Investigation of the Profitability of the Methods of Selecting for Predicting the Risk of Stock Price Fall in the Supply Chain of Companies Listed in Tehran Stock Exchange

Hassan Mohammadi[1], Alireza Zarei Soudani[2]
[1]*Department of Accounting, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran*
Alireza Zarei Soudani[2*].
[2]*Department of Accounting, Falavarjan Branch, Islamic Azad University, Falavarjan, Iran*

***Abstract*- The aim of this study was to investigate the profitability of variable reduction methods for predicting the risk of the stock price drop of companies listed in Tehran Stock Exchange. To achieve this, the literature review was conducted and 24 primary variables were selected which were most frequently used in the literature and the required data for measuring them was available. The optimum variables were selected or extracted among the primary variables using variable selection methods (the correlation-based method and the relief method) and variable extraction methods (the factor analysis and the principal component analysis). Then, the risk of stock price fall for 101 companies listed in Tehran Stock Exchange was predicted for 2001-2015 using linear regression. In order to evaluate the performanceof the variable reduction methods, the evaluation criteria resulting from prediction using variables selected or extracted by these methods were compared with criteria resulting from prediction using all variables. Findings of the research indicated the profitability of variable reduction methods and significant differences between profitability levels of different methods. The results obtained from the investigation of the performance of different methods of prediction and variable reduction in the industry group indicated the effect of the type of industry on the prediction performance. Furthermore, the results of the prediction of returns during 2001-2015 showed that the performance of prediction was higher in some years and lower in some other years compared to the results of the collective investigation of the supply chain of companies in this period.**

Keywords: *risk, stock price fall, variable reduction methods, supply chain.*

1. Introduction

Based on the significance of considering the risk of stock price fall in decision making by investors and creditors as well as the significant role of selection and extraction of optimum predictor variables in predictions, the present paper examined the performance of different nonlinear methods and compared them in respect to predicting the risk of the stock price fall. To achieve this, the profitability of the factor analysis method, principal component analysis method, and the correlation-based method for selecting and extracting the optimum predictor variables were investigated and compared. So far, no study has been conducted on predicting the risk of the stock price fall using these methods in Tehran Stock Exchange. Moreover, in studies conducted on the prediction of return, the main aim and emphasis was on presenting suitable and precise models for prediction and variable reduction and their desirable methods received less attention (selection and extraction of variables or factors). In contrast, in most domestic and foreign studies in this regard, the predictor variable reduction stage has been ignored and the predictor variables have been selected without any standard but by merely relying on the previous studies. This may lead to the selection of non-optimal predictor variables, and in some cases, improper predictor variables. [1] showed that the selection of predictor variables and the extraction of predictor variables (factors) and their methods have more influence on the average prediction precision compared with the selection of a predictor model.

International Journal of Supply Chain Management
IJSCM, ISSN: 2050-7399 (Online), 2051-3771 (Print)

2. Theoretical Foundations and Literature Review

2.1. Variable Reduction (Selection and Extraction of Predictor Variables)

Most researchers are more interested in achieving the best prediction of the dependent variable by some predictors while it is necessary to concentrate more on hypothesis test or evaluating the relative significance of the predictor variables. In such conditions, the researcher spends most of his/her time on achieving the highest multi-variable root correlation. Since in the behavior sciences, most variables have mutual correlations, it is often possible to select a smaller set among the whole set of variables and achieve the same R^2 which results from the whole set of variables. If it is supposed to select some variables among the available variables, it is usually considered that the selected variables should be of the least number and also consider the same variance value considered by the whole set of variables. However, the practical considerations (including the relative expenses of data collecting and simplicity of management) often interfere in the selection process. In such conditions, the selected variables might be more than the minimum number necessary for considering a variance which is almost equivalent to the variance considered in the whole variables set. In this case, the researcher may select a higher number of variables (e.g. five variables) which considers the same R^2, instead of selecting the least number of variables (e.g. three variables). The variable reduction stage (selection and extraction of the predictor variables) is often conducted before learning the predictor models. However, in most studies in the field of accounting, this stage is often ignored and the predictor variables are not selected systematically. This may lead to the selection of non-optimum, and in some cases, improper predictor variables. In these studies, the predictor variables were selected without considering any standard but by merely relying on previous studies.

