Stockholder overreaction and mean reversion: Evidence from Tehran Stock Exchange

Mohammad Khodaei Valahzaghard\textsuperscript{a} and Amin Shakourloo\textsuperscript{b}

\textsuperscript{a} Assist. Prof. & Faculty Member, Department of Accounting, School of Management and Human Sciences, Tehran North Branch, Islamic Azad University (IAU), Tehran, Iran
\textsuperscript{b} M.Sc. Student, Department of management, School of Management and Human Sciences, Tehran north branch, Islamic Azad University (IAU), Iran

\textbf{ABSTRACT}

In this paper, we address the stockholder overreaction and mean reversion in specified major industry groups in Tehran Stock Exchange (TSE). This paper investigates this issue with panel data analysis and with particular attention to the Box-Jenkins Approach for stationary diagnosis with appropriate order and modeling stock prices with regard to specific industries. The study processes modeling of panels where stationary and mean reversion takes place in complementary analysis. The sampling intervals are explored monthly within the past few years. The results reveal that mean reversion presence in three industry group stock prices and industry stock prices would not behave in certain pattern.

1. Introduction

The efficient market hypothesis (EMH), initiated at the University of Chicago's business school, states that stock prices fully reflect all available information, and no market participant is able to systematically make abnormal profit (Fama, 1970). In fact, the EMH is integrated by coupling the evolutionary principle with the notion of bounded rationality (Simon, 1955). A bounded rational investor is normally considered to present satisfying rather than optimal behavior. When the information set is limited to only historical prices, the market is considered to be in weak form efficient, and asset return is completely unpredictable from historical prices. While most finance academics state that the market is called weak-form efficient (Doran et al., 2010), there are several critics from behavioral finance who document irrational but anticipated investor behavior such as overreaction and overconfidence (De Bondt & Thaler, 1985; Barber & Odean, 2001) and the momentum impact (Jegadeesh & Titman, 1993). The theory that financial asset markets are efficient has been under tremendous investigation among contemporary finance researchers. While the first
version of this theory states that financial information is disseminated effectively and stock prices cannot be predicted, many question the universality of this theory. More specifically, researchers have provided evidence of negative autocorrelation, or “mean reversion,” in stock returns over long time horizon. The presence of mean reverting process in stock index returns has been investigated, significantly and it was first analyzed for the U.S. market. DeBondt and Thaler (1985) reported that past losing stocks over the previous 3-5 years substantially outperform past winning stocks over a 3-5 years of holding period. Their results indicated that stock prices would not follow a random walk, but contained a strong mean reverting component. The autocorrelation in market returns may be either positive or negative. Negative very short term, positive short-term autocorrelations as well as long-term reversal of stock returns have been widely investigated (Koutmos, 1999). The presence of positive autocorrelation implies partial adjustment of stock prices towards their intrinsic values. The negative autocorrelation, on the other hand, indicates that stock price changes are followed by predictable volatility in the opposite direction, which is consistent with the view that market agents overreact irrationally to price shocks from their long run values (De Bondt & Thaler, 1985, 1987).

Stock return autocorrelation and its positive/negative effects have been an interesting issue among investors for utilizing appropriate trading strategies. For example, some evidence for significant autocorrelation represents predictable return behavior and the present autocorrelation is negative, a contrarian trading strategy can be profitable. On the other hand, in spite of its popularity in practice, industry analysis has received limited academic attention in finance research specifically in emerging and developing markets such as Tehran Stock Exchange (TSE). Microeconomics state that the market supply can be determined by a group of firms that produce homogenous products (i.e., industry). However, in financial economics, the supply (or demand) of an asset is infinitely elastic, because all assets are perfect substitutes. Popular models, whether rational or behavioral, simply grant no effect to industries. Nevertheless, researchers commonly control for the “industry effect” in empirical studies, without any theoretical basis for doing its (e.g., Kahle & Walking, 1996). Theoretically, an industry is associated with a group of firms producing homogenous products or close substitutes; specifically, a firm rarely produces just one product. Broad industry classifications, such as standard industrial classification (SIC) codes have been implemented to identify homogeneous groups of firms, which engage in practice in close businesses. These classification schemes merely reflect broad attributes, such that business units of the same industry could be competitive when they generate close substitutes but cooperative when their products are complements. Whether it is market wide, industry specific, or firm specific in nature, it is difficult to predict how firms could react to relevant information. The amount of research devoted for analyzing mean reversion using industry sorted portfolios has been limited. Only Fama and French (1988) applied industry groups to test for the existence of mean reversion.

