Advances in mathematical finance & applications, 4 (2), (2019), 43-64
DOI: 10.22034/amfa.2019.584793.1180

Published by IA University of Arak, Iran
Homepage: www.amfa.iau-arak.ac.ir

The Integration of Multi-Factor Model of Capital Asset Pricing and Penalty Function for Stock Return Evaluation

Ali Akbar Farzinfar[a], Hossein Jahangirnia[a,]*, Reza Gholami Jamkarani [a], Hasan Ghodrati Ghazaani[b]

[a]*Department of Management and Accounting, Qom Branch, Islamic Azad University, Qom, Iran*
[b]*Department of Management and Accounting, Kashan Branch, Islamic Azad University, Kashan, Iran*

ARTICLE INFO

Article history:
Received 11 October 2018
Accepted 23 March 2019

Keywords:
Multi-factor models
Capital asset pricing
Penalty function
Sudden shocks
Stock return evaluation

ABSTRACT

One of the main concerns of investors is the evaluation of the return on investment, which is conducted using various models such as the CAPM (single-factor model), Fama-French three/five-factor models, and Roy and Shijin's six-factor model and other models known as multi-factor models. Despite the widespread use of these models, their major drawbacks include sensitivity to unexpected changes, sudden shocks, high turbulence of price bubble, and so on. To eliminate such negatives, the multi-factor model using the penalty function method is used, in which, instead of averaging, the optimization and avoidance of the effects of abnormal changes and other factors affecting the capital market are considered. In order to evaluate stock returns, it is possible to select effective factors, to simulate and develop a model appropriate to the conditions governing the capital market in Iran. In the present study, by forming portfolios of investments and identifying and refining effective factors, the classification and estimation of the hybrid model of penalty and multi-factor (P and PCA) functions were performed based on the functional data during 2007-2017. The results of this study indicated that the extensive use of the simulation algorithm for the penalty function in the form of P and PCA estimation method improves the efficiency of multi-factor methods in stock return evaluation, and that the use of the hybrid algorithm of penalty and multi-factor functions, compared to the exclusive use of multi-factor models, brings a higher accuracy in estimating stock returns.

1 Introduction

One of the methods to evaluate assets (e.g., stock assessment) is using multi-factor models. Some techniques such as maximum likelihood are often applied to estimate these models, according to which the most relevant factors are selected (e.g., Fama-French and Carhart models). Due to the ability to expand useful information from a large number of related variables, multi-factor models and their derivatives are widely exploited in economic forecasts and analyses. In these models, information production processes are often based on the linear combination of common related factors and error conditions [1]. The estimation of multi-factor models, in cases where the conventional relevant

*** Corresponding author Tel.: *+989123531318***
E-mail address: *Hosein_jahangirnia@yahoo.com*

factors are intangible, can be faced with some problems. Therefore, one of the fundamental goals of estimating such models is to identify the conventional latent factors and their information load. Some of the common techniques to estimate the models include maximum likelihood estimation (MLE), Markov chain Monte Carlo (MCMC), and principal component analysis (PCA). Accordingly, PCA is more applicable, compared to MLE and MCMC, mainly due to its less computation volume. Despite the east of computation by PCA, it is incapable of accurate estimation of latent factors and their information load in case of appearance of several common factors [2]. Despite the extensive use of multi-factor models in areas such as asset evaluation, their problems must not be neglected. In this regard, the proposed algorithm (integration of the multi-factor model and penalty function) has taken steps toward the elimination of some of these issues. Firstly, the basic problem of multi-factor models is sensitivity to unexpected changes (e.g., sudden shocks), and in a sense, unconventional fluctuations [3]. Secondly, a specific number of factors in the multi-factor models are used in the estimation of stock returns as explanatory variables, relying on which in various asset markets might not provide a suitable explanation. Thirdly, unlike the commonly used research in Iran, the proposed model applies one of the multi-factor asset pricing model or comparing some of these models. Rather, the current process research is based on modelling of stock valuation as a capital asset in the Iranian capital market. The present research aimed to recognize and classify the factors affecting the value of asset in the capital market of Iran based on underlying research, evaluation of factors based on factor analysis, estimation of the proposed multi-factor model using penalty function or modelling, as well as testing and validating the model.