The selection and extraction of proper variables to achieve the best results in prediction are considered as the challenging topics in the last two decades. In the theoretical viewpoint, learning based on the number of predictor variables may lead to more precise predictions. However, the empirical evidence have shown that this is not always the case because all variables are important for diagnosis and prediction, or some of them are generally irrelevant in prediction. Since many factors (including data quality) are effective in the success of a learning algorithm, if data include repetitive and irrelevant information, or include uncertain or parasitic information, it would be hard to harvest any knowledge form that data. Moreover, reducing the number of irrelevant or excess predictor variables may decrease the performance time of the learning algorithm and also lead to a more general concept. Other potential advantages of selection and extraction of the predictor variables include facilitating perception and incarnation of data, reducing the requirements of measurement, data storage, and the course of dimensionality as well as improving the performance of prediction and providing a better insight into the fundamental concept from the classification of the real world.

There are two important aspects in different methods of dimension (variable) reduction:

- Extraction of the predictor variables (factors): extraction of predictor variables, or in other words, changing the predictor variables is a process which results in K new variables which are the result of combining N primary predictor variables. The most well-known algorithms for extracting the predictor variables include the principal component analysis, factor analysis, and auditing analysis. The principal component analysis and factor analysis are considered as the most important methods of extracting the predictor variables, which is also used in the present study.
- Selection of predictor variables: contrary to the predictor variables extraction algorithms, the predictor variable selection algorithms contribute to the selection of the best K variables among N primary variables and other less important variables are omitted.

It is to be mentioned that in the variable selection, the main variables are selected without change, but in variable extraction, the variables are used in their changed form. The correlation-based method and the relief method are considered as the most important methods for selecting the variables in predicting the continuous variables, which is also employed in the present study. The reasons for selecting these methods are:

1. The findings of previous studies [2] showed better performance of the above-mentioned methods compared with other variable selection methods.
2. The above-mentioned methods are among the variable selection methods in the prediction field in which the primary variables with no change would be selected; however, in the variable extraction methods, the variables are used in their changed form. Moreover, the mentioned methods are used in the prediction issues (with dependent continuous variable) while some of the variable selection methods are just used in classification issues (with nominal dependent variables such as bankruptcy). For example, despite the better performance of the fuzzy rough set [3], the mentioned method is just used in classification issues.

2.2. Factor Analysis

Factor analysis is a general name for some multi-variable statistical methods whose main aim is to brief the data. This method concentrates on investigating the inner correlation of many variables and finally classifies and explains them in the form of limited general factors. In the factor analysis method, all variables are considered simultaneously and each variable is considered as a dependent variable. The factor analysis is one of the multi-variable methods in which the dependent or independent variables have the same weight because this method is considered as a co-dependent technique and all variables are considered as dependent on each other, and it is tried to brief high number of variables in a few factors. To conduct the factor analysis, four fundamental steps are to be taken:

A) Creating a correlation matrix from all variables used in analysis and estimating the partnership
B) Factor extraction
C) Selection and rotation of factors to simplify the factor structure
D) Interpreting the results

2.3. Methods

The relief method of variable selection is one of the predictor variable selection methods which is based on the criterion of distance. In the relief method, the weight which shows the relationship between each variable and category is determined by the Euclidian distance between the samples, and the weight of each variable shows the ability to separate the categories by the given predictor variable. In this method, if a variable has the same value per samples in a class and has different values per other samples of the class, then, it gains higher weight. Relief method selects a sample among the training data randomly and obtains the Euclidian distance of that sample to the nearest sample in the same class and the nearest sample in a different class, and then uses these distances for updating the weight of each variable. Finally, it selects the algorithm of the variables whose weight is higher than a threshold predefined by the user. The Relief method, which was initially presented by [4], can only be used in classification issues with two groups (e.g. bankrupt vs. non-bankrupt). [5] extended the relief method for being used in continuous output data. In the present study, the extended relief method (RRelieff) was used for selecting the predictor variables.

