Suggested Approach to Stocks' Selection for Enhancing Portfolio Performance

Mona N. Abdel Bary
Mona_nazihali@yahoo.com

Abstract—financial time series data are a sequence of prices of some financial assets over a specific period of time [10]. In this paper we will focus our research on the most volatile and challenging financial market; the stock market. All the financial analysts carry out many kinds of research in the hope of accomplishing one goal; to beat the market. To beat the market means to have a rate of return that is consistently higher than the average return of the market while keeping the same level of risk as the market. This paper introduces approach to stock'selection for enhancing portfolio performance.The process of selection of a portfolio can be divided into two main stages; namely evaluation of the performance of available stocks and the choice of the portfolio. Markowitz (1952), (1959) is concerned with the second stage in which the portfolio is selected only without evaluation and stocks' selection [7], [8]. Hence, an emphasis should be given to the first stage to enhance the second stage performance. There are two major problems affecting the optimization portfolio process; outlier and multicollinearity problems. To explain this, the outlier problem leads to wrong expectation of both the rate of the return and the rate of the risk. Therefore, this problem may result in wrong selection of the optimal portfolio. Additionally, the multi-collinearity problem causes increase of risk in a certain manner in which the drop down of one stock in the portfolio leads to the drop down of the whole portfolio. It is necessary to introduce the approach of stocks' selection in a portfolio to avoid the portfolio drop problem, and to avoid the bad effect of the stock market movement, and to gain an un-expectation return. Thus, this paper aims at introducing complete analysis of the stock market which presents a useful analysis for all investors of the stock market in order to be able to analyze the outlier problem, determine the multiicollinearity stocks groups, determine the stocks group controlling the stock market movements, and the outstanding stocks group, constrain the sum weights of each group in the portfolio frame to enhance the portfolio performance.

Index Terms: Modern Portfolio Theory, Outliers, Multi-collinearity, Principle Component, Principle Component Regression, Egyptian Stock Market.

1 Introduction

The process of selection of a portfolio can be divided into two main stages; the first is performance evaluation of the available stocks and the second is the choice of the portfolio. As an illustration of this, Markowitz (1952) is concerned mainly with the second stage in which the portfolio is selected based on Modern Portfolio Theory, also known as mean- variance portfolio optimization, which was introduced by Markowitz [7]. This approach provides the techniques for creating a set of portfolios that realizes optimal return for a given level of portfolio risk.

Markowitz focuses only on the selection of the portfolio without stocks' evaluation and selection. Markowitz refers to the fact that if an investor invests in a portfolio which positively yields correlated returns, then it does not at all lower his risk, since the returns move in only one direction and the investor in such a portfolio can suffer significant losses. If in case, portfolio has negatively correlated return, then the returns have an inverse movement. Time series analysis is essential to study the stock market behavior to define its characteristics [10]. This study focuses mainly on analysing the outlier and suggesting an adequate solver for it, studying of multicollinearity among the stocks and determining of multicollinearity groups, defining the most stocks affected on the stock market movement. Finally, this study sheds light on defining of the outstanding stocks. Thus, an approach for putting constrains on the selection of the stocks to construct optimal portfolio is introduced as the end to this paper and start to the approach of construct portfolio. The newly born and non-efficient markets have a good advantage that it is possible to gain an exceptional return by studding the market and determining good chances.

2 Literature Review

There are main theories discussing portfolio such as Modern Portfolio Theory, also Known as mean-variance portfolio optimization, which is introduced by Markowitz (1952). It provides the techniques for creating a set of portfolios that realize optimal return for a given level of portfolio risk. According to Markowitz's theory the investor is risk averse, the investor will create a portfolio with the aim of achieving the largest return for the minimum risk. There turn of a portfolio according to Markowitz's theory worked at first from determining the expected return of one asset and then from the expected return of the whole portfolio..

$$E(R_i) = \sum_n P_n R_n \qquad (1)$$

Where $E(R_i)$ is the expected return of the asset i; and P is the probability of occur the return R. The total portfolio return is the weighted average of the individual returns of the stocks in the portfolio,

$$E(R_i) = \sum_{i=1}^{N} W_i R_i \qquad (2)$$

$\sum W_i = 1$ and R_i is the expected rate of return for asset i. An investor is invested not only in the rate of the return but also in the risk. In measuring risk Markowitz at first works from

ISSN 2229-5518

the risk of one asset and then from the risk of the portfolio. Since the risk of each asset i in the portfolio is calculated as follows:

$$\sigma_i = \sqrt{\sum_i^n (R_n - E(R_i))^2 P_n}, \quad (3)$$

where P_n is the probability of the possible rate of return R_i. The risk of a portfolio however is not simply a weighted average of the risks of individual stocks in the portfolio. The degree of risk of the portfolio is influenced also by other variables, in particular by the mutual relation between the returns of individual stocks

$$\sigma^2 = (R_p - E(R_p))^2 \quad (4)$$

$$\sigma^2 = \sum_{i=1}^N \sum_{j=1}^N W_i W_j \sigma_{ij}$$
$$\sigma^2 = \sum_{i=1}^N \sum_{j=1}^N W_i W_j \sigma_i \sigma_j \rho_{ij}$$

where σ_{ij} is the covariance between the returns of the *i* and *j* stock, and ρ_{ij} is the correlation between returns of the *i* and *j* stocks. This covariance in relation to the standard deviation of the *i* and standard deviation of the *j* stock gives us the correlation between the returns of these stocks.