In order to solve the problem of accurate estimation of latent factors and their information load, a simple and applicable method is proposed in this research, goal of which is reducing errors in the calculation of conventional latent factors and information loads through simultaneous consideration and estimation of conventional and unconventional components. First, the estimation issue was formulated in the form of a least squares model along with a penalty function for unconventional components. Afterwards, the estimation issue was solved through creating a rational and regular algorithm in order to solve the PCA and single-factor estimation issues. The proposed model, which is an integration of PCA and penalty function, is called PandPCA, meaning an integration model of penalty function via least squares method and analysis of essential components. One of the fundamental differences between multi-factor models and the proposed model (PandPCA) is being based on a method of data production process, which exists in some of the unconventional factors. In data processing, the facts are overshadowed due to severe turbulences in the stock market, such as economic bubble. However, such turbulences cannot be attributed to constant changes in factors or their factor loads. In general, one can estimate factors and sustainable factor loads by using PCA under appropriate assumptions regarding specific and unconventional components [4]. The main question raised in the current study is: what is the result of designing and using a suitable multi-factor model for evaluation of efficiency based on integration of penalty function and multi-factor analysis model in companies accepted in the stock exchange market of Tehran?

2 Theoretical Background of the Research

Research conducted by Alaleh et al. [5] and Ebrahimi and Saeedi [6] demonstrated that Iran's capital market is not efficient even at weak levels. In inefficient markets, there is a distance between daily prices of security and its intrinsic price. Multi-factor models and their derivatives are extensively applied in economic forecasts and analyses owing to their ability to expand information. In these models, information production processes are often based on the linear combination of mutual related factors and error conditions [1]. Evolution of multi-factor models in the evaluation of capital assets is shown in Table 1.

Table 1: Theoretical extension of CAPM

Type of model	Asset valuation model	Valuation model initiator
Static pricing models	Markowitz mean-variance model	Markowitz [7,8]
	Sharpe-Lintner CAPM	Sharpe [9], Lintner [10], and Mossin [11]
	Black's zero-beta CAPM	Black [12]
	CAPM with human capital unrelated to market	Mayers [13]
	CAPM based on multiple consumer goods	Breeden [14]
	International CAMP	Solnik [15], Adler and Dumas [16]
	Arbitrage pricing theory	Ross [17]
	Fama-French three-factor model	Fama and French [18]
	Attribute-based attitude model	Hogan and Warren [19], Bawa and Lindenberg [20], and Harlow and Rao [21]
	Three-attribute CAPM	Rubinstein [22], Kraus and Litzenberger [23]
	Four-attribute CAPM	Fang and Lai [24], and Dittmar [25]
Dynamic pricing models	Temporary CAPM	Merton [26]
	Consumption-based CAPM	Breeden [14]
	Production-based CAPM	Lucas [27], and Brock [28]
	Investment CAPM	Cochrane [29]
	Liquidity CAPM	Acharya and Pedersen [30]
	Conditional CAPM model	Jagannathan and Wang [31]
	Four-factor model of Carhart	Carhart [32]
	Four-factor model of Hu et al.	Hu et al. [33]
	Five-factor Fama-French model	Fama and French [34]
	Five-factor model of IVGMM	Racicot and Rentz [35]
	Six-factor model	Roy and Shijin [36]

Some empirical evidence that is explained and supported based on behavioral finance included high volume anomalies Shiller [37], (Equity Risk Premium puzzle of Mehra and Prescott [38], returns volatility Shiller [37], and returns predictability Fama and French [39]. One of the main issues of behavioral finance is the psychological phenomenon faced by people Shiller [40]and includes overconfidence Daniel et al. [41], overreaction De Bondt and Thaler [42], optimism Statman [43], accessibility to pattern approximation Barberies and Thaler [44], regret avoidance Statman [43], functionality approximation Tversky and Kahneman [45], normative approximation Tversky and Kahneman [45], the inability to use optimization in practice Benartzi and Thaler [46], false detection Saunders [47], and social events Shiller [37].