2.4. The Correlation-based Method

Generally, a variable is appropriate if it is related to the dependent variable (class in the classification), but it is not excessive relative to other related predictor variables. If the correlation between two variables is considered as the criterion of being appropriate, then, the mentioned definition would be changed to this one: A variable is proper if it has a high correlation with the dependent variable (class) and low correlation with other predictor variables. In other words, if the correlation between a predictor variable and the dependent variable (class) is high enough to be related to predicting the dependent variable (class) and the correlation between it and other related predictor variables does not reach a definite level so that the mentioned variable cannot be predicted by other related variables. Then, the given variable is considered as a proper variable for prediction (classification). In this situation, the main issue would be selection of the variable, searching a proper criterion for correlation between variables and logical manner for selecting proper variables based on this criterion. The correlation based method calculates the correlations between predictor variables and also the predictor variables and dependent variable and then searches the variables subset spaces. The subsets found by searching which has the highest profitability would be used for reducing the dimensions of the primary training data and test data.

2.5. The Principal Component Analysis

The main idea of the principal component analysis is to reduce the dimensions of a set of data which has a high number of correlated variables while keeping the variability present in the dataset as much as possible. This reduction is done through changing to a set of new variables (the principal components) which are not correlated and they are to be ordered in a way that a few variables which remain in the beginning keep the main part of the variability present in the whole primary main variables. Calculating the special values and special vectors from the principal component, a linear combination of the principal variables can be found which causes the highest variance. The first principal component explains the variability in the dataset as much as possible and each of the following components explains the remaining variability as much as possible. Therefore, defining and calculating the principal components are straightforward. The factor analysis is a generalization of the principal component analysis. The difference between the factor analysis and the principal component analysis is that the latter considers the total variance for all general and unique variances (special plus error) in the dataset while the former just considers the general variance.

2.6. The Research Hypotheses

Based on the questions, theoretical basics and research background, the following hypotheses have been presented:

1. The selected or extracted optimum predictor variables predict the risk of stock price fall significantly better than all primary predictor variables.

1-1. The optimum predictor variables extracted by the factor analysis method predict the risk of the stock price fall significantly better than all primary predictor variables.

1-2. The optimum predictor variables extracted by the principal component analysis predict the risk of the stock price fall significantly better than all primary predictor variables.

1-3. The optimum predictor variables selected by the correlation-based method predict the risk of the stock price fall significantly better than all primary predictor variables.

1-4. The optimum predictor variables selected by the relief method predict the risk of the stock price fall significantly better than all primary predictor variables.

2. There is a significant difference between the profitability of the methods of selecting and extracting the variable in predicting the risk of the stock price fall.

2-1.There is a significant difference between the profitability of the methods of factor analysis and principal component analysis in predicting the risk of the stock price fall.

2-2.There is a significant difference between the profitability of the factor analysis method and the correlation-based method in predicting the risk of the stock price fall.

2-3.There is a significant difference between the profitability of the factor analysis method and relief method in predicting the risk of the stock price fall.

2-4.There is a significant difference between the profitability of the principal component analysis method and the correlation-based method in predicting the risk of the stock price fall.

2-5.There is a significant difference between the profitability of the principal component analysis method and the relief method in predicting the risk of the stock price fall.

2-6.There is a significant difference between the profitability of the correlation-based method and the relief method in predicting the risk of the stock price fall.

3. Methods

This is an applied research. It used a quasi-experimental design based on the ex-post facto approach (by previous information). The ex-post facto approach is used when the researcher investigates an event after its occurrence. Moreover, the manipulation of the independent variables is not possible [6].

3.1. Data Collection Method

In the present study, the library and field methods were used for collecting the required data. The theoretical basis of the research was collected from texts, journals, and websites. The financial data were collected through referring to the website of Tehran stock exchange and financial statement of the supply chain of companies and also TadbirPardaz and Rahavard Novin software. In the first stage, a literature review was conducted (including 250 Farsi and English papers) and approximately 150 primary predictor variables were identified. The primary papers were mainly selected from reliable papers and the thesis available on websites such as Science Direct, Springer, JStore and Proquest [8]. Among the identified variables, 24 variables were selected which were used in the literature for explaining or predicting the risk of the stock price fall, and the required data for their analysis was available through the website of stock exchange organization and also software such as TadbirPardaz and Rahavard Novin. Afterward, the optimum variables were selected or extracted among 52 mentioned variables using the methods of selecting the predictor variables (the correlation-based method and the relief method) in Weka software. [Table 1] shows seven optimum variables selected by the correlation-based method. The relief selection method rates all predictor variables. In the present study, the seven top (better) variables rated by the mentioned method were used in order to better comparison ability with correlation-based methods. The mentioned variables are depicted in [Table 1] based on their rating. Moreover, the factor analysis method and the principal component analysis method led to the extraction of 20 and 12 factors, respectively.

