

## **Predicting the Stock Return Direction Using Artificial Neural Network: The Case of Amman Stock Exchange**

*Basheer Ahmad Khamees*

### **ABSTRACT**

This study aims at measuring the ability of artificial neural networks (ANNs) to predict the direction of one-year-ahead stock return. It applies fundamental analysis on the Jordanian industrial and service corporations over the period 1999 to 2009. Five independent variables are used to achieve the objective of this study, which are earnings change, earnings level, the ratio of market value of equity to book value of equity, dividend yields, and return on equity. Descriptive measures, correlation analysis, and regression analysis are also used to explore data before (ANNs) are used to predict stock return. The above variables were modelled, as lag variables, with stock return. The findings of the study show that (ANNs) are able to predict stock return with accuracy of 80.2% in the training sample and 58.1% in the test sample.

**Keywords:** Stock Return, Artificial Neural Networks, Regression Analysis, Fundamental Analysis.

### **INTRODUCTION**

Efficient market hypotheses assume that the price of a security fully reflects all available information at that point in time. Thus stock returns are non-predictable (Malliaropulos, 1996). It was until mid-eighties when the non-predictability result of stock return has been overturned. For example, Keim and Stambaugh, 1986 and Fama and French, 1988 have noticed that some financial ratios such as dividend yield and price-earnings explain part of the variation in stock returns measured over intervals of several years. Many other research papers also indicate that accounting earnings and some of its components capture information that is contained in stock prices. It is assumed that these variables can discover values that are not reflected in stock prices. In other words, they can discover overpriced and underpriced stock.

In addition, traditional fundamental analysis assumes that firms' values are indicated by information in financial statements. It also assumes that stock prices

depart at times from these values and only slowly move towards the fundamental values. Therefore, Ou and Penman, 1989 argue that analysis of published financial statements can discover values that are not reflected in stock prices. Patelis, 1997 also concluded from reviewing previous literature that expected stock returns are time-varying. In other words, it means that assets returns are, to some extent, predictable.

Malliaropulos, 1996 summarizes several explanations of the possibility of stock returns predictability or mean reversion of stock prices. It can be explained for example by speculative bubbles, asset pricing model with time varying required returns, models of noise traders, or feedback traders.

Stock returns were predicted using several types of analyses, such as multivariate adaptive regressive splines (Safer, 2003), logit analysis (Ou and Penman, 1989) and (Holthausen and Larcker, 1992), and recursive least squares (Clare et. al., 1997). A more powerful tool that can be used to improve stock prediction is artificial neural networks, which is a nonparametric technique useful for analysing nonlinear data sets such as those that characterize stock price information. Several studies have used neural networks to analyse stock market data

---

Received on 15/1/2012 and Accepted for Publication on 16/4/2012.

[For example: (Safer, 2003), (Faria et. al., 2009), and (AVCI, 2009)]. This study will expand previous studies in order to examine the predictive ability of neural networks in emerging markets.

### 1.1 Study Problem

The problem of the study is that it is unclear whether stock return of the Jordanian industrial and service corporations can be predicted ex-anti utilizing financial variables. Although there are several studies that have been conducted to develop models to predict the failure of Jordanian corporations, there is no model that predicts stock return. Examples of these studies are (Gharaibeh and Abdelatif, 1987); (Khamees and Gharaibeh, 1990); (Juhmani and Al-Daoud, 2004) and (Alawi and Gharaibeh, 2008). In addition, the studies that relate stock return to financial ratios of the Jordanian corporations did not examine the predictive ex-anti ability of these variables (i.e. they did not use financial ratios of a specific year to predict the stock return of the following year).

Therefore, the purpose of this paper is to examine the ability of artificial neural networks to predict the direction of one-year-ahead stock return. This may provide an indirect indicator of how ASE is efficient. In addition, it may help investors in taking their investment decisions. It should be emphasized that examining all of the possible ratios is considered as a limitation of the study. Thus, only the common ratios calculated from the published financial statements are used.

The paper consists of six sections including an introduction and a summary. The second section reveals the objective and importance of the study. The third section provides a literature review of stock return prediction. The methodology and data of the study is presented in section four. Section five reports the results. Finally, section six summarises and concludes the paper.

## 2. Significance of the Study

Predicting stock returns has some value for both the institutional and individual investors. For the

institutional investors predicting stock returns can provide some clues for their portfolio decision. It is also of great importance for individual investors, who are not capable of diversification (Faria et. al., 2009). Therefore, implementing new and powerful tools to predict stock returns may enhance the ability of different types of investors in taking their investment decisions. One should keep in mind that using these tools by all investors, the matter which is highly unlikely due to lack in information technology resources and knowledge among investors, may reduce their ability to achieve any abnormal return.

Unfortunately, the relationships between financial ratios and stock returns are not so clear to the investor (Ou and Penman, 1989). That makes it difficult to predict stock returns in advance. The object of this study is to examine to which extent ANNs are capable to predict stock return in an emerging market. The prediction ability will be examined in terms of a percentage of correct predicts.

## 3. Literature Review

In this part some of the previous studies that have been conducted on the prediction of stock returns are reviewed.