2.1. Valuation Models based on Effective Factors

Manuscripts must be written in English. Authors whose native language is not English are recommended to seek the advice of a native English speaker, if possible, before submitting their manuscripts. In the text no reference should be made to page numbers; if necessary, one may refer to sections. Try to avoid excessive use of italics and bold face. Also, the Editors reserve the right to return manuscripts that do not conform to the instructions for manuscripts for manuscript preparation or papers that do not fit the scope Journal, prior to referring.

Studies conducted on asset pricing models are related to the CAPM based on Sharpe research [9], in which the returns of each portfolio are merely the result of a systematic risk. After that, the three-factor model of Fama and French [48, 49] was proposed with the increase of two new factors of opportunity to grow and company size, which justified the inability of the single-factor CAPM but was unable to justify the momentum's strategy based on purchasing and maintaining stocks with high returns and selling those with low returns, which was invented by Jegadeesh and Titman [50].

Afterwards, the four-factor model of Carhart [32] was able to reduce irregularities raised in criticism of previous models by developing a three-factor model. Black and Scholes [51], Lintner [10], Mossin [11], and Sharpe [9] developed the relationship between risk and return in explaining the asset pricing models. Fama and French [52] proposed a five-factor model by adding two new factors of investment and profitability to the three previous factors to match the diversity of asset returns in the portfolio, using the mentioned factors in the pricing model as variables that can provide more explanations. Based on theoretical arguments and empirical evidence, Chiah et al. [53] demonstrated that in the global competitive environment, the Fama-French five-factor model [52] had a better performance in explaining changes in the returns and evaluating capital assets, compared to the other single or multi-factor models. Kubota and Takehara, [54] marked that the Fama-French five-factor model [52] provides less realistic explanation of changes in asset returns. Campbell [55] was interested in finding valuation strategies and determining the size associated with the components of human capital. According to the results obtained by Kim et al. [56], the human capital component affects the predictive power of value and size strategies. Mayers [13] identified the role of the component of human capital as a valuable component of total wealth in predicting asset returns. Kuehn et al. [57] evaluated the dynamism of interaction between the work market and financial markets for identifying the factors that determine the cross-sectional stock returns, concluding that human capital represents a significant part of variability of asset returns. In the study by Roy and Shijin [36], the dynamics of human capital component, common factors and most of the financial variables in asset pricing and changes in the prediction of profit return on assets were evaluated at the international level. According to their results, the human capital component provides the explanatory power of the company's size and value strategies in predicting the return on investment (ROI). After the five-factor Fama-French model [52], a five-factor model was tested by Racicot and Rentz [58] based on panel data and with the use of functional data of the Fama-French five-factor model [52]. According to the results obtained by these researchers, the mentioned model can be significant only in evaluation of portfolio market. Roy and Shijin [36] evaluated the ability to explain changes and predictability of their proposed six-factor model in the capital market of the United States for the first time. Applying the functional data used in the Fama-French five-factor model [52], they compared explanatory power of the invented six-factor model with the Fama-French five-factor model and even the five-factor Racicot and Rentz model [58].

Khani et al. [59] examined the expected returns of Carhart model compared to the capital asset pricing model and the implicit capital cost model based on cash and capital returns of growth and value stocks. The results showed that in the case of growth stocks, the expected returns on the basis of Carhart model are closer to real returns compared to expected returns based on the capital asset pricing model. But about value stock, the expected returns on the basis of Carhart model are not closer to actual returns compared to expected returns based on the capital asset pricing model and the cost of capital, and ultimately for growth stocks, expected returns based on Carhart model compared with expected returns, the implicit capital cost model is closer to actual returns. Matevž et al. [60] introduced the eight-factor asset pricing model as an extension of the Fama and French five-factor model. In addition to former factors, they propose three additional factors that represent momentum, liquidity and default risk. They find that the incorporation of additional factors improves the model's explanatory power. Cao et. al [61] documented the existence of five investment-related anomalies in the Australian market. Cross-sectional stock returns are negatively related to each of asset growth, net operating assets, inventory growth and investment-to-assets, and positively related to asset tangibility. While the investment-return relation is theoretically motivated by q-theory, there is only support for the q-theory explanation in relation to the investment-to-assets effect. Limits to arbitrage appear to be a factor in the asset-tangibility effect, where the mispricing can be traced to the over-pricing of stocks with high levels of goodwill.