3.2. Independent (Predictor) Variables

The first twenty predictor variables used in this study are presented in [Table 1] and their selection method was explained in section 3.5.

Table 1. Definition of the independent variables used in this study

Symbol	The variable under study
InsOwn	Institutional ownership
CentOwn	Central ownership
ManOwn	Management ownership
OPAQUE	Lack of clarity of financial information
BrdIndep	Board independence
BrdDobl	Board double
OverconfidentCEO	Over confidant CEO
FRQ	Financial Report Quality
BIG	Auditing institute size
SPECIALIST	Specialty of auditor in the industry
TENURE	Tenure period of auditor
DISXT	Disturbing the real activities through unusual optional expenses
PROD	Disturbing the real activities through unusual production expenses
CFO	Disturbing the real activities through unusual operational cash flow
DA	Interest management based on deliberate items
HHI	Herffndal Hirschman Index
QN	Tobin`s Q ratio
LI	Learner indicator
TI	Adjusted learner indicator
CC_SCORE	Conditional conservatism
UC_SCORE	Unconditional conservatism
PERSIST	Persistence of the interest
PREDICTABILITY	Interest predictability
VOLATILE	Interest smoothing
DUALITY	Investors duality
STD	Standard deviation of stock monthly return
RET	Mean stock monthly return
Size	Size of the company
ROE	Profitability Index
MTB	The ratio of market value to office value of stakeholders
LEV	Financial leverage

3.3. The Dependent Variable

3.3.1. Measuring the Stock Price Fall

In order to measure the stock price fall, the skewness coefficient model [7] and the model introduced by [9] were used and equation (1) was used for its calculation.

$$CRASH_{it} = \text{NCSKEW}_{it} = -(n(n-1)^{\frac{3}{2}}\sum W_{J,\theta}{}^{3})/((n-1)(n-2)(\sum W_{J,\theta}{}^{2})^{\frac{3}{2}}) \quad (1)$$

NCSKEW_{it} is the negative coefficient of the skewness of special monthly return of company i in year t; $W_{J,\theta}$ is the special monthly return of company j in month θ during the financial year which is estimated by model (2) and n is the number of observed months of return during the financial year. In the above-mentioned model, when the negative skewness coefficient is high, it means that the supply chain of company is at a higher risk of stock price fall.

3.3.2. Population and Sample of the Study

The statistical population of the present study included supply chain of companies listed in Tehran stock exchange from 2001 to 2015. The purposeful sampling method (systematic elimination) was used for sampling. To achieve this, all companies of the population having the following conditions were selected as the sample and the remaining ones were eliminated: In order to have comparable information, 29th of Esfand (20th March) was selected as the end of the financial year. In order to have homogenous information, the manufacturing companies were selected. The transactions of their stock have not been stopped in Tehran stock exchange for more than three months in the study period.

3.3.3. Data Analysis Method and Hypothesis Testing

After determining the variables and optimum factors using either selection method or extraction method, the linear regression method was used for predicting the stock return. It is to be noted that in this study, the data from the previous year of the supply chain of companies were used for predicting the risk of the stock price fall [9]. In order to evaluate the performance of different methods of prediction, the evaluation criteria (including mean absolute percent error, root mean square error and determination coefficient) related to the prediction of the stock price fall were used in each of the prediction methods. The mentioned criteria are considered as the most prevalent criteria for evaluating the performance in prediction issues, which are depicted in [Table 2].

Table 2. The criteria used for the prediction performance assessment

Criterion	Measuring Method
(RMSE) Root mean square error	$\sqrt{\frac{\sum_{P=1}^{P}(dp - zp)^2}{P}}$
Determination coefficient (R^2)	$1 - \frac{\sum_{P=1}^{P}(dp - zp)^2}{\sum_{P=1}^{P}(dp - \bar{d}p)^2}$
Mean absolute percent error (MAPE)	$\frac{100}{P} \times \sum_{P=1}^{P}\left\lvert\frac{dp - zp}{dp}\right\rvert$
Z_p : Predicted value d_p: Real Value $\bar{d}$: Mean Value	

Source: Smith and Gupta (2002: 9) and Azar and Karimi (2009: 8)