$$\rho_{ij} = \sigma_{ij} / \sigma_i \sigma_j \quad (5)$$

The coefficients of covariance range from -1 and +1 and express the direction of the correlated movement of investments in the portfolio. If the covariance has a positive value, it means that the returns of investments have the same direction of movement. An inverse relationship between the returns of investments exists in the case where the covariance has a negative value. The covariance has a zero value where there turns moving independently. Markowitz refers to the fact that if an investor invests in a portfolio which perfectly positively has correlated returns, then it does not at all lower his risk, because the returns move in only one direction and the investor in such a portfolio can suffer significant losses. In case, if a portfolio has negatively correlated return, then the returns have an inverse movement. Assets with non-correlated returns create a portfolio in which the returns have no relation to one another. The benefits of diversification lie in the fact that a more efficient compensation effect of risk and return will be achieved through an appropriate combination of assets, the correlation of which does not extend to a form of completely positive correlation. Diversification lowers risk quickly also in the case of a small number of stocks gradually with an increasing amount of stocks the effectiveness declines. Since, from Markwitz's selection model, if an investor wants to reduce the overall risk of the portfolio, he/she must combine those assets which are not perfectly positively correlated. Moreover, they should not be also perfectly negatively correlated. Markowitz highlights the fact that the investor can select a portfolio within the market various combinations of stocks with various returns and risks. This means that, there is feasible set of all possible combinations of investment, which an investor is faced with in the market. From the set of Pareto optimal combinations of expected returns and variances, investors will according to Markowitz select portfolios which give the maximum expected rate of return at various levels of risk or offer minimum risk in the case of various levels of expected rates of return. The set of portfolios fulfilling these two conditions is known as the efficient set or efficient frontier. This limit depicts the points with the maximum rate of return for a given level of risk, which are measured by the standard deviations of the portfolio's returns, by applying indifference curves, which from the aspect of the theory of frontier utility express the various combinations, in the case of which an investor tries to achieve the same utility. As Markowitz states, indifference curves have different slope in the case of a risk averse investor. The indifference curves of an investor seeking risk have more moderate slope. Since in selection model an investor gives preference to the risk averse investor, the contact of his indifference curve with the efficient frontier of the portfolio creates the optimal portfolio for the investor. Thus, Modern Portfolio Theory defines an efficient frontier of optimal portfolio which is a set of portfolios that maximize the expected return for a given level of risk or that minimize risk for a given level of return. When the portfolio return equation is solved to obtain the maximum return of the portfolio, the portfolio risk of the portfolio return is held constant. The final step in portfolio construction is to combine an investor's utility function with the efficient frontier graph. The optimal portfolio for the investor is the indifference curve that is tangent to the efficient frontier. Markowitz believes that quadratic programming algorithm can be used to find the portfolio that provides the maximum expected return for standard deviation of return [13].

Modern portfolio theory determines risk through the use of variance, standard deviation of expected [7]. Markowitz believes that quadratic programming algorithm can be used to find the portfolio that provides the maximum expected return for standard deviation of return [13]. CAPM is the most widely recognized explanation of stock prices and expected return. It states that systematic risk is the main factor to expected return. The CAPM allows for separation of risk from return and it is regarded for determining the required rate of return for assets and portfolio. Hence, the CAPM predicts the expected return of security given the expected return on the market, the security's beta, and the risk free rate. The market has beta value of one. A risk free asset has beta value of zero. The introduction of risk free security to the Markowitz model changes the efficient frontier from a curved line to a straight line called the Capital Market line (CML). CAPM is not adequate for the non efficient market because of the CAPM assumptions. Sharp (1992) aims a finding the best a set of asset class exposures by using of quadratic programming for the purpose of determining a fund's exposures to changes in their turns of major asset classes is termed style analysis [12]. In addition to that, the style identified in such an analysis is an average of potentially changing styles over the period covered. The deviations of the fund's return from that of style itself can arise from the selection of specific securities within one or more asset classes, or rotation among asset classes, or both stock selection and asset class rotation. According to Markowitz (1952) determining the efficient set from the investment opportunity set, these to fall possible portfolios, requires the formulation and solution of a parametric quadratic program [7]. Among the studies addressing the issue of portfolio was the study [6] using linear programming method for investment allocation that would enable creating an optimal portfolio under an effective investment strategy to prove the inefficiency of the Egyptian