Ou and Penman (1989) is one of the studies that initiated rigorous academic research on using financial ratios to predict future earnings and intern predicting stock returns. They utilized logit analysis to predict the direction of stock return indirectly through annual earnings per share during the period 1973 - 1983. They used traditional financial statements ratios as independent variables to predict unexpected earnings. Then these earnings were intern used to predict stock returns. Depending on the prediction of unexpected earnings, which are likely to be positive or negative, the trading strategy they follow takes a long or short position in the common stocks of firms. Depending on their strategy, the average market adjusted return is 8.3% for a 12-month holding period. It appears from their results that financial statement analysis captures equity

values that are not reflected in stock prices.

Holthausen and Larcker (1992) implemented Ou and Penman trading strategy during the period 1978 – 1988. They produced as a result of their implementation, on average, annual returns of between -0.1% and 1.6%. In other words, they did not achieve excess return. Therefore, they developed a logit model that utilizes financial statement information to predict the sign of subsequent twelve-month excess stock returns directly. They based their model on three different measures of 12-month excess returns. These measures are: market-adjusted returns, excess returns computed using the Capital Asset Pricing Model (CAPM), and size-adjusted returns. Depending on their model, the achieved average annual excess returns were between 4.3% and 9.5%.

Morton and Shane (1998) argue that if the success of financial statement analysis achieved by Holthausen and Larcker (1992) is due to inefficiency, then the prediction of small firms stock returns should be better than larger firms. They based their argument on the amount of information processing in the firms information environment. They assume that small firm environment provides less information than large firms. Therefore, they examined the effects of the information environment on abnormal returns. Their results do not show greater predictable power associated with the analysis of small firms financial statements. Morton and Shane clarify that excess return may be due to omitted risk in the measurement of its risk.

Cao et. Al. (2005) used artificial neural networks to predict stock price returns for firms traded on the Shanghai stock exchange. They also used three independent variables, which are: the market risk factor, the firm's market capitalization, and the book-to-market value of the firm. Cao et. al. Found that neural networks outperform the linear models compared. They concluded that neural networks are a useful tool for stock price prediction in emerging markets.

Several local studies have been conducted to examine the explanatory power of several variables to stock return. Most of these variables used regression analysis to connect these variables with stock return.

One of these recent studies is (Abdelqader et. al. 2008), in which the authors examined the effect of expanding the measurement window of the relationship between accounting earnings and stock returns on the explanatory power of the model and the earnings response coefficient. They found that expanding the measurement window enhances the explanatory power of accounting earnings in regard to variations in stock market prices. They also found that the value of the earnings response coefficient increases as the covariance between accounting earnings and unrecorded goodwill increases. They estimated goodwill by subtracting book share value from market share price.

Saleh, 2007 applied the one-factor and the multi-factor models to examine the ability of several variables to explain the cross-sectional stock returns over the period 1980 – 2000. Saleh found that sorting Jordanian firms year asset growth rates is not superior over earnings-to-price or dividend-to-price. Saleh concluded that dividends-to-price ratio appears to be the best strategy that helps in predicting the cross-section of Jordanian stock returns.

Al-Shiab and Al-Alawneh, 2007 used the security market line (SML) to evaluate the common stock performance of the Jordanian corporations. They also adapted the Augmented Dickey-Fuller Unit Root for testing the stationarity assumption. They found that companies listed on Amman Stock Exchange (ASE), on average, are over-priced. They concluded that making abnormal profit is fairly possible and consequently ASE efficiency is questionable (Share prices for companies listed on ASE were not fully fairly priced).

Al-Shiab and Al-Ali, 2006 examined the relationship between investment performance of equity securities and their P/E ratios using risk-adjusted measures. They found that P/E ratios, due to exaggerated investor expectations, may be indicators of future investment performance. Their Ordinary Least Squares (OLS) regression results suggest that gaining abnormal return is a function of changes in Portfolio's P/E over the period considered.

Al-Rjoub et. al., 2006 used GARCH in mean and

Exponential GARCH models to examine the relationship between stock returns and risks. They found that the theoretical relationship between returns and risks exists even after adjusting for sudden change in variance of market return.

Also Haddad and Abu Al-Ulla, 2004 Found a significant relationship between economic value added, and refined economic value added, and abnormal returns.

Al-Khadash and Al-Abbadi, 2005 examined the relationship between the accounting rate of return to equity, the cash flow to equity and the influence on stock market prices of the industrial Jordanian corporations. They found that both tested models have significant relationships with stock prices. However, the cash flow to equity is more related to stock prices relative to the accounting rate of return to equity.

Nour and Al-Fadil, 2003 examined the effect of firm size, profitability, and degree of non systematic risk on the relationship between dividends and stock market returns. They applied their study on industrial, agricultural, and service corporations listed in both Amman and Baghdad stock markets. Their results show that firm profitability does not affect the relationship between dividends and market returns. However, the results show that this relationship is affected by firm size and degree of non systematic risk.

In order to search for risk proxies, Haddad, 2003 found no relationship between four variables and stock returns. These variables are: book to market, debt to equity, sales to the company's market value, and firm size. Therefore, he concluded that it is not possible to use these variables as proxies for risk.

Abu Nassar and Al-Debi'e, 2002 applied three regression models to examine the relationship between the annual stock return and per share data for earnings, sales, cost of sales, administrative and selling expenses, interest expenses, and income taxes. Their first, second, and third model include earnings per share, the other data elements, and all per share data, respectively. They found that stock returns are affected by earnings per share. They also found that the other data elements have

no incremental effect on stock returns above that of the earnings figure. Instead, their results show that earnings per share have additional information content above the other data elements.