When the determination coefficient is higher and the two other criteria are lower, the prediction has a better performance. Although there are some other common criteria for evaluating the performance in this filed, they are not presented here because the prediction performance can be easily calculated using the above-mentioned criteria. For example, the root mean square error (EMSE), the mean square error (MSE), the normalized mean square error (NMSE), and the determination coefficient are complementary to each other. Moreover, in order to evaluate the performance of different methods for selecting and extracting the optimum variable, the evaluation criteria (mean absolute percent error, root mean square error, and determination error) resulting from each of the methods of selecting the variable were compared with each other and also with evaluation criteria resulting from ignoring the phase of selecting the predictor variables in each of the linear and nonlinear methods. Ignoring the predictor variable selection stage means prediction using all predictor variables (before reducing the number of variables). Reduction of prediction variables and prediction of stock return (except for factor analysis which is done using SPSS software) were done using different linear and nonlinear methods in Weka software (version 3-7). ANOVA test (and the non-parametric Kruskal-Wallis test if the parametric assumptions are not confirmed) and paired t-test (and the nonparametric Wilcoxon test if the parametric assumptions are not confirmed) were used to test the principal and secondary hypothesis of the research, respectively, based on 100 precisions resulting from executing the ten-part mutual validity with ten repeats in each prediction method in SPSS software (version 21).

4. Findings of the Study

[Table 3] shows the mean evaluation criteria (including mean absolute percent error, root mean square error and determination coefficient) related to predicting the risk of the stock price fall based on the linear regression method in five cases (using 24 predictor variables with selective variables in the correlation-based method (Corr) and the relief method (R), and also with the extracted factors in the principal component analysis (PCA) and factor analysis (FA)). The mean criteria result from ten repetitions of the ten-part mutual validity (the ten-part mutual validity with 10 repetitions) which lead to the creation of 100 precision for each prediction method.

Table 3. Mean performance of different methods of prediction using five methods of variable reduction

Prediction method / Performance criterion	FA	PCA	R	Corr	ALL
RMSE	/895 86	/524 77	/617 75	/895 84	89/247
MAPE	0.476	0.374	0.342	0.467	/512
R^2	0.102	0.142	0.167	0.114	0.041

In order to test the first main hypothesis, based on [Table 4], the mean determination coefficient of prediction related to using the variables and factors selected or extracted by four methods of selection and extraction of variables (the correlation based method, the relief method, the principal component analysis method, and the factor analysis method), and using all variables in the linear regression method through using variance analysis method (and nonparametric Kruskal Wallis if the parametric assumptions are not confirmed) were compared. In order to test the secondary hypothesis related to the first main hypothesis, the performance criteria of each prediction method in the case of using 24 predictor variables were compared with the case of using variables selected through the correlation-based method, the relief method, factors extracted through the principal component analysis method, and the factor analysis method (two by two) using the paired t-test (and nonparametric Wilcoxon if the parametric assumptions are not confirmed). If the performance of each prediction method (based on [Table 4]) in the case of using variables selected through variable selection methods and factors extracted through variable (factor) extraction method are better than using the 24 variables, and their difference is statistically significant (based on [Table 4]), it can be inferred that the methods of variable selection and factor extraction have positive and significant influence on the performance of prediction method. Because of the normality of determination coefficient distribution in investigating and comparing each couple of determination coefficient, the paired t-test was used. The results of the paired t-test related to the comparison of the mean determination coefficient of each prediction method in the case of using 24 variables and using the selected or extracted variables are shown in [Table 4]. The reason for using determination coefficient for hypothesis testing is that the mentioned criterion is the most well-known and most frequently used criterion for evaluating the prediction models. It is to be noted that in this regard, the precision resulting from ten repetitions of the ten-part mutual validity (the ten-part mutual validity with 10 repetitions) were used which led to the creation of 100 precisions for each prediction method.