ISSN 2229-5518

Stock Market. Deng et al focuses on the problem of optimal portfolio and equilibrium when the target is to maximize the weighted criteria under the worst possible evolution of the rates returns on the risky assets [4]. The optimal portfolio was analytically presented, which can be obtained using linear programming technique. Lai et al focuses on the use Genetic algorithm to identify good quality assets in terms of asset ranking [5]. Zhang, et al discuss the portfolio selection in which there are exit both probability constraint on the lowest return rate of the portfolio and upper bounds constraints on investment rates to assets [14]. Mitra et al illustrate that mean-variance rule for investor behavior that implies justification of diversification, is affected by risk adverse investors [9]. Chi-Ming Lin et al introduce the following model to obtain the optimal portfolio [3]:

$$max \sum_{i=1}^{n} \mu_i w_i \qquad (6)$$

$$min \sum_{i=1}^{n} \sum_{i=1}^{n} \sigma_{ii} w_i w_i \qquad (7)$$

$$s.t. \sum_{i=1}^{n} w_i = 1, w_i \geq 0, i = 0, 1, 2, \ldots, n.$$

The objective function (6) is to maximizing the total expected return. And the objective function (7) is to minimizing the total risk of the portfolio.

From the literature review; most of the studies deal with evaluation and selection of the portfolio as mentioned in modern theory but there is no theory for selecting the stocks. So, this study sheds intensive light on introducing an approach to select the stocks to enhance the portfolio performance. So, we need to approach for selecting the stocks can be do the following aims:

1. Selection of stocks should avoid the drop of the portfolio problem by solving the multi-collinearity among the stocks into the portfolio. The multi-collinearity among the stocks leads to that drop one of the stocks in the portfolio leads to drop all the portfolio. Therefore, studding the multi-collinearity among the stocks, determining the multi-collinearity stocks groups and putting investment boundary in it are very essential.
2. Studding the stock market movement to determine the important stocks which control on the stock market movement as a package and increasing or decreasing the boundary of the investment in it in spot of the state of the market.
3. Studding the chances of gaining an un-expectation return; gaining the most return and at the same time the lowest risk. To do this the evaluation stocks measures are required. The stock which has the most return is attractive stock and the stock which has less of the risk is attractive stock. But the stock which has the most return and at the same time lowest risk is an outstanding stock. To do a good evaluation for selecting the stock, we suggest the ratio between the return and the risk to avoid the problem of there are not equivalent between the increasing of the return of the stock and the increasing in its risk.

3 The Study Problem:

The main problem of this study is lack of approach for selecting the stocks can be solved the following problems:

1. Risk of drop of the portfolio.
2. Risk of the dynamic movement of the stock market.
3. Loss of gain un-expectation return.

4 Objective of the Study:

This paper aims to introduce approach for selecting the stocks into the portfolio which includes the following elements:

1. Introducing solve for the drop portfolio.
2. Introducing solve for the risk of the dynamic movements of the stock market.
3. Introducing measurement for the stock performance.

5 Importance of the Study:

This paper is important because of the following reasons:

1. Lack of studies related to stocks selection into the portfolio.
2. The outliers and multi-collinearity are the major problems which case wrong on the evaluation of the stocks behind to case a higher on the invest risk. Our suggested approach introduces analysis and solving to both of them.
3. Avoid the risk of portfolio drop, risk of the market movements, and loss of the un-expectation return.
4. Defining the characterizes of the stock market is useful for both of individual stocks investors and portfolio investors.
5. The majority of studies focus mainly on the efficient markets which aren't adequate the newly born market and noneffective market.

6 Planning of the Study:

This paper is organized as follows; the following section introduces selection of data set. The next section deals with the problem to be solved and implemented algorithms to find the solution and characterize the solution and its implications. The paper ends with conclusions.

7 Data of the Study:

Our study includes 45 stocks from the highest 100 stocks on the Egyptian Stock Market which have continuous price time series. It focuses on the close price for a period extension from January /2000 to April/2008 and uses the monthly data. The source of data is Egyptian Stock Market**.**

8 Primaries Time Series Analysis:

This study begins with presenting the graphs for the time series of the stocks' prices. The outlier is fundamental problem but it differs from one case to another. Increase of stock return to any level do not present problem for the investor but decrease of stock return is a dangerous problem. So, the price time series is essential to define the nature of the price time series movements. It is obvious from the graphs. There is a large drop at the price of some stocks such as Asset

ISSN 2229-5518

4 , Asset 5 , Asset 12, Asset 13 , Asset 15 , Asset 22, Asset 25 , Asset 26 , Asset 32 , and Asset 33 . This drop leads to changing the level of the stock price after this drop. This drop appears as outlier return point. By defining why this outlier occurred, it is found that that these companies use a strategy which splits the stocks. Stock split leads to lessening the price of the stock without lessening the investment capital. Then, this strategy does not lead to loss of any investment capital and may lead to increase the rate of the return. So, it is not present any risk but if we don't take care in the analysis this outlier leads to bad forecasting for prices and rate of the return and the risk. So, this wrong leads to wrong on stock's rate of the return, leads to wrong on the stock's expected return and leads to wrong on the stock's risk. All the previous lead to wrong on the valuation of stocks and finally wrong on the portfolio construct. For all the pervious, the analysis of the stock price time series is important. In case of splitting the stock, we suggest that there is not need to change the rate of the return and the risk for this observation "outlier by split the stock", and let the rate of the return and the risk of this observation equal to the previous observation's rate of the return and risk. The rate of the stock is calculated as:

$$R_S = (P_{n+1} - P_n)/P_n \quad (8)$$

1. There is large increase on the price such as Asset5, Asset 12, Asset 16, Asset 36, Asset37, Asset 39, Asset 40, Asset 43, and Asset45which doesn't problem but may be attract the investor to invest in these stocks. This kind of outlier return does not present problem for the investor but may cause problem on forecasting both of the stock prices and return calculation of the risk. In this case, price time series should be divided into two samples; the first sample is before the outlier and the second sample is after the outlier. It should be consider with the sample after the change if it is considered with the current price forecasting.
2. There are stocks's price increase and decrease into narrow limit such as Asset1, Asset14, Asset27, Asset31, Asset 34 , Asset 38 , and Asset 42 or systematic increasing such as Asset 29 and Asset 30. So, these stocks don't include outlier return points.

9 Suggested Approache to Selection Stocks:

9.1 The Drop Portfolio Problem:

As previously mentioned in the above sections; Markowitz (1952, 1959) provides the techniques to create a set of portfolios that are optimal in the sense that they maximize portfolio return for a given level of portfolio risk. Markowitz refers to the use of the variance or the standard deviation of the portfolio as a measure for the risk: Markowitz indicates to use the variance or the standard deviation of the portfolio as measure for the risk.

According to Markowitz's theory, the risk of a portfolio.

$$E(R_p) = \sum_{i=1}^{N} W_i R_i \quad (9)$$

$$\sigma^2 = \sum_{i=1}^{N}\sum_{i=1}^{N} W_i\, W_j \sigma_{ij} \quad (10)$$

For simply assume, we have three stocks only:

$$R_p = W_1^2\sigma_{11} + W_2^2\sigma_{22} + W_3^2\sigma_{33} + 2W_1W_2\sigma_{12} + 2W_1W_2\sigma_{13} + 2W_2W_3\sigma_{23} \quad (11)$$

So, the risk of the portfolio is:

$$R_p = \sum_{i=1}^{n} W_i^2\,\sigma_{ii} + \sum_{i=1}\sum_{i<i\in N} W_i W_i + \sum_{i=1}\sum_{i<i\in P} W_i W_i \quad (12)$$

Where $N = \{\sigma_{ij}|\sigma_{ij} > 0\}$ and $P = \{\sigma_{ij}|\sigma_{ij} < 0\}$. From the previous; we object to decrease the volatility of the portfolio return so that we object to decrease the weights of the stocks that have R1, increase the weights of the stocks that have negative covariance and decrease the weights of the stocks that have positive covariance to avoid the drop portfolio problem so that we object to increase R2, and decrease R3. But, the minimization of the portfolio risk avoids the pair wise correlation only. Computing all pair wise correlation coefficient is necessary but not sufficient for the detection of collinearity. So that, multi-collinearity groups should be defined and the sum of the weights of each group on the portfolio should be constrained. By our suggested we avoid the drop portfolio problem. We will do regression for Market Index with all the 45 stocks prices.

TABLE 1
THE REGRESSION PARAMETERS (LR)

Stocks	B	Stocks	B	Stocks	B
Stock1	0.0640	Stock16	11.8000	Stock31	8.2800
Stock2	0.9450	Stock17	2.83000	Stock32	-1.180
Stock3	-2.460	Stock18	7.46000	Stock33	0.1390
Stock4	-0.272	Stock19	4.70000	Stock34	-6.940
Stock5	2.2800	Stock20	-5.2100	Stock35	-2.070
Stock6	-0.930	Stock21	7.6200	Stock36	-2.860
Stock7	9.4000	Stock22	-2.190	Stock37	19.300
Stock8	1.0700	Stock23	0.5820	Stock38	2.3600
Stock9	0.9600	Stock24	-9.370	Stock39	1.5000
Stock10	-3.230	Stock25	-0.144	Stock40	-7.460
Stock11	-2.720	Stock26	-0.490	Stock41	-3.180
Stock12	-1.370	Stock27	1.7200	Stock42	15.400
Stock13	0.1700	Stock28	-10.40	Stock43	-0.721
Stock14	2.0700	Stock29	5.8200	Stock44	-3.180
Stock15	2.8000	Stock30	0.7590	Stock45	1.2800

$$R^2(adjusted) = 0.997$$

Assumptions of the Ordinary Least Squares Regression which should be validated:

1. **Linearity:** The model should be linear in theparameters:
$$Y = XB + \varepsilon$$

ISSN 2229-5518

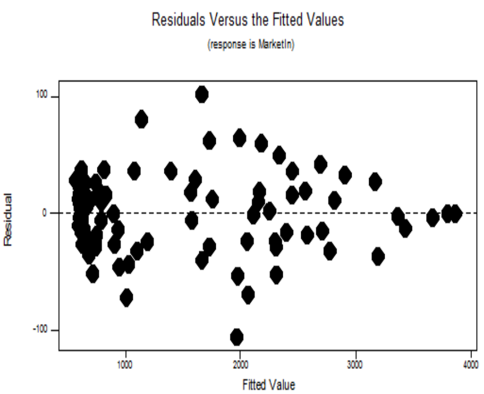

Fig. 1. Fig1: The plot of Residual vs. fitted values results in a random scatter plot which validates linearity and homogeneity assumptions.