In summary, we identify several unique industry related patterns in our analysis in TSE market, which appear to be new in the literature. The empirical findings indicate that industry stock prices in TSE reflect both significant rational and behavioral components, differently, but neither rational nor behavioral theories alone can fully describe stock price fluctuation. The results indicate that industries play a specific role in explaining stock prices, which deserves further exploration with various methods. In other words, the testable implications of the EMH and then stockholder overreaction are twofold. First, the degree of market efficiency fluctuates over time. Second, the degree of market efficiency is governed by market conditions. This paper examines the first implication by tracking the evolution of price over 48 lag and predictability of the TSE stock market in this period.

In this paper, we first examine for random walk phenomenon in stock prices of TSE known as developing market; using monthly data over the period 2009-2013. We provide useful information from this independent sample, and locate in the existing studies on developing markets. Our research operation occurs in first ten major industries in TSE that have maximum number of stock holder compared with other industries in TSE, so we can measure this phenomenon specially and exactly in
terms of major number of real stockholders in TSE. In particular, we investigate the relationship between the overreaction parameter and market industry group using Box-Jenkins Approach for Time Series Models. The results of our research operation suggest that inverted AR root exist in price time series in seven numbers of ten industries and rest of industries has a less inverted AR root amounts, thorough this method for unit root tests the results implying that mean reversion take place in three industry groups.

2. Data and methodology

The proposed study of this paper uses the Box-Jenkins methodology based on ARIMA models for industry stock price panels modeling and means reversion presence determination thorough stationary diagnosis. Auto Regressive (AR) models were first introduced by Yule (1926). Slutsky (1937) presented Moving Average (MA) schemes by supplemented them, consequently. Wold (1938) who combined both AR and MA schemes showed that ARMA processes could be applied to model all stationary time series as long as the appropriate order of p, the number of AR terms, and q, the number of MA terms, was appropriately specified. This means that any time series can be modeled as a combination of past values and/or past errors. Thorough Box-Jenkins methodology, this study investigated whether or not price shocks in industry group had temporary impact on stock price and lose their effect over subsequent lags. If price shocks have temporary effect, mean reversion takes place in price group and stockholder overreacted in relative industry group. The primary objective of this paper is to examine stockholder overreaction in specified ten major industry group thorough mean reversion method, so we focus panels data stationary diagnosis and price time series modeling operations on series that mean reversion take place in these. For this purpose, we use routine Box-Jenkins approach that shown in Fig. 1 as follows,

![Fig. 1. Schematic representation of the Box-Jenkins Approach](image)

Finally, we use a model to forecast criteria estimation for accuracy in stock price panels, where mean reversion takes place in their process and predictable, and makes autocorrelation test on this model. Before these operations, we brought descriptive analysis relative to each industry group prices in time of our research. We apply the setup to TSE financials stock market data over the period 2009-2013 and explain how the price time series for each industry groups can be modeled. We employ monthly industry price stock in our analysis. For these panels the Box-Jenkins procedure will be used to analyze the series and select an appropriate model that can be used for forecasting. There are three
types of theoretical Box-Jenkins models. Eq. (1) shows moving average model of order q, moving average (MA) is the simplest of time series models. Eq. (2) shows autoregressive model (AR) and we use in our analysis for price panels modeling. Finally we bring mixed autoregressive moving average (ARMA) model in Eq. (3).

\[ Y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q}, \] (1)

\[ y_t = \mu + \sum_{i=1}^{p} \varphi_i y_{t-i} + u_t, \] (2)

\[ Y_t = \mu + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q}. \] (3)