Table 4. Results of the t-test and the related probability level using linear regression

	FA	PCA	R	Corr	All
All					
Corr					3.163 (0.000)
R				3.159 (0.000)	3.879 (0.000)
PCA			3.128 (0.000)	3.167 (0.000)	3.791 (0.000)
FA		3.148 (0.000)	3.627 (0.000)	1.364 (0.176)	3.528 (0.000)

5. Discussion and Conclusion

The results of the first main hypothesis showed that the selected or extracted optimum prediction variables were significantly better in predicting the risk of the stock price fall than all primary prediction variables. Therefore, the first main hypothesis (the profitability of variable reduction methods in predicting the stock price fall) was accepted. The results of the secondary hypothesis of the first main hypothesis (the significant superiority of the performance of predicting the stock price fall using optimum prediction variables selected or extracted by each of the variable reduction methods compared with using all predictor variables) are as follows:

- In the prediction method, the performance of predicting the risk of the stock price fall using factors extracted by factor analysis was significantly better than the prediction by all variables. Therefore, the secondary hypothesis 3-1, which emphasized on the significant superiority of performance of predicting the risk of stock price fall using the factors extracted by factor analysis method compared with using all prediction variables, was accepted.
- In the prediction method, the performance of predicting the risk of stock price fall using factors extracted by the principal component analysis method was significantly better than prediction using all variables. Therefore, the secondary hypothesis 3-2, which emphasized on the significant superiority of the performance of predicting the risk of the stock price fall using components extracted by the principal component analysis method compared with using all predictor variables, was accepted.
- In the prediction method, the performance of predicting the risk of the stock price fall using variables selected by the correlation based method was significantly better than prediction using all variables. Therefore, the secondary hypothesis 3-3, which emphasized on the significant superiority of the performance of predicting the risk of stock price fall using variables selected by correlation based method compared with using all predictor variables, was accepted.
- In the prediction method, the performance of predicting the risk of stock price fall using variables selected by the relief method was significantly better than prediction using all variables. Therefore, the secondary hypothesis 3-4, which emphasized on the significant superiority of the performance of predicting the risk of stock price fall using variables selected by the relief method compared with using all predictor variables, was accepted.

The results of testing the second main hypothesis showed a significant difference between the profitability levels of different methods of reducing the variables number in predicting the risk of the stock price fall. Therefore, the second main hypothesis of the study (there is a significant difference between the profitability levels of different methods of variable reduction) was accepted. The results of testing the secondary hypothesis of the second main hypothesis (there is a significant difference between profitability levels of each couple of variable reduction methods in predicting the risk of stock price all) are as follows:

- In the prediction method under study, factors extracted by the principal component analysis method predicted the risk of stock price fall significantly better than the factors extracted by the factor analysis method. Therefore, hypothesis 4-1, which says that there is a significant difference between the profitability levels of the factor analysis method and principal component method, was accepted.
- Regarding the prediction through linear regression, a significant difference was found between using factors extracted by the factor analysis method and variables selected by the correlation based method.
- Regarding the prediction method under study, factors extracted by the factor analysis method predicted the risk of stock price fall significantly better than variables selected by the relief method. Therefore, hypothesis 4-3, which says there is a significant difference between the profitability levels of the factor analysis method and the relief method, was accepted.
- In the prediction method under study, factors extracted by the principal component analysis method predicted the risk of the stock price fall significantly better than variables selected by the correlation based method. Therefore, hypothesis 4-4, which says there is a significant difference between the profitability levels of the principal component analysis method and the correlation based method, was accepted.
- Regarding the prediction through linear regression, using the variables selected by the relief method predicted the risk of stock price fall significantly better than the factors extracted by the principal component analysis method.
- In the prediction method under study, variables selected by the relief method predicted the risk of stock price fall significantly better than variables selected by the correlation based method. Therefore, hypothesis 4-6, which says there is a significant difference between the profitability level of the relief method and the correlation based method, was accepted.

In general, results of testing the first and fourth main hypothesis (and the related secondary hypotheses) showed better performance of variables selected or extracted by variable reduction methods, compared with using all primary predictor variables, in predicting the risk of stock price fall, and also the significant difference between the profitability levels of different methods of variable reduction. In other words, when the predictor variable

reduction methods were used, the mean determination coefficient increased and the mean absolute percent error and root mean square error decreased. The reason for the superiority of the performance evaluation criteria in the case of conducting the variable reduction phase, compared with ignoring this phase, is the curse of dimensionality. It seems that adding more variables may increase the parasites (noise) and finally the errors, and adding the variables may contribute to improving the prediction up to a threshold, and adding the variables more than that threshold may lead to the curse of dimensionality. Moreover, the findings of the present study showed the superiority of the relief method for selecting the variables compared over other methods of variable reduction. After the relief method, the principal component analysis method had a better performance compared with the factor analysis method and the correlation-based method in reducing the predictor variables. It is to be noted that the factor analysis method, which is commonly used in the financial and accounting projects, had weaker performance compared to the other three methods introduced in the study. However, using this method was proved to be better than ignoring the phase of variable reduction (and using all primary predictor variables).