Al-Khalayleh, 2001 applied ordinary least squares (OLS) to examine the relationship between market-based performance measures and accounting performance measures in the long run. He found a significant relationship between stock return and both of the return on assets and the return on equity.

Omet and Bino, 2000 examined the relationship between the risk and stock returns during the period 1987-1997. They utilized simple and multiple regression analyses to investigate whether or not risk measures including systematic risk are significant in explaining the changes in stock returns. They found that CAPM is not applicable to the Jordanian Market and furthermore, other risk measures do not affect portfolios' returns.

The current study is distinguished from the previous literature in being the first study, according to the researcher knowledge, that predicts one-year ahead the direction of Jordanian corporation stock returns. However, this study builds upon and completes the previous efforts, which justifies reviewing these previous studies.

#### **4. The Methodology and Data**

This study examines the ability of neural networks to predict stock return. It is an empirical study. It reveals the prediction performance of neural networks and that of the regression analysis. It also examines the colorations between the independent and dependent variables.

##### **4.1 The Prediction Models**

Prediction models of the study are going to be developed using artificial neural networks. However, regression analysis will be used to explore data before applying networks. A brief discussion of each technique is provided in this section.

#### 4.1.1. Artificial Neural Networks

Artificial neural networks (ANNs) are parallel information processing techniques that simulate living nervous systems. They are computer-based systems that work differently than conventional computing which processes its input one thing at a time and works sequentially. It is much harder for conventional programmes or computing to detect patterns that become obvious when one can see the whole picture at once (Nelson and Illingworth, 1991).

An ANN can be defined as a collection of simple, analog processors (nodes), connected through links called connections. It is a directed graph schema, where the nodes represent the processing elements, the arcs represent the modulating connections, and arrowheads on the arcs indicate the normal direction of signal flow. The nodes are arranged into layers, which are an input layer, one or more hidden layers, and an output layer. They are used to estimate the parameters (weights) of the data. Modifying these weights enable the network to learn the application data patterns. The number of nodes within each layer and the number of hidden layers can vary depending on the size and nature of the data set (Skapura, 1995).

For more clarification, the number of neurons in the input layer equals the number of the input variables. But there is no satisfactory method to define the number of neurons that should be used in hidden layers. The size of hidden layers is usually found by a trial and error method. It should be emphasized that if more neurons are used, more complicated shapes can be mapped and networks increasingly lose their generalization ability (Safer, 2003). Therefore, when designing a neural network, a researcher should determine the number of layers, the number of nodes in a given layer, number and type of interconnections, data representation (distributed or local representation), type of paradigm, type of learning model (supervised or unsupervised), learning rates, and training techniques (Nelson and Illingworth, 1991).

ANNs have strong mathematical basis, inherent parallelism, no specific memory locations, fault

tolerance, the ability to self-adjust, and pattern-recognition skills. They have powerful ability to detect complex, non-linear relationships among a number of different variables. On the other hand, they cannot justify answers (not human plausible), not good if precise answers are required (Nelson and Illingworth, 1991). They are able to model nonlinear processes without priori assumption about the nature of the generating process (Hagen et al. 1996).

In general, the typical components of neural networks are nodes, inputs and outputs, weighting factors, neuron functions, activation functions, transfer functions, learning functions, combining elements, combining layers, connectivity options, and filters (Nelson and Illingworth, 1991). The pattern of connections between these components and the structure of them characterize neural networks. The various possible patterns and structures result in numerous types of neural network architectures that have been developed in the literature. One type of neural networks often used for classification problems is feed-forward neural network which is adopted in this study. In this type of neural networks, information moves in one direction. It starts from the input nodes passing through the hidden nodes down to the output nodes. It should be also emphasized that connections between nodes do not form a directed cycle. Other less common networks are recurrent networks, probabilistic networks, and fuzzy networks. The most common algorithm that has been used to perform training on the feed-forward neural network is the backpropagation.

Backpropagation, as a training algorithm, trains a neural network using a gradient descent algorithm to minimize the mean square error between the network's output and the desired output. This creates a global cost function which minimized iteratively by backpropagating the error from the output nodes to the input nodes. When the network's error has reached the specified threshold, the network has converged and is now trained. However, the convergence rate of backpropagation as an algorithm is a very poor rate (Safer, 2003).

In this study, the network parameters are determined through the following procedures. Firstly, the values of the input variables were pre-processed by normalizing them within a range of -1 and +1 to minimize the effect of magnitude among the inputs, and to increase the effectiveness of the learning algorithm. Secondly, a sigmoid hyperbolic tangent function was selected as the activation function to generate an even distribution over the input values. Thirdly, the resilient backpropagation training algorithm was used to train the neural network. This optimization method is generally much faster than the standard steepest descent algorithm, and it also requires only a modest increase in memory requirements (Thawornwong et al. 2003). Finally, a single hidden layer is used, because it has been successfully employed for financial classification and prediction (Swales and Yoon 1992).

The above procedures resulted in a feed-forward neural network with three layers, including the input layer, hidden layer, and output layer. However, determining the number of nodes for the input layer and the hidden layer is a major issue in constructing neural networks. It should be noted that the number of input nodes maps the complexity (linear and/ or non-linear) of the autocorrelation structure in the data. As pointed by (Jasic and Wood, 2004), no systematic procedure or standard statistical test for non-linear dependence exists to determine this number. In this study, five input variables are used as inputs. Each of them was assigned a separate input neuron to the input layer. It is thereby assumed that these variables are sufficient to capture the dynamics of stock prices movements.