2. **Errors:** The errors should be normally distributed, independent of each other, have zero mean and constant (but unknown) variance; i.e.:
$$\varepsilon \backsim N(0, \sigma^2 I)$$

The normal probability plot results in a straight line through the origin with a slope equal to 1. Then, the assumption of the normality of residuals holds.

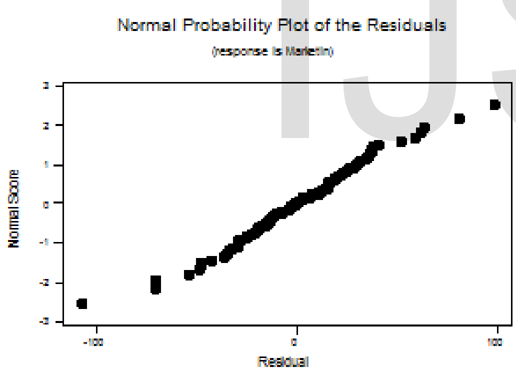

Fig. 2. Normal Probability Plot of the Residuals.

The index plot of residuals results in a randomly scatered horizontal band around the zero which means that theresiduals are not autocorrelated i.e.: independent, their mean is zero with constant variance.

3. **Unusual observations:** The Ordinary Least Squares Regression assumes that the observations are equally reliable and have equal influence in determining the Least Squares results. Furthermore, some of the observations seem to have high residuals. Namely; observations number 60, 67, 69, 71, 72, 77, 81, 83, 84, 87, 89. These observations are outliers which have a large standardized residual. Observationnumber 100 is outlier X-space which has large influence.

TABLE 2
UNUSUAL OBSERVATIONS

Observation	Asset1	MarketIndex	Residual	St Resid
60	89	1221.17	80.48	2.42R
67	97	1768.71	97.89	2.95R
69	144	1996.48	-71.04	-2.37R
71	144	2055.64	61.03	2.59R
72	149	2258.05	-53.69	-3.28R
77	145	1860.11	-107.33	-3.30R
81	155	2239.65	63.19	3.81R
83	139	2280.15	-28.95	-2.09R
84	141	2380.84	51.97	3.13R
87	150	2483.02	35.04	2.11R
89	166	2744.25	-35.52	-2.36R
100	299	3857.92	1.01	0.25X

R: denotes an observation with a large standardized residual and X denotes an observation whose X value gives it large influence.There are measures based on remoteness of points in X-Y space such as the ith diagonal element of the prediction matrix. It should examine the high leverage points by calculating leverages and plotting them to more details Chatterjee and Hadi (1988). The average leverage represents the number of the variables, and n represents the number of the observations. $P_{ii} = 45/100$. The observation with $P_{ii} > 2 * (k/n)$ =0.9 considers a high leverage point. Hence, we conceder the observation with $P_{ii} > 0.8$ considers a high leverage point. Then, we detect 19 high leverage points.

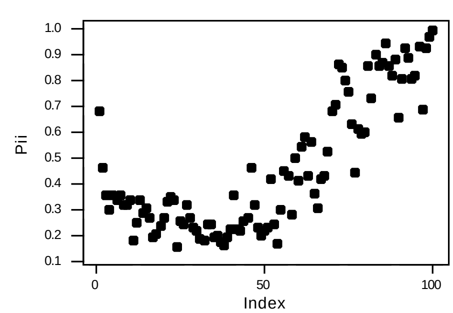

Fig. 3: Leverages value

4. **Independence of stocks' prices:** Stocks' prices are supposed to be linearity independent. Computing all pairwise correlation coefficient is necessary but not sufficient to detect collinearity. Hence, testing the multi-collinearity is essential to be sure about the independence of the stocks. Calculation Kappa to discover whether there is multi-collinearity between the stocks' prices or

not. The principle Component Analysis is used to get the Eigen Values and Eigen Vectors. The large values of Eigen values are selected and kappa values are calculated.

$$K_i = \sqrt{\lambda_i/\lambda_i} \qquad (13)$$

If K_i > 10 then, there are multi-collinearity. From the values of Kappa, there are sets of collinearity existing in the data. To define which of the variables are involved in each set, the Eigen Vectors that it is from V_{16} to V_{45} are used and the large values on them are chosen.

TABLE 3
THE KAPPA'VALUE

Eigen Value	Kappa value	Eigen Value	Kappa value	Eigen Value	Kappa value
28.08	01.00	00.20	11.85	00.00	Inf
10.50	01.64	00.20	11.85	00.00	Inf
07.00	02.00	00.20	11.85	00.00	Inf
05.70	02.22	00.20	11.85	00.00	Inf
03.20	02.92	00.10	16.76	00.00	Inf
02.00	03.75	00.10	16.76	00.00	Inf
01.70	04.06	00.10	16.76	00.00	Inf
01.30	04.65	00.10	16.76	00.00	Inf
01.20	04.84	00.10	16.76	00.00	Inf
01.00	05.29	00.10	16.76	00.00	Inf
00.60	06.84	00.10	16.76	00.00	Inf
00.50	07.43	00.00	16.76	00.00	Inf
00.40	08.38	00.00	Inf	00.00	Inf
00.40	08.38	00.00	Inf	00.00	Inf
00.30	09.68	00.00	Inf	00.00	Inf

If $K_i = \sqrt{\lambda_i/\lambda_i} > 10$, then there are collinearity and if Eigen Value is near to zero it is perfect collinearity.