We adopted these regression models because these models have been adapted specifically for time series variables that main comprise of this research subject. In addition, time-series regression models are useful when the parameters describing the trend, seasonal, or cyclic components of a time series are deterministic. To select optimum ARMA model with appropriate order we use information criteria as known Schwarz (1978) Bayesian Information Criterion (SBIC) & Hannan-Quinn Criterion (HQIC) in Eq. (4) and Eq. (5) as follows,

\[ SBIC = \ln(\hat{\sigma}^2) + \frac{k}{T} \ln(T), \] (4)

\[ HQIC = \ln(\hat{\sigma}^2) + \frac{2k}{T} \ln(\ln(T)), \] (5)

where \( \hat{\sigma}^2 \) represents residual estimation variance. After developing the proposed model for the trend, we have to confirm that the chosen model is the most accurate by examining the measures of accuracy for fitted models. The mean absolute percentage error (MAPE), mean absolute error (MAE), and Theil Inequality Coefficient (TIC), are three measures of accuracy. Whatever these criteria are smaller they are better indicate the fit of the model. All three indicators measure the accuracy of the fitted time series values, but they express the results in different units. MAPE measure the accuracy of fitted time series values and expresses accuracy as a percentage. We show these predict criteria in Eq. (6) to Eq. (8). The mean absolute error can be defined as:

\[ MAE = \frac{\sum_{t=T+1}^{T+m} |Y_t^f - Y_t|}{m}, \] (6)

where \( m \) is total sample size that’s from \( T + 1 \) to \( T + m \) and \( Y_t^f \) is predicted values and \( Y_t \) is real values. MAE accuracy criterion provides a quadratic loss function, and so may be particularly useful in situations where large forecast errors are disproportionately more serious than smaller errors. This may, however, also be viewed as a disadvantage if large errors are not disproportionately more serious. In addition, it has the attractive additional property compared with MSE interpreted as a percentage error. Another criterion that we apply in this research is MAPE shown in following equation:

\[ \text{MAPE} = \frac{100}{m} \sum_{t=T+1}^{T+m} \frac{|Y_t^f - Y_t|}{Y_t} \] (7)

Adjusted MAPE or symmetric MAPE corrects for the problem of asymmetry between the actual and forecast values. Theil Inequality Coefficient (TIC) is another accuracy criterion that applies in this paper. The amount of this criterion always varies between one and zero. The zero amounts indicate lack of forecasting error in estimation. Eq. (8) shows TIC accuracy criteria structure:
\[ TIC = \frac{\sqrt{\sum_{t=T+1}^{T+m} (Y_t^f - Y_t)^2}}{m} \]

\[ \sqrt{\frac{\sum_{t=T+1}^{T+m} (Y_t^f)^2}{m} \sqrt{\frac{\sum_{t=T+1}^{T+m} Y_t^2}{m}}} \]

(8)

2.1. Descriptive statistics

Table 1 provides descriptive statistics of the price panels in each ten industry. The Jarque-Bera test statistics for normality soundly indicate deviations from the normal distribution in three industry groups under our study; these groups include Food products, Construction, and Computer. As reflected by the standard deviation, the Petroleum products, Pharmaceutical & Computer industry groups have the highest volatility while the Banking, Construction & Construction materials industry groups record the least volatility. These can be clearly seen in price plots that presented in Fig. 2. All industry group prices except the Chemicals, Petroleum products and Banking groups are positively skewed. The kurtosis statistics, which are substantially higher than 3, indicate excess peakness of the industry stock price distribution in Food products, Construction, Computer & Construction materials group.

![Price plots for various industries](image)

