship exists between lagged overnight and lagged intraday returns. An important criterion for any two variables to be random is that they should not be correlated. The implication of the statistic is that the overnight returns can be somewhat predicted on the basis of intraday returns and vice versa which is in contrary to the weak form of Efficient Market Hypothesis.

**Conclusion**

Undoubtedly the Efficient Market Hypothesis (EMH) has found many takers but over the period of time the concept is fading due to non-applicability in the contemporary capital markets. This study is also an attempt in this regard and we try to find out that if the overnight and intraday returns move as per the theory of random walk. All the three econometric techniques used i.e. the ADF test, Correlation analysis and the Regression analysis clearly demonstrate that the variables used (ON, ID, lagID, lagON & CC) are related and do not behave as warranted by the Efficient Market Hypothesis. All in all, the past prices (technical analysis) can be employed by speculators to take advantage of the non-random behavior of these variables with the objective of earning better profits than as justified in an efficient market.

**References**


Website: http://www.nseindia.com/products/content/equities/indices /historical_index_data.html