9.2 Stocks Control the Stock Market Movement:

Sometimes stock market move without any apparent news. We have to identify the most stocks reflect the movement of the stock market. By controlling the sum weights of these stocks in the portfolio can be avoided the risk of stock market move. So, we aim to define stocks control the stock market Move. Principle Components Regression is one of the ways to deal with collinearity:

1. Get the standardize values for all variables (dependent and independent)
2. Regress the Market Index on the standardize Assets Prices.

TABLE 4
THE REGRESSION PARAMETERS(PCR)

zStocks	B	zStocks	B	zStocks	B
zst01	0.004	zst16	0.149	zst31	0.221
zst02	0.048	zst17	0.070	zst32	-0.07
zst03	-0.04	zst18	0.228	zst33	0.024
zst04	-0.01	zst19	0.031	zst34	-0.04
zst05	0.034	zst20	-0.09	zst35	-0.02
zst06	-0.02	zst21	0.052	zst36	-0.03
zst07	0.069	zst22	-0.02	zst37	0.092
zst08	0.007	zst23	0.033	zst38	0.012
zst09	0.018	zst24	-0.06	zst39	0.009
zst10	-0.05	zst25	-0.01	zst40	-0.09
zst11	-0.05	zst26	-0.01	zst41	-0.03
zst12	-0.02	zst27	0.043	zst42	0.113
zst13	0.002	zst28	-0.09	zst43	-0.06
zst14	0.309	zst29	0.062	zst44	-0.06
zst15	0.049	zst30	0.121	zst45	0.089

3. Get the Principal Components Analysis for the independent standardize variables that to get Eigen Vectors and Eigen Values.
4. Even though Asset's is a set of highly collinear variables, it could be transforme to W's which is completely orthogonal which leads to complete independence. To find W's , we calculate the Principle Components
$$W_i = Z\,(Stock\,i) * V_i \qquad (14)$$
Where *Z(Stock i)* is standardize values of *(Stock i)* and V_i is the Eigen Vectors.
5. Get regress zMarket Index on W_i.

TABLE 5
THE REGRESSION PARAMETERS(PCR)

Stocks	B	Stocks	B	Stocks	B
W01	-0.185	W16	-0.148	W31	0.0872
W02	0.0285	W17	0.1292	W32	0.0675
W03	-0.063	W18	0.0090	W33	0.0927
W04	0.0589	W20	-0.040	W34	-0.062
W05	-0.042	W20	-0.009	W35	0.2098
W06	-0.048	W21	-0.072	W36	-0.109
W07	0.0171	W22	757.4e-6	W37	0.0329
W08	-0.025	W23	-0.065	W38	0.0289
W09	-0.044	W24	0.0748	W39	-0.016
W10	-0.026	W25	-0.144	W40	-0.029
W11	0.0678	W26	-0.054	W41	0.0183
W12	-0.089	W27	-0.003	W42	0.1469
W13	0.0286	W28	-0.056	W43	0.2003
W14	-0.003	W29	-0.041	W44	-0.089
W15	-0.077	W30	0.0426	W45	-0.185

If we regress zMarket Index on the significant W's only we will get the same W's coefficient because W's are orthogonal.

6. The estimated parameters of Principal Components Regression:
$$\hat{\beta}_{PCR} = V * \hat{\alpha} \qquad (15)$$
where V is the Eigen Vector of the zAsset i and $\hat{\alpha}$ is the significant coefficients of the regression *zMarket* Index on $W_1, \dots, W_{45}$
7. Test significant of $\beta's$ by calculation of:
$$V(\hat{\beta}) = \sigma^2\,(V\Lambda V^T)^{-1} \qquad (16)$$
To find W's the Principle Components are calculat-

ISSN 2229-5518

ed.

TABLE 6
THE EXPECTED RETURN OF THE STOCKS

Stocks	E(R)	Stocks	E(R)	Stocks	E(R)
Stock1	0.0411	Stock16	0.0539	Stock31	0.0852
Stock2	0.0340	Stock17	0.0200	Stock32	0.0711
Stock3	0.0433	Stock18	0.0800	Stock33	0.0839
Stock4	0.2000	Stock19	0.0649	Stock34	0.0400
Stock5	0.0400	Stock20	0.0150	Stock35	0.0340
Stock6	0.0320	Stock21	0.0600	Stock36	0.0533
Stock7	0.0260	Stock22	0.1029	Stock37	0.0460
Stock8	0.0360	Stock23	0.0502	Stock38	0.0411
Stock9	0.0463	Stock24	0.0569	Stock39	0.0600
Stock10	0.0232	Stock25	0.0321	Stock40	0.0629
Stock11	0.0849	Stock26	0.0371	Stock41	0.0529
Stock12	0.0310	Stock27	0.0430	Stock42	0.0809
Stock13	0.0800	Stock28	0.0759	Stock43	0.1400
Stock14	0.0364	Stock29	0.0652	Stock44	0.0460
Stock15	0.0839	Stock30	0.0473	Stock45	0.0749