Moreover, the reduction percentage of the predictor variables in case of using different methods of variable selection or extracting the predictor factors is an important criterion in evaluating a predictor variable reduction method. In general, the logical low number of predictor variables is one of the most important criteria in evaluating the model quality, and a model is considered as a valuable and important one when explains a high extent of variations just through a low number of variables. Totally, the variable selection methods (the correlation-based model and the relief model) had preference over variable extraction methods because of maintaining the primary variables without any change (and not in the altered form). Furthermore, based on the findings of the present study, the factor analysis method, which was frequently used in the previous studies, was found to be the weakest method for reducing the variables.

The results of the study, which showed the profitability of variable reduction methods and a significant difference between profitability levels of different variable reduction methods, are in line with the findings of [9].

6. Suggestions of the Study

Based on the findings, the following suggestions are presented:

1. Because of the positive effect of using the predictor variable reduction methods, compared with ignoring it, on the stock price fall prediction operation, investors and other users are recommended to conduct the variable reduction phase in predicting the risk of the stock price fall, and not just select the predictor variables based on the previous studies. According to the better performance of other methods (especially the relief method) compared with the factor analysis, it is recommended to use this method for selecting the optimum variables.
2. With regard to the optimum variables identified in this research for predicting the risk of the stock price fall [Table 3], the stock exchange organization is recommended to obligate the companies to present and disclose these variables (at least every three months).
3. According to the theoretical foundations of financial reporting, the relevance (and predicting value) of the accounting information is one of the most important qualitative features and relevance is interpreted regarding the prediction of the risk of the stock price fall. Hence, it is proposed that the Accounting Standards Committee specially take into consideration the selected optimum variables in this study which are very important in predicting the risk of the stock price fall so as to develop the accounting standards.

7. Limitations of the Study

1. Many political, ecological and social conditions of Iran (especially price inflation and not presenting the adjusted financial statements) might have influenced the findings which were out of the control of the researcher.
2. The lack of required and reliable data for calculating the variables regarding some companies or some years might have led to discarding them from the statistical sample, which might influence the ability to generalize the results to the statistical population. In the absence of these limitations, a higher number of companies could be studied and the results could be generalized to the whole population with higher confidence.

References

[1] Ahmed, A., & Duellman, S. *"Managerial overconfidence and accounting conservatism"*, Journal of Accounting Research, 51, 1–30, 2012.

[2] Andreou, P., Antoniou, C., Horton, J., & Louca, C. *"Corporate governance and firm-specific stock price crashes*", 2013. Available at SSRN: http://ssrn.com/abstract=2029719.

[3] Francis, B, Pandit, S, Li, L. *Abnormal real operations, real earnings management, and subsequent crashes in stock prices*, 2014.

[4] Habib, A. *"Managerial talent, investment efficiency and stock price crash risk: Working paper"*, Massey University, 2014c.

[5] Hamm, S.J.W., Li, E.X., & Ng, J. "*Management earnings guidance and stock price crash risk"*, Singapore Management University School of Accountancy Research Paper, No. 10, 2014. Available at SSRN: http://ssrn.com/abstract=2055008**.**

[6] Hu, J., Li, A.Y., & Zhang, F. *"Does accounting conservatism improve the corporate information environment?"* Journal of International Accounting, Auditing and Taxation, 23, 32–43, 2015.

[7] Robin. A., and Zhang. H. *"Do Industry-Specialist Auditors Influence Stock Price Crash Risk?"* The Accounting Review, Vol 83, No. 6: 1571-1603, 2016.

[8] Trkman, P., and McCormack, K. *"Supply chain risk in turbulent environments—A conceptual model for managing supply chain network risk,"* International Journal of Production Economics, Vol 119, No. 2, 247-258, 2009.

[9] Khan, O., and Burnes, B. *"Risk and supply chain management: creating a research agenda,"* The international journal of logistics management, Vol 18, No. 2, 197-216, 2007.