**Fig. 2.** Three highest and the lowest price volatility industry groups

<table>
<thead>
<tr>
<th>Industry</th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Std. Dev.</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>J-B Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machinery</td>
<td>1584.791</td>
<td>2430.500</td>
<td>1092.857</td>
<td>358.7740</td>
<td>2.618102</td>
<td>0.798129</td>
<td>3.142870</td>
<td>0.207747</td>
</tr>
<tr>
<td>Food products</td>
<td>2355.714</td>
<td>3311.714</td>
<td>1826.429</td>
<td>328.0418</td>
<td>4.000462</td>
<td>0.934676</td>
<td>7.492332</td>
<td>0.023608</td>
</tr>
<tr>
<td>Chemicals</td>
<td>4935.932</td>
<td>6326.182</td>
<td>3317.455</td>
<td>949.8026</td>
<td>1.810032</td>
<td>-0.289097</td>
<td>2.917219</td>
<td>0.232559</td>
</tr>
<tr>
<td>Computer</td>
<td>1792.842</td>
<td>2228.375</td>
<td>1578.375</td>
<td>160.2796</td>
<td>4.116158</td>
<td>1.249563</td>
<td>12.48573</td>
<td>0.001944</td>
</tr>
<tr>
<td>Petroleum products</td>
<td>11436.28</td>
<td>31500.00</td>
<td>7040.000</td>
<td>5574.888</td>
<td>6.357150</td>
<td>1.945955</td>
<td>44.02905</td>
<td>0.000000</td>
</tr>
<tr>
<td>Pharmaceutical</td>
<td>9765.625</td>
<td>13321.00</td>
<td>5358.500</td>
<td>2466.024</td>
<td>1.898130</td>
<td>-0.484054</td>
<td>3.585588</td>
<td>0.166494</td>
</tr>
<tr>
<td>Fundamental metals</td>
<td>3727.507</td>
<td>5133.429</td>
<td>2365.286</td>
<td>737.4794</td>
<td>2.273090</td>
<td>0.271281</td>
<td>1.371284</td>
<td>0.503767</td>
</tr>
<tr>
<td>Petroleum products</td>
<td>3053.650</td>
<td>3804.000</td>
<td>2534.375</td>
<td>274.3516</td>
<td>0.397209</td>
<td>0.379198</td>
<td>0.974357</td>
<td>0.614357</td>
</tr>
</tbody>
</table>

**Table 1**

Industry stock price panels descriptive statistics
2.2. Price model identification

As we can see from the results of Table 2, price panels in the seven industry groups have a unit root. This shows us that these groups have random walk in their process and exhibiting non-stationary behavior obtained by assuming optimal difference order in these time series. In other word AR(1) coefficient assumes to follow a random walk in Foods, Chemicals, Pharmaceutical, Construction, Petroleum products, Banking & Fundament metals industry groups.

Also the results of this study in Table 2 show that mean reversion takes place in Machinery, Computer and Construction material industry groups and price panels in these groups behave stationary, so stockholder overreaction occurred in these groups with various intensity according to Box-Jenkins time series modeling with appropriate order. There is a strong tendency that the null hypothesis of no price predictability is rejected for these three industries group. In other groups inverted AR root absolute value greater-than or equals to sign one, so price time series process in these industry groups behave non-stationary in Box-Jenkins price time series modeling method and mean reversion do not take place in these groups. Therefore, price time-series modeling and sequence forecasting operation are impossible for these groups.

Next, we determine whether there is any trend and/or seasonal component. Therefore, we attempt to produce a model based solely on previous stock prices, which are correlated but can be heavily impacted by endless possible extraneous factors. From a statistical stand point broad aggregates such as national consumption, income and savings are not considered and make it hard to produce a model based only on previous stock prices. However, to determine whether there is any underlying trend or seasonality, an autocorrelation function (ACF) of the differenced series is analyzed. The autocorrelations for stationary series are large for low order autocorrelations but dies out rapidly as lag length increases. If the series is trended, autocorrelations are large with positive for short lags, decreasing slowly as the lag increases. We show these concepts for price time series in each industry groups in three figures and we could look at which of the ACF or PACF is cutting off more abruptly to tentatively identify the model and look at the estimations, diagnostics, and forecasts to select the best model.

After analyzing time series for optimum model with appropriate order, we show in Table 2, the AR(1) model adopted for Computer and Construction material groups and only Machinery group have been ARMA(1,1) structure in our analysis. So the resulting models for nearly all the series were AR(1) models. For stationary stock price panels, the models were all AR(1) and they behaved similar to a random walk because the AR parameter was very close to one.