In this study, the number of hidden neurons was determined according to a usual rule of thumb in statistical modelling. This rule specifies that for a set of  $N$  observations, the degree of freedom or the number of weights in the model should not exceed  $\sqrt{N}$  (Jasic and Wood, 2004). Given that the used training sample consists of 566 observations (firm/year), an upper bound would be around 24 weights. Taking into account that (feed-forward neural network) used in this study is a fully connected network (i.e. every output from one

layer is passed along to every node in the next layer), previous choice of number of input variables and the above upper bound value, the number of hidden nodes is defined as half the number of the input nodes which approximately equals 3 neurons.

In the output layer one neuron was used. That is because there is one output, which is the sign of the stock return. The predicted direction was determined according to the highest value of the two output neurons. Now because there are 5 neurons in the input layer, 3 neurons in the hidden layer, and 1 neuron in the output layer, then the network architecture that is used in the current study is 5-3-1. The number of the network architecture weights is usually calculated according to the following equation: (Number of neurons in the input layer + 1)  $\times$  the number of the neurons in the hidden layer + (Number of the neurons in the hidden layer + Number of neurons in the output layer). Applying of this equation on this study results in  $(5+1) \times 3 + (3+1) = 22$  weights.

A stopping criterion is implemented to prevent over fitting problem, which usually occurs during network training. Another stopping technique is early stopping, but this approach needs more critical attention as this problem is harder than expected (Lawrence et al., 1997). According to the stopping criterion, a validation set is used during training instead of using only the training data set. After 100 iterations the network is tested with the validation data (i.e. the training results are observed in steps of 100 iterations). The training is then stopped as soon as the error on the out-of-sample data increases rapidly higher than the last time it was checked (Prechelt, 1998).

#### 4.1.2 Regression Analysis

Regression analysis consists of all techniques used to examine the relationship between a dependant variable and one or more independent variables. It shows how the value of the dependent variable changes as a result of changing any one of the independent variables, while the other independent variables are held fixed (Freedman, 2005). (Osborne and Waters 2002) discusses the

assumptions of multiple regression analysis. The authors identify four assumptions that are not robust to violation. The four assumptions required that variables are normally distributed, the relationship between the independent and dependent variables is linear, variables are measured without error (reliably), and the variance of errors is the same across all levels of the independent variables (homoscedasticity).

#### 4.2 The Study Sample and Period

The initial study population includes all of the industrial and service corporations listed in Amman Bourse during the period 1999-2009. The models of the study are estimated based on data pooled over firms and time. A firm-year observation to be included into the sample should satisfy the following conditions: (1) its financial statements are available for at least two consecutive years and trading data is available for the third consecutive year, For example, the financial statements of 1999 and 2000 and the trading data of 2001. It should be noted that training the neural networks and estimating the regression model is started in 2000. Financial statements of 1999 were used due to the computation of earnings change variable. Then, financial statements data of 2000 is used to predict the stock returns of 2001. These procedures were repeated for all years until 2009.

(2) No merger, acquisition, or consolidation occurred for any of the firms during the study period.

(3) Financial period of each firm ends on the 31 of December each year.

(4) The firm did not change its financial year-end.

(5) No stock stock split during the year.

(6) None of the corporations had any preferred stocks.

Applying the above conditions resulted in a final sample of 652 firm-year observations for the period from 1999 to 2009.

The sample of the study is divided into two periods. The first from 1999 – 2008 (566 observations), and the second period is the year 2009 (86 observations). During the first period ANN is trained validated and regression

model is estimated. The second period (2009) was used to test (ANNs).

Secondary data sources are used in this study, examples of which are published financial statements, annual reports, stock prices, company guide and disclosures published.

#### 4.3. Variables

The variables of this study consist of five independent variables and one dependent variable. The independent variables are:

- (Level of Earnings Per Share/ Share Price at the beginning of the period) (%) = (Net Income Pertains to Shareholders/ No. Of Subscribed Shares) × 100/ Share Price at the beginning of the period
- (Change in Earnings Per Share × 100 / Share Price at the beginning of the period) (%) = (EPS in Year<sub>t</sub> – EPS in Year<sub>t-1</sub>) × 100/ Share Price at the beginning of the period
- Dividend Yield % = (Proposed Stock Dividends × 100)/Market Capitalization
- Price to Book Value (Times) = Market Capitalization/ Total Shareholders Equity
- Return on Equity % =( Net Income Pertains to Shareholders × 100)/ Total Shareholders Equity

The dependent variable is the annual cumulative stock return, which is defined as follows:

$$CR_{it} = \left\{ \prod_{m=1}^t (1 + R_{im}) \right\} - 1$$

Where:

CR<sub>it</sub>: the annual cumulative stock return of company i for year t

π: represent victorial

R<sub>im</sub>: the monthly stock return of company i for month m

Where the monthly stock return is calculated as following:

$$R_{im} \% = \frac{(P_{im} - P_{im-1})}{P_{im-1}} \times 100$$

Where the P<sub>im</sub> is the market share price of company i at month m, and P<sub>im-1</sub> is the market share price of company i at month m-1

The above ratios are chosen according to their

usefulness to explain stock returns and to their popularity in the literature. They consist of three dimensions, which are profitability, dividends, and risk. Profitability is measured by three variables, which are EPS level, change in EPS, and ROE. Price to book value ratio is used as a proxy for risk. The five variables are expected to indicate positive stock-return.