9.3 Evaluation Performance of the Stocks:

The process of evaluating the stocks is important because it is the main step for obtaining optimal portfolio and the managing of it. One of the most important indicators of measuring the performance of mutual funds or prediction systems is the return of the system. The one period simple net return R_t of a financial time series is calculated with the following formula:

$$R_t = (P_t - P_{t-1})/P_{t-1} \qquad (17)$$

where R_t is the stock rate and P_t is the stock price. For a *n* period financial time series, the return at the end of the period is a cumulative return that is composed all the periods

$$E(R_i) = \sum_n P_n R_n \qquad (18)$$

where $E(R_i)$ is the expected return of the asset *i*; and *P* is the probability of occur the return *R*.

TABLE 7
THE RISK OF THE STOCKS

Stocks	E(V)	Stocks	E(V)	Stocks	E(V)
Stock1	0.1051	Stock16	0.2381	Stock31	0.2431
Stock2	0.0955	Stock17	0.1093	Stock32	0.2333
Stock3	0.1854	Stock18	0.2059	Stock33	0.2210
Stock4	0.1404	Stock19	0.2121	Stock34	0.1548
Stock5	0.1298	Stock20	0.0842	Stock35	0.1465
Stock6	0.0946	Stock21	0.1843	Stock36	0.0532
Stock7	0.1391	Stock22	0.4810	Stock37	0.2058
Stock8	0.1445	Stock23	0.0502	Stock38	0.1701
Stock9	0.1419	Stock24	0.2157	Stock39	0.1882
Stock10	0.0929	Stock25	0.1093	Stock40	0.2041
Stock11	0.2747	Stock26	0.1612	Stock41	0.2198
Stock12	0.1150	Stock27	0.1394	Stock42	0.2512
Stock13	0.2500	Stock28	0.2224	Stock43	0.3990
Stock14	0.1070	Stock29	0.1966	Stock44	0.2363
Stock15	0.3059	Stock30	0.1491	Stock45	0.2541

Another measure is the rate of the risk. In the case of using the rate of the risk measure, it is selected the stocks with smallest rate of the risk. The risk of each asset i.

$$\sigma_i = \sqrt{\sum (R_n - E(R_i))^2 P_n} \qquad (19)$$

where P_n is the probability of the possible rate of return R_i.

TABLE 8
THE SUGGESTED RATIO OF THE STOCKS

Stocks	Ratio	Stocks	Ratio	Stocks	Ratio
Stock1	0.39	Stock16	0.23	Stock31	0.35
Stock2	0.36	Stock17	0.18	Stock32	0.31
Stock3	0.32	Stock18	0.39	Stock33	0.38
Stock4	1.42	Stock19	0.31	Stock34	0.26
Stock5	0.31	Stock20	0.18	Stock35	0.23
Stock6	0.34	Stock21	0.33	Stock36	0.26
Stock7	0.19	Stock22	0.21	Stock37	0.22
Stock8	0.25	Stock23	0.30	Stock38	0.24
Stock9	0.33	Stock24	0.27	Stock39	0.32
Stock10	0.25	Stock25	0.29	Stock40	0.31
Stock11	0.31	Stock26	0.23	Stock41	0.24
Stock12	0.27	Stock27	0.31	Stock42	0.32
Stock13	0.32	Stock28	0.34	Stock43	0.35
Stock14	0.34	Stock29	0.33	Stock44	0.19
Stock15	0.27	Stock30	0.32	Stock45	0.29

In the case of using the rate of the return measure, the stocks with largest rate of the return are selected. In case of using the rate of the risk measure, the stocks with smallest rate of the risk are selected. The risk of each asset i. There are two main major measures for the stocks' performance; the first measure is the selection of the stocks with the highest expected value of the rate of the return but perhaps these stocks have the highest rate of the risk. So, it is not a successful way to select the stocks; the second measure is the selection of the stocks with the lowest expected rate of risk but perhaps these stocks have the lowest rate of the return. Perhaps the decrease in the rate of the risk is not equivalent to the decrease in the rate of the return and perhaps the increase in the rate of the return is not equivalent the increase in the rate of the risk. So, the below ratio is suggested.

$$M_i = E(R_i) / \sigma_i \qquad (19)$$

where M_i is the suggested measure.

Some stocks have the largest rate of return but at the same time have the largest rate of risk. Additionally, there are some stocks have the smallest rate of risk but have the smallest rate of return.