Table 2
Stock price panels model identification results

<table>
<thead>
<tr>
<th>Industry</th>
<th>Box and Jenkins model</th>
<th>coefficient</th>
<th>t-statistics</th>
<th>prob</th>
<th>R^2</th>
<th>HQIC</th>
<th>D-W stat</th>
<th>Inverted AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machinery</td>
<td>AR(1)</td>
<td>0.989707</td>
<td>56.43842</td>
<td>0</td>
<td>0.863463</td>
<td>17.98645</td>
<td>2.008531</td>
<td>0.99</td>
</tr>
<tr>
<td>Foods</td>
<td>AR(1)</td>
<td>0.997608</td>
<td>76.58354</td>
<td>0</td>
<td>0.623216</td>
<td>17.40238</td>
<td>1.911512</td>
<td>1</td>
</tr>
<tr>
<td>Chemicals</td>
<td>AR(1)</td>
<td>1.015694</td>
<td>97.42368</td>
<td>0</td>
<td>0.879422</td>
<td>19.41877</td>
<td>2.066306</td>
<td>1.02</td>
</tr>
<tr>
<td>Pharmaceutical</td>
<td>AR(1)</td>
<td>0.999849</td>
<td>108.326</td>
<td>0</td>
<td>0.856679</td>
<td>20.58532</td>
<td>1.707378</td>
<td>1</td>
</tr>
<tr>
<td>Computer</td>
<td>AR(1)</td>
<td>0.957423</td>
<td>22.03551</td>
<td>0</td>
<td>0.581994</td>
<td>20.60559</td>
<td>2.104433</td>
<td>0.96</td>
</tr>
<tr>
<td>Construction</td>
<td>AR(1)</td>
<td>0.999968</td>
<td>99.65381</td>
<td>0</td>
<td>0.453582</td>
<td>16.59872</td>
<td>1.612573</td>
<td>1</td>
</tr>
<tr>
<td>Oil</td>
<td>AR(1)</td>
<td>0.995031</td>
<td>79.15986</td>
<td>0</td>
<td>0.709556</td>
<td>18.64248</td>
<td>1.88122</td>
<td>1.01</td>
</tr>
<tr>
<td>Fundament metal</td>
<td>AR(1)</td>
<td>1.006611</td>
<td>66.2983</td>
<td>0</td>
<td>0.709956</td>
<td>18.64248</td>
<td>1.88122</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Schwarz-Bayesian Information Criterion (SBIC) and Hannan-Quinn Criterion (HQIC) indicated in Table 2 represent the minimum amount in compare with another Box-Jenkins time series model, this
fact confident us that we adopt optimum model with appropriate order for price modeling in our analysis. In this paper, the order of the ARMA model is found by examining the autocorrelations and partial autocorrelations of the stationary panels. For this purpose, once the trend has been modeled appropriately, we compute the autocorrelation and partial autocorrelation coefficients for the residuals of the trend model.

In Machinery price panel group, the autocorrelations are highly correlated for lower lags and then decreases as the lags increase, but for this panel the ACF continues to increase negatively after 10 lag, also the increase in autocorrelation starts to decrease positively after 19 lag. After this analysis and also with respect to Hannan-Quinn criterion we can confirm the model for this series is AR(1).

ARMA structure for construction materials industry group plots do not show any intense apparent fluctuation, and do not indicate any drastic rise or fall in the price of the stock except after 2 lag occurs. After this analysis and also with respect to Hannan-Quinn criterion, we can confirm the model for this panel as AR(1).

Fig. 3. Machinery group ARMA structure

Fig. 4. Construction materials group ARMA structure

Fig. 5 shows the case of computer industry group stock price panel, overall, the results are similar to the case of Machinery group and price panel behave similarly in Computer industry group to machinery group but with different lag. As we can see in Fig. 5 the autocorrelations in Computer industry group are highly correlated for lower lags and then decreases as the lags increase, but for this series the ACF continues to increase negatively after 8 lags. It is indicating the presence of overall
negative stock price autocorrelation, which represents a higher degree of panic or overreaction by market participants than in past crises. In fact, in this time series, differencing caused the ACF to die off faster and price shocks have temporary effects. In addition, this increase in autocorrelation starts to decrease positively after 18 lags. After this analysis and also with respect to Hannan-Quinn criterion we can assume the appropriate model for this panel is AR(1).