It should be noted that artificial neural networks techniques put some limitations on the number of predictive variables relative to the number of observations used as explained in section 4.1.1 above. In addition, increasing the number of variables will result in reducing the number of observations available to be used in the study. That is because firm observations are excluded if merely one item is missing or not available for a firm.

It is assumed, but not confirmed, that all study variables can be calculated for each company four months after the fiscal year-end, since financial statements are due by that time. It is not confirmed because as a matter of fact some companies submit their financial statements after the first four months of the year. Therefore, a 12-month excess return measures are based on returns from May 1 to April 30 of the next year. For example, the financial statements of Dec 31,

2000 is assumed to be released on April 30, 2001. Then the financial ratios, which are calculated based on these statements, are used to predict the cumulative annual stock return of the period from May 1, 2001 to April 30, 2002. The following year's stock return is specified as a binary outcome, a stock return increase or a stock return decrease.

Because some companies release their annual reports to ASE more than four months after the fiscal year-end, it is possible that the prediction accuracy associated with the estimated returns is understated. That is because the financial statements would not be available for use at the assumed time.

## 5. Study Findings

This section reveals the results of the study. It presents the descriptive statistics of the study variables in the training and test samples. In addition, it reveals the results of the correlation, regression and neural networks results.

### 5.1. Descriptive Statistics

Table (1) reveals the descriptive statistics for all variables used in the training sample.

**Table I**  
**Descriptive Statistics for Training Sample**

|                   | Minimum | Maximum | Mean  | Median | Std. Deviation |
|-------------------|---------|---------|-------|--------|----------------|
| Change in EPS %   | -45.35  | 141.77  | 2.14  | 0.41   | 15.13          |
| EPS Level %       | -82.54  | 127.12  | 4.06  | 5.19   | 13.62          |
| Dividend Yield %  | 0.00    | 14.97   | 2.75  | 1.65   | 3.20           |
| P/B Value (times) | 0.18    | 7.53    | 1.54  | 1.24   | 1.04           |
| ROE %             | -74.00  | 39.25   | 5.93  | 6.31   | 12.38          |
| Stock Return %    | -76.94  | 444.83  | 16.93 | 4.71   | 59.39          |

Based on the table above, although the values of the mean and median for most of the variables are close to each other, standard deviation is high. That indicates the non-normality of the variables. In addition, the range

between the minimum and maximum values indicates that variables have high dispersion.

Table (2) reveals the descriptive statistics for all variables used in the test sample.

**Table 2**  
**Descriptive Statistics for Test Sample**

|                   | Minimum | Maximum | Mean  | Median | Std. Deviation |
|-------------------|---------|---------|-------|--------|----------------|
| Change in EPS %   | -56.17  | 31.05   | -0.97 | -0.43  | 10.07          |
| EPS Level %       | -44.21  | 15.43   | 1.15  | 2.94   | 8.89           |
| Dividend Yield %  | 0.00    | 12.06   | 2.18  | 0.00   | 3.10           |
| P/B Value (times) | 0.31    | 4.90    | 1.66  | 1.31   | 1.03           |
| ROE %             | -53.47  | 57.21   | 4.01  | 3.96   | 15.39          |
| Stock Return %    | -75.00  | 91.03   | -5.45 | -6.48  | 31.25          |

It can be seen from the above table that most of the variables are far from normality. They also have high dispersion.

### 5.2. Correlation Analysis

Table (3) reveals the correlation between stock return and fundamental accounting variables in the training sample. The results reveal a low correlation between stock return and the accounting indicators. Only the correlation between price to book value ratio and stock return, and the correlation between return on equity are

statistically significant. However, it is noticed that both variables have negative coefficient signs. One explanation of this unexpected result is that investors and traders of share stock do not rely on the values of these ratios when deciding the price of the stock. This result is in accordance with the results of (Al-Ra'i, 2001), who found that ROE has the wrong sign in many cases. Al-Ra'i also concluded that investors do not use the ROE ratio in their market valuation of common stock.

**Table 3**  
**Correlation Coefficients of the Training Sample**

|                | Change in EPS | EPS Level | Dividend Yield | P/B Value | ROE    | Stock Return |
|----------------|---------------|-----------|----------------|-----------|--------|--------------|
| Change in EPS  | 1             | 0.38**    | -0.01          | -0.04     | 0.27** | -0.04        |
| EPS Level      |               | 1         | 0.40**         | 0.11**    | 0.73** | -0.08        |
| Dividend Yield |               |           | 1              | -0.02     | 0.46** | 0.05         |
| P/B Value      |               |           |                | 1         | 0.41** | -0.22**      |
| ROE            |               |           |                |           | 1      | -0.11*       |
| Stock Return   |               |           |                |           |        | 1            |

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

Table (4) reveals the correlation between stock return and fundamental accounting variables in the test sample. The results reveal that only the correlation between Dividend Yield ratio and stock return and the correlation between ROE and stock return are significant. Both

ratios have the right sign. This indicates that correlation coefficients are not stable through time. It should be clarified that the fundamental accounting variables, used in this study, are lagged variables.