Ramzi Radchobeh et al. [62] represented a research method that was correlative descriptive and statistical sample consisted of corporate accepted in Tehran Stock Exchange during 2012-2017. To test the hypotheses, regression analysis has been utilized. Results revealed the existence of ambiguity in Tehran Stock Exchange, which affects the asset pricing negatively.

2.2. Penalty Function in Estimation of Valuation Model

Penalty function is a type of algorithm used to solve constrained optimization problems (mathematics). The penalty function technique replaces a constrained optimization issue with a set of unconstrained issues. Unconstrained issues are created by adding a condition to a goal function that consists of a penalty parameter and a level of constraints. When the constraints are violated, the level of the amount of violation is opposite zero, and when the constraints are not violated, the violation level is equal to zero. In 1991, parameters of negative penalty were introduced in the modeling of ranges of structural systems in order to estimate the natural frequencies applying Reyleigh-Ritz Method. In a general model, penalty function can be defined in an optimized form as shown below. If we hypothesize that the constrained minimization function below is assumed:

Min f(x)

s. t:

$$c_i(x) \geq 0 \ \forall i \in I. \tag{1}$$

The penalty function can be defined in the form of an unconstrained penalty function model and within the framework of a data-mining model based on meta-heuristic algorithms and according to the findings of Ando and Bai [63], as presented below:

$$min\, \emptyset_k(x) = f(x) + \sigma_k \sum_{i \in I} g(c_i(x)) \tag{2}$$

In this new model, we have:

$$g\left(c_i(x)\right) = min(0, c_i(x))^2 \tag{3}$$

Therefore, g(Ci(x)) is a penalty function in the new model, in which δk is penalty coefficients. In each K repetition of the penalty function optimization model, the penalty coefficient is elevated (e.g., with coefficient of 10, compared to before), the new problem is solved with constraints and the response obtained in each repetition is used as a basic response for the next repetition until reaching the optimal response. The use of penalty function as a method for estimation of pricing model is based on the principal idea that the error function is defined as the squared difference between the actual value of stock returns and the estimated value of the parametric function based on the factors affecting the stock returns. Moreover, identification of effective factors is initiated by counting penalty in estimation of a range of factors and reaches the most effective factors and the rational estimation pattern by step-by-step removing of factors. One of the methods for finding optimal values is the use of meta-heuristic algorithms, such as genetic algorithms. In addition, penalty function is applied to consider limitation. Since the penalty function reduces the value of the target function in each repetition, the possibility of selecting a point that is not within the limit of the problem for the next step is reduced but not eliminated. This is mainly due to the fact that the justified point close to the optimal point provides more useful information, compared to the justified point away from the optimal point. There are one or several constraints in each optimization problem. For instance, in capital asset valuation, efficiency must not be negative or exceed the maximum defined in the data domain. In such an issue, production of unjustified points is possible. Therefore, the main question is finding a solution to prevent the emergence of such points. Four methods have been proposed to solve this problem and eliminate responses outside the justified zone. In this regard, one of the most important techniques is application of penalty function to solve the problem of unjustified points Takács [64]. This function maintains the unjustified points and reduces their target function amount, which can be carried out through various techniques. Accordingly, the most common method for reducing target function is multiplying the value of a function in one number or adding it to a number. However, the number must be selected in a way that first: increased distance between the point and the justified border be associated with increased amount of penalty, and second: the value of penalty be in a way that the points outside the border are not completely eliminated and are not replaced by the points inside the justified border Francis [65]. By combining the Francis's proposed barrier function with the aim of increasing the performance of penalty function, Merton [26] marked that the new penalty function had a suitable applicability in solving problems with slight changes. Cho et al. [3] relied on performance data of capital market of Shanghai in evaluation of using the penalty function in estimation of stock returns. These researchers exploited the Fama-French three-factor model [49] along with a number of other factors in the form of a penalty function to estimate the stock returns. The results obtained by these scholars were indicative of higher efficiency of the model based on penalty function, compared to the three-factor model.