![Fig. 5. Computer group ARMA structure](image)

2.3. Forecast criteria investigation

Table 3 with various accuracy criteria shows a high degree of accurate price predict based on ARMA panel models over the period 2009-2013 price panels for three industry groups with stationary panel process with statistical significance. For more confidence, we use another accuracy criteria including Bias Proportion, Covariance Proportion & Variance Proportion. Whatever BP & VP proportions tends to zero, it means that predict operation acts with more accuracy.

<table>
<thead>
<tr>
<th>Forecasting criteria</th>
<th>Machinery</th>
<th>Industry Computer</th>
<th>Construction material</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>1353.787</td>
<td>3521.821</td>
<td>1549.500</td>
</tr>
<tr>
<td>MAPE</td>
<td>6.185208</td>
<td>14.08803</td>
<td>6.273608</td>
</tr>
<tr>
<td>TIC</td>
<td>0.042500</td>
<td>0.145637</td>
<td>0.041348</td>
</tr>
<tr>
<td>BP</td>
<td>0.002216</td>
<td>0.016560</td>
<td>0.001800</td>
</tr>
<tr>
<td>VP</td>
<td>0.000044</td>
<td>0.004672</td>
<td>0.000096</td>
</tr>
<tr>
<td>CP</td>
<td>0.997739</td>
<td>0.978768</td>
<td>0.998103</td>
</tr>
</tbody>
</table>

The MAPE for the forecasts in three industry groups are 6.185208, 14.08803, and 6.273608, respectively, which indicates they follow the same overall behavior, approximately. As we can observe from the results of Table 3, the forecasts are very similar to the actual prices. In other words, none of the forecasts is far off from the actual values and forecast analysis indicates that these model adequate.

It is clear from Table 3 that, in general, the degree of price predictability is higher in Construction material group because of BP & VP accuracy criteria amounts and they tend to zero compared with amounts of these two criteria in other two industry groups. We show industrial stock price panel reflects the behavior of the stock with mean reversion without any known imposed manipulations in Fig. 6 for Construction material industry group. This series consist mainly of fluctuation with significant mean reversion that occurs in the price during lags.
In Fig. 7, we show industrial stock price panel reflects the behavior of the stock with mean reversion without any known imposed manipulations for computer industry group. This series consist mainly fluctuation with significant mean reversion that occurs in the price.

As we can see in Fig. 7, computer industry group maintains two upward trends in stock price in 2010. In addition, we notice that the industry stock price panel continues cutting off with much faster rate after previous increase and adjust to long run values in subsequent period. This decrease did not influence the data sampled on the subsequent periods of the research as much and the overall behavior is very similar. For third industry groups with stationary price panel, we see stock price panel that reflects the behavior of the stock with mean reversion without any known imposed manipulations in Fig. 8 for Machinery industry group. This panel consists mainly of fluctuation with significant mean reversion that occurs in the price during lags. Although the actual prices increased overall in 2010 last part but they experienced a major drop in 2011 and stock price panel back to the intrinsic value after subsequent lags.

3. Conclusion

This paper has investigated industry stock price panel mean reversion analysis by using the Box-Jenkins approach with regard to specific industries. We have estimated an overreaction parameter
governing the dynamics of industry stock prices panels in ten separate industry groups. In addition, we examined random walk process in stock price panel for these groups. Our tests have failed to detect a random walk behavior in three industry price panels and stock prices did not follow a random walk in three industry groups, but contained a strong mean reverting component. Panel tests for the existence of mean reversion in asset prices have provided definitive evidence for the presence of mean reversion and stockholder overreaction in these three industries for short-term 48 periods in this case. We have shown that a temporary shock to an asset’s prices takes no constant in three industry-group thorough Box-Jenkins price time series modeling in fact thorough this method significance amount during the periods and price process tend to revert towards their intrinsic values. The fact that stock prices experienced mean reversion was robust and so we have rejected capital market efficiency for TSE market in our study for three major specified industry groups. The study does not suggest that the Random Walk investment strategy could be useful for this market in short-term periods.

References