**Table 4**  
**Correlation Coefficients of the Test Sample**

|                | Change in EPS | EPS Level | Dividend Yield | P/B Value | ROE    | Stock Return |
|----------------|---------------|-----------|----------------|-----------|--------|--------------|
| Change in EPS  | 1             | 0.65**    | 0.12           | .12       | 0.51** | -0.11        |
| EPS Level      |               | 1         | 0.48**         | 0.28**    | 0.73** | -0.01        |
| Dividend Yield |               |           | 1              | 0.14      | 0.52** | 0.36**       |
| P/B Value      |               |           |                | 1         | 0.40** | 0.08         |
| ROE            |               |           |                |           | 1      | 0.23*        |
| Stock Return   |               |           |                |           |        | 1            |

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 5.3. Regression Analysis

Now to examine how accounting variables can explain the stock return. I apply multivariate regression analysis. Table (5) shows the model summary of the training sample. It can be seen from the results that

adjusted R square is equal to 5%; this means that 5% of the variation in stock return is attributed to the five fundamental lagged accounting variables used in the study.

**Table 5**  
**Model Summary - The Training Sample**

| R     | R <sup>2</sup> | Adj. R <sup>2</sup> | F     | Sig.  |
|-------|----------------|---------------------|-------|-------|
| 0.242 | 0.058          | 0.050               | 6.947 | 0.000 |

Table (6) shows the results of the training sample regression model. It can be seen from Table (6) that there is almost no statistically significant relationship between stock return and the accounting variables, except for the P/B. This result is consistent with the correlation results.

Table (7) shows the model summary of the testing sample. It can be seen from the results that adjusted R square is equal to 18.5%. The difference between the explanatory of the training sample model and the testing sample model may be due to the effect of pooling samples.

**Table 6**  
**Regression Analysis results of Training Sample**

|                | Unstandardized Coefficients |            | Standardized Coefficients | T      | Sig.  |
|----------------|-----------------------------|------------|---------------------------|--------|-------|
|                | B                           | Std. Error | Beta                      |        |       |
| (Constant)     | 33.719                      | 5.217      |                           | 6.464  | 0.000 |
| Change in EPS  | -0.066                      | 0.179      | -0.017                    | -0.369 | 0.712 |
| EPS Level      | -0.407                      | 0.284      | -0.093                    | -1.434 | 0.152 |
| Dividend Yield | 1.447                       | 0.901      | 0.078                     | 1.607  | 0.109 |
| P/B Value      | -12.748                     | 2.781      | -0.223                    | -4.584 | 0.000 |
| ROE            | 0.114                       | 0.351      | 0.024                     | 0.325  | 0.745 |

**Table 7**  
**Model Summary - The Test Sample**

| <b>R</b> | <b>R<sup>2</sup></b> | <b>Adj. R<sup>2</sup></b> | <b>F</b> | <b>Sig.</b> |
|----------|----------------------|---------------------------|----------|-------------|
| 0.483    | 0.233                | 0.185                     | 4.869    | 0.001       |

Table (8) shows the results of the testing sample regression model. It can be seen from Table (8) that there are three statistically significant relationships between stock return and the accounting variables. These relationships are between stock return and EPS level,

Dividend Yield, and ROE. This result is consistent with the correlation results, except for EPS level. The comparison between the results of the training sample model and the testing sample model reveals that regression models are not stable through time.

**Table 8**  
**Regression Analysis results of Test Sample**

|                | Unstandardized Coefficients |            | Standardized Coefficients | T      | Sig.  |
|----------------|-----------------------------|------------|---------------------------|--------|-------|
|                | B                           | Std. Error | Beta                      |        |       |
| (Constant)     | -15.288                     | 6.457      |                           | -2.368 | 0.020 |
| Change in EPS  | -0.210                      | 0.430      | -0.068                    | -0.489 | 0.626 |
| EPS Level      | -1.473                      | 0.604      | -0.419                    | -2.438 | 0.017 |
| Dividend Yield | 3.775                       | 1.247      | 0.375                     | 3.026  | 0.003 |
| P/B Value      | 0.032                       | 3.267      | 0.001                     | 0.010  | 0.992 |
| ROE            | 0.763                       | 0.327      | 0.376                     | 2.334  | 0.022 |

The results of the regression models above show a low explanatory power for the accounting variables. Based on this power, artificial neural networks will be examined to see whether it is able to predict the sock return one year ahead.

networks, first it was applied on the training sample. The number of observations in the training sample is 566 observations. They were divided into two groups according to the stock return. There were 306 observations with positive return and 260 observations with negative return. Table (9) shows the results of the prediction power of the neural networks in the training sample.

#### 5.4. Artificial Neural Networks

To examine the prediction power of the neural

**Table (9)**  
**Prediction Output in the Estimation Sample**

|                             |          | Predicted Sign of Stock Return |              | Total       |
|-----------------------------|----------|--------------------------------|--------------|-------------|
|                             |          | Positive                       | Negative     |             |
| Actual Sign of Stock Return | Positive | 283<br>92.5%                   | 23<br>7.5%   | 306<br>100% |
|                             | Negative | 89<br>34.2%                    | 171<br>65.8% | 260         |
| Total                       |          | 372                            | 194          | 566         |

It can be seen from Table (9) that neural networks are able to predict the direction of stock return with

accuracy of 92.5% for the positive return companies and 65.8% for the negative stock return companies. The

overall accuracy is 80.2%. This percentage was calculated by dividing the correct predicted observations (283+171) over total observations (566). It can be noticed that error type I (Predicting an actual negative return as a positive return) is higher than error type II (Predicting an actual positive return as a negative return). Error type I equals 34.2% and error type II equals 7.5%. It should be noted that error type I is more costly than error type II. In the case of error type II investors lose income. Meanwhile, in the case of error

type I they lose their money.