3 Research Model and Materials

This article is based on a theoretical deduction approach following the integration of a multi-factor model of asset pricing and penalty function in the Tehran Stock Exchange to assess stock returns. Therefore, it is a theoretical research. The design and application of the model are in the form of field experiment with the goal of helping investors and activists of the capital market in the area of better capital decisions. It is notable that the current research is applied in terms of goal. A part of the statistical population of stock exchange companies was selected randomly as a generalizable sample for testing the model. The present study is descriptive in terms of the method of deducing the expression of sample observations and is inductive in generalization to the statistical community of stock exchange companies. The research design is post-event, experimental-field, retrospective, or descriptive-analytical based on past experiences due to the use of functional statistical data related to the last periodic intervals and the past timespan. General framework (model and algorithm) of estimation of multi-factor model, which is based on stock value and risk and is a function of unconventional factors, was discussed. It is assumed that the n-dimensional time series including the selected variables of X_t for prediction of dependent variable of Y_t are limited to the data produced through the process below:

$$X_t = \wedge F_t + J_t + e_t \tag{4}$$

$$Y_{t+h} = \beta_F^T F_t + \beta_W^T W_t + \varepsilon_{t+h} \tag{5}$$

Where t: 1, 2,.., is defined for T. In this equation, K is the matrix of factors with rate of N*r, F_t is vector with rate of r*1 of latent factors, ε_t is a vector with rate of r*1 of errors in measurement, and J_t is a vector with rate of r*1 of special latent factors. In the mentioned equations, b_F is a vector with the rate of r*1 of the coefficients of regression estimation for latent factors defined, W_t is a vector with the rate of m*1 of evident endogenous variables (external), and b_w is a vector with the rate of m*1 of regression estimation coefficients for W_t. In addition, the h index is indicative of forecast horizon and W_{t+h} and ε_{t+h} demonstrates the prediction and error variables in estimation of the future time horizon. The mentioned equations are determined based on the multi-factor model applied in a research by Stock and Watson [66] and modified variables of Cho et al. [3]. However, the difference is that in this research, the vector of X_t variables encompasses especial J_t unconventional components of variables (factors) in the defined equation. Based on the assumption that the appearance of unconventional variables rarely occurs, J_t is usually the variables of a sparse vector, meaning that some of the elements of this vector are equal to zero and can be specifically a representation of increase in values.

3.1 Algorithm of Estimation of Multi-factor Model

The following stages were carried out in factor analysis method to explain the process of the multi-factor model or PCA for estimation of Ft latent factors and uploading of K uploaded factors.

First Stage, Factor Recognition: Explanatory variables or possible factors affecting stock returns are recognized using the literature (or applying background research model)

Explanatory variables or possible factors affecting stock returns are identified by using compilation of literature (or by applying a background research pattern), and are listed in the form of vector of variables of $x = (x_1 \ldots x_T)^T$.

Second stage, classification of factors: Explanatory variables or effective factors based on factor analysis model are divided into two categories of latent and obvious variables. In this classification, variables or effective factors are divided into categories of effective $F = (F_1 \ldots F_T)^T$ and hidden $J = (J_1 \ldots J_T)^T$ factors.

Third Stage, Model Estimation: Assuming that the number of variables or factors possible is at least equal to the number of effective factors (N ≥ T) and the maximum number of selected factors were recognized as effective endpoints or unknown explanatory variables (r). In this case, we can estimate the Factor matrix F and matrix K based on the optimization problem solving without using J_t:

$$min_{F.\Lambda} \frac{1}{TN} \|X - F\Lambda^T\|_F^2, subject\ to\ \frac{F^T F}{T} = I_r \tag{6}$$