TABLE 9
THE MULTI-COLLINEARITY GROUPS

Groups	Stocks
Group1	Stock1, Stock21, Stock24
Group2	Stock11, Stock29, Stock31
Group3	Stock18, Stock19, Stock27
Group4	Stock20, Stock23, Stock30
Group5	Stock16, Stock35, Stock36, Stock40

10 Conclusions

In this paper the approach to evaluation and selection of the stocks is introduced. The suggested approach solves two main problems; the first problem is the outlier problem and the second problem is multi-collinearity problem. Additionally, stocks prices time series analysis defines the characteristics of the stocks and determines the multi-collinearity stocks groups. Determination of the multi-collinearity groups enables the investors to construct various construct portfolios adequate with his opinion. This study sheds light on the importance of using Robust Regression for financial time series analysis. The most important result of our study is understanding the case of outliers on the time series and introducing solution to it as indicated. Additionally, the groups of the multi-collinearity stocks are introduced. From all the previous, this paper introduces approach to evaluation and selection of the stocks as the first step to construct portfolio. This suggested approach enables us to do the following: **the suggested approach presents the following:**

1. Avoid the drop portfolio risk by solving the multi-collinearity problem among the stock into the portfolio; determining the multi-collinearity groups by using the Principle Component Regression and constraining the sum allocated of the investment in each group to avoid increasing the allocate in each stock of each group.
2. Avoid the market movement risk by determining the market's stocks and conceder it as group; Increasing the sum of allocated of the investment in this group in the case of active state and decreasing the sum of allocated of the investment in this group in the case of week state.
3. Suggesting the ratio of return/risk as measure of the stock performance and determining the best stocks which have outstanding performance.

From the previous analysis; the constrain selection stocks in the optimal portfolio as following:

1. **Constrain of the multi-collinearity groups:**

$$W_G \le P \qquad (20)$$

where W_G is the allocation of the multi-collinearity stocks group *G* and p is the allocation of the stocks into the portfolio.

$$W_1 + W_{21} + W_{24} \le P_1 \qquad (21)$$
$$W_{11} + W_{29} + W_{31} \le P_2 \qquad (22)$$
$$W_{18} + W_{19} + W_{27} \le P_3 \qquad (23)$$
$$W_{20} + W_{23} + W_{30} \le P_4 \qquad (24)$$
$$W_{16} + W_{35} + W_{36} + W_{40} \le P_5 \qquad (25)$$

2. **Constrain of the market'stocks:**

$$W_{SM} \le P \qquad (26)$$

where W_{SM} is the allocation of the market' stocks into the portfolio.

$$W_1 + W_2 + W_{14} + W_{16} + W_{18} + W_{27} + W_{30} + W_{31} + W_{39} \le P_7 \qquad (27)$$

3. **Constrain of the outstanding stocks into the portfolio.**

$$W_{OS} \le P_7 \qquad (28)$$

where W_{OS} is the allocation of the outstanding stocks into the portfolio.

References

[1] Chatterjee, S.Hadi, A.S. (1988). Sensitivity Analysis in Linear Regression, NewYork: Wiley.
[2] Chatterjee, S.Hadi, A.S. (2000). Regression Analysis by Examples, Third Edi-tion, NewYork: Wiley.
[3] Chi-Ming, L., Mitsuo, G. (2007). An E_ective Decision-Based Genetic Algorithm Approach to Multi-objective Portfolio Optimization Problem, Applied Mathematical Sciences, 5, vol (1), 201-210.
[4] Deng, X.T., Li, Z.F., Wang, S.Y. (2005). A Minimax Portfolio Selection Strategy with Equilibrium, European Journal of Operational Research, vol (166), 278-292.
[5] Lai, K.K., Yu, L., Wang, S.A., Zhou, C. (2006). A Double Stage Genetic Optimization Algorithm for Portfolio Selection, Kingetal.Eds.:ICONIP 2006, PARTIII, LNCS4234, Springer-Verlag Berlin Heidelberg 2006, 928-937.
[6] Lashean, A. (2001). Buiding Optimal Portfolio for Egyptian Mutual Funds ,Magazine of Financial and Strategic Decisions, Zagazeg University,2 ,vol(23), 207-238.
[7] Markowitz, H. (1952). Portfolio Selection, the Journal of Finance, 1, 77-91.
[8] Markowitz, H. (1959). Portfolio Selection Efficient Diversi_cation of Investments, John Wiley & Sons, London.
[9] Mitra, G., Ellison, F., Scoweroft, A. (2007). Qudratic Programming for Portfolio Planning, Journal of Asset Management, 8, 200-214.
[10] MATLAB, Version 7.5. (2007). http://www.mathworks.com/patents.
[11] Mona, N. Abdel Bary (2004). Use of Time Series Models and Neural Networks in the Prediction of Stocks Prices in the Egyptian Stock Market, The 39th Annual Conference of Statistics, Computer Science and Operations Research Second Part, Applied Statistics, Cairo University, Egypt, 1—24.
[12] Sharpe, W.F. (1992). Asset Allocation: Management Style and Performance Measurement, the Journal of Portfolio Management, 7-19.
[13] Sharpe, W.F. (2007). Expected Utility Asset Allocation, Financial Analysis Journal, 63, vol (5), 18-30.
[14] Zhang, W.G.Chen, W., Wang, Y.L. (2006). The Adaptive Genetic Algorithms for Portfolio Selection Problem, International Journal of Computer Science and Network Security, 1,vol(6),196-200.