In order to examine the stability of the prediction accuracy rates, a test sample was used to examine how it works outside the period of training. The number of observations in the test sample is 86. They were divided into 33 observations with positive stock return and 53 observations with negative stock return. Table (10) shows the results of the prediction power of the neural networks in the test sample.

**Table (10)**  
**Prediction Output in the Test Sample**

|                             |          | Predicted Sign of Stock Return |             | Total      |
|-----------------------------|----------|--------------------------------|-------------|------------|
|                             |          | Positive                       | Negative    |            |
| Actual Sign of Stock Return | Positive | 29<br>87.9%                    | 4<br>12.1%  | 33<br>100% |
|                             | Negative | 32<br>60.4%                    | 21<br>39.6% | 53<br>100% |
| Total                       |          | 61                             | 25          | 86         |

It can be seen from Table (10) that the accuracy rates of the positive stock return companies are approximately stable. However, error type I increased dramatically. The overall accuracy rate is 58.1%. This percentage was calculated by dividing the correct predicted observations (29+21) over total observations (86). A justification of low accuracy percentage of error type I may be due to unclear pattern of the relationships between independent variables and dependent variable in the negative stock return observations. It should be clarify that neural networks used in this study works on learning the pattern exist between variables.

### 6. Summary and conclusion

The purpose of this study was to examine the ability of artificial neural networks to predict the direction of one-year-ahead stock return for the industrial and service Jordanian corporations. In addition, it was aimed to examine the stability of the prediction accuracy rates. Descriptive analysis, correlation analysis, regression analysis and artificial neural networks were applied on a

sample contains 652 observations for the years 1999 - 2009.

The results show that the performance of neural networks in predicting direction of stock return is promising. It was shown that most of the correlations between the variables of the study are low. In addition, it was seen that independent variables explain the dependent variable poorly. Despite that the overall prediction accuracy of the neural network in the training sample was 80.2%. In addition, it should be emphasized that no actions were done to manipulate or improve data, such as deleting outliers, transforming data, or approaching normality.

In order to examine the stability of the prediction accuracy rates of neural networks, a test sample lays outside the training period was used. The results of the test show that error type I increases dramatically. This limits the possibility of using neural network model developed in a specific period to predict the direction of stock returns in another period. It should be emphasized that the real achievement of any prediction tool, not only neural networks, is to predict correctly the desired output

outside the training sample.

Therefore, based on the study results, the researcher concludes that it is not difficult to utilize artificial neural networks to predict the direction of stock return. It should be emphasized that information availability in the

market does not mean that all investors will handle this information in the same way. Therefore, new techniques and technology may be helpful for investors who utilize it until it becomes common.

## REFERENCES

- Abdelqader, M., M. Al-Debi'e, and F. Gharaibeh. 2008. The Relation between Earnings and Stock Returns in the Long Run: Replication and Extension. *Jordan Journal of Business Administration*, 4(1): 24-44.
- Abu Nassar, M. and M. Al-Debi'e. 2002. The Information Content of the Main Elements of the Income Statement. *Dirasat: Administrative Sciences*, 29(2): 278-296.
- Alawi, S. and F. Gharaibeh. 2008. Predicting Corporate Distress in Jordan: A Multidimensional Scaling Approach. *Dirasat: Administrative Sciences*, 35(2): 377-397.
- Al-Khadash, H. and M. Al-Abadi. 2005. The Relationship between Accrual-Based Values, Cash-Based Values and Stock Prices. *Dirasat: Administrative Sciences*, 32(1): 141-153.
- Al-Khalayleh, M. A. 2001. The Association Between Market Based Performance Measures and Accounting Performance Measures in the Long Run. *Dirasat: Administrative Sciences*, 28(1): 93-110.
- Al-Ra'i, Z. 2001. The Impact of the Accounting Rates of Return on Equity and Market Valuation of Common Stocks "Empirical Investigation". *Dirasat: Administrative Sciences*, 28(2): 464-477.
- Al-Rjoub, S., Kh. Alzubi, and M. A. 2006. Stock Market Returns, Volatility and Volatility Persistence: Evidence from Emerging Markets. *Jordan Journal of Business Administration*, 2(2): 326-342.
- Al-Shiab, M. and A. Al-Alawneh. 2007. Common Shares Performance Evaluation for Companies Listed at Amman Stock Exchange. *Dirasat: Administrative Sciences*, 34(1): 210-220.
- Al-Shiab, M. and A. Al-Ali. 2006. Common Stock Appraisal in Relation to Their Price-Earning Ratios Using Risk-Adjusted Measures: An Emerging Market Perspective. *Jordan Journal of Business Administration*, 2(1): 115-133.
- AVCI, Emin. 2009. Stock return forecasts with artificial neural network models. *Mamara Universitesi*, 26 (1): 443 – 461.
- Cao, Q., K. B. Leggio, and M. J. Schniederjans. 2005. A Comparison between Fama and French's Model and Artificial Neural Networks in Predicting the Chinese Stock Market. *Computers & Operations Research*, 32: 2499 – 2512.
- Clare, A. D., R. Priestley, and S. H. Thomas. 1997. Stock return predictability or mismeasured risk?. *Applied Financial Economics*, 7: 679 – 687.
- Fama, Eugene F., and Kenneth R. French. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics*, 22: 3 – 25.
- Faria, E. L. de, Marcelo P. Albuquerque, J. L. Gonzalez, J. T. P. Cavalcante, and Marcio P. Albuquerque. 2009. Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. *Expert Systems with Applications*, 36: 12506 – 12509.
- Freedman, David A. 2005. *Statistical Models: Theory and Practice*, Cambridge University Press.
- Garaibeh, F. and R. Abdelatif. 1987. Using Financial Ratios to Predict the Failure of Industrial Jordanian Corporations. *Dirasat: The Humanities*, 14(3): 33-66.
- Haddad, F. 2003. Using Financial Ratios for Measuring Firms' Systematic Risk Empirical Study on Jordanian Manufacturing and Servicing Firms. *Dirasat: Administrative Sciences*, 30(2): 314-321.
- Haddad, F. and Kh. Abu Alulla. 2004. The Relationship between Economic Value and Abnormal Returns:

- Empirical Study on Jordanian Industrial Firms. *Dirasat: Administrative Sciences*, 31(1): 1-9.
- Hagen, M. T., H. B. Demuth, and M. Beale. 1996. *Neural Network Design*. Boston, MA: PWS Publishing Company.
- Holthausen, R. W. and D. F. Larcker. 1992. The Prediction of Stock Returns Using Financial Statement Information. *Journal of Accounting and Economics*, 15: 373 – 411.
- Jasic, T. and D. Wood. 2004. The Profitability of Daily Stock Market Indices Trades Based on Neural Network Predictions: Case Study for the S&P 500, the DAX, the TOPIX and the FTSE in the period 1965-1999. *Applied Financial Economics*, 14: 285 – 297.
- Juhmani, O. and Al-Daoud. 2004. Predicting Industrial Corporate Failure in Jordan Using A Multidimensional Scaling Approach. *Dirasat: Administrative Sciences*, 31(2): 209-233.
- Keim, Donald B., and Robert F. Stambaugh. 1986. Predicting returns in the stock and bond markets. *Journal of Financial Economics*, 17: 357 – 390.
- Khamees, B. and F. Garaibeh. 1990. Using Financial Ratios to Predict Highly-successful and less-successful Industrial Jordanian Corporations. *Dirasat: The Humanities*, 17(3): 202-228.
- Lawrence, S., C. L. Giles, and A.C. Tsoi. 1997. Lessons in Neural Network Training: Training May be Harder than Expected. *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, AAAI-97, (pp. 540-545), Menlo Park, California: AAAI Press.
- Malliaropoulos, D. 1996. Are long-horizon stock returns predictable? A bootstrap analysis. *Journal of Business Finance & Accounting*, 23(1): 93 – 106.
- Morton, Richard M. and Philip B. Shane. 1998. The Information Environment and The Ability of logit-based Financial Statement Analysis to Predict Abnormal returns. *Accounting and Finance*, 38 (2): 71 – 89.
- Nelson, M. and W.T. Illingworth. 1991. *A Practical Guide to Neural Nest*, Addison-Wesley Publishing Company, New York.
- Nour, A. N. and M. Al-Fadil. 2003. Relationship Between Dividends and Abnormal Market Returns for Shares and Its Profitability, Volume and Degree of Non-systematic Risk: A Comparative Study Between Iraqi and Jordanian Corporations. *Dirasat: Administrative Sciences*, 30(1): 183-197.
- Omet, G. and A. Bino. 2000. On the Assessment of Risk and Return in Amman Financial Market. *Dirasat: Administrative Sciences*, 27(1): 233-241.
- Osborne, Jason & Elaine Waters. 2002. Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, 8 (2). Retrieved September 17, 2010 from <http://PAREonline.net/getvn.asp?v=8&n=2>
- Ou, J. and S. Penman. 1989. Financial Statement Analysis and the Prediction of Stock Returns. *Journal of Accounting and Economics*, 11: 295 – 329.
- Patelis, Alex D. 1997. Stock return predictability and the role of monetary policy. *The Journal of Finance*, 52 (5): 1951 – 1972.
- Prechelt, L. 1998. Early Stopping-but when? *Neural Networks: Tricks of the trade*, pp. 55-69. Retrieved August 14, 2010 from World Wide Web: <http://www.wipd.ira.uka.de/~prechelt/Biblio/>
- Safer, Alan M. 2003. A comparison of two data mining techniques to predict abnormal stock market returns. *Intelligent Data Analysis*, 7: 3 – 13.
- Saleh, W. 2007. Earning-to-Price, Dividend-to-Price, Firm Growth and Stock Returns. *Jordan Journal of Business Administration*, 3(1): 85-105.
- Skapura, David M. 1995. *Building neural networks*, Addison-Wesley Publishing Company, New York.
- Swales, G. S., and Y. Yoon. 1992. Applying Artificial Neural Networks to Investment Analysis. *Financial Analysts Journal*, 48: 78 – 80.
- Thawornwong, S., D. Enke, and C. Dagli. 2003. Neural Networks as a Decision Maker for Stock Trading: A Technical Analysis Approach. *International Journal of Smart Engineering System Design*, 5: 313 – 325.

:

\*

.2009

1999

%58.1

%80.2

:

\* The University of Jordan, Faculty of Business  
Amman, Jordan  
E-mail: basher@ju.edu.jo

---

.2012/4/16

2012/1/15