In this equation, $\|.\|_F$ is the Frobenius norm in this estimation. The optimization problem defined in equation 3 is largely close to the multi-factor estimation method. The estimated factor matrix of $\hat{F}$ can be obtained from multiplying $\sqrt{T}$ into a matrix consisting of an Eigen vector related to the largest r of special factors of XX^T matrix with rate of T*T. By obtaining $\hat{F}$, factor loading matrix of $\hat{\Lambda}$ can be obtained applying least squares method in the form of $\bar{F} = ((\bar{\Lambda}^T\bar{\Lambda})^{-1}\bar{\Lambda}^T X^T)^T = X\bar{\Lambda}/N$. On the other hand, in cases where N≥T, solving of the mentioned model can lead to determining both the effective and final factors through replacing the limitation of $F^TF/T = I_r$ with the statement of $\Lambda^T\Lambda/N = I_r$. Factor loading matrix of K can be obtained by multiplying $\sqrt{T}$ in the matrix consisting of a special vector related to the largest r of special factors of XX^T matrix with rank of T*T. By considering the z, $\hat{Z}$, and $\bar{Z}$ matrixes as $Z=F\Lambda^T$, $\hat{Z} = \hat{F}\hat{\Lambda}^T$, and $\bar{Z} = \bar{F}\bar{\Lambda}^T$, the $\hat{Z}$, and $\bar{Z}$ matrixes were considered as the lowest estimation rates per x matrix. Therefore, the target function of $\|X - F\Lambda^T\|_F^2/(TN)$ will reach two optimal values of $\hat{F}, \hat{\Lambda}$) and $\bar{F}, \bar{\Lambda}$).

3.2 Combination of Penalty Function and Model

If the special vector of unconventional components of J_t is used in estimation of F and K during data production process based on equation 1 and in determining the effective and final factors involved in prediction of stock returns, it might lack the necessary efficiency in cases where the multi-factor estimation method is directly applied to estimate the XX^T and X^TX matrixes. If J is known, F and K can be better estimated based on a multi-factor model according to the CC^T and C^TC matrixes. Meanwhile, C = X – J. If J is unknown, F and K can be estimated at first by calculating $\hat{J}$ instead of J and using the multi-factor model in estimation of $\widehat{C\hat{C}}$ or $\widehat{C^T\hat{C}}$ matrixes while $\hat{C} = X - \hat{J}$ is a matrix that provides a different estimation of $\hat{J}$ for X matrix. The strategy used in the estimation of J based on the application of a property of J is assuming the presence of scatter matrix. One of the conventional methods for estimation of J parameter is attributing the penalty coefficient of l_p to J while 0≤p≤1. Therefore, the focus is on penalty level of l_1 in the penalty function. It is assumed that N > T and the number of r factors is unknown. Therefore, F, K, and J will be estimated by solving the penalty model below based on the penalty level or l_1 penalty:

$$\min_{F.\Lambda J} \frac{1}{TN} \left\|X - F\Lambda^T - J\right\|_F^2 + \frac{\delta}{TN} \|J\|_1, \text{subject to } \frac{F^T F}{T} = I_r \tag{7}$$

In this model, $\delta \in R^+$ is a penalty parameter and $\|J\|_1$ is equal to total absolute values of each ele-

ment in the J vector. In this model, the penalty level of l_1 might allocate the level of penalty function in estimation of scattering of data to itself in the majority of cases.
Examples of application of this model can be observed in studies such as the research by Cho et al. [3]. The penalty level of l_1 was a convex function of Eigen vector of J. This property of the Eigen vector causes the moderated estimation problem to be achievable, even if T and N are very large. The proposed model is named with the symbol P-PCA in the multi-factor model of valuation of stock returns and risk because it is a combination of multi-factor estimation and penalty function. In solving the model presented in the recent equation, a mathematical algorithm is proposed which, in each repetition of which one of the components of the multi-factor analysis model is obtained, considered and solved in each repetition as a single-valued function. The summary of this algorithm is presented in the estimation of the multi-factor model based on the integration function and the penalty function as follows.

3.3 Selection of Number of Factors

The adaptation of the proposed algorithm and, during the simulation process, the criterion ICp developed by Bai and Ng [4]has been used to estimate the number of the parameters affecting the prediction of return *r*. Assume that N> T and$\hat{F}(r)$ is an estimate of the special factor matrix with the rank T * r (a matrix which contains $\hat{F}(r)$as the number of times for the first value *r* of the Eigen vector $\hat{F}(r)$). In addition, suppose that:

$$V\left(r, \hat{F}(r)\right) = min_{\Lambda} \frac{1}{NT} \left\| \hat{C} - \hat{F}(r)\Lambda^T \right\|_F^2 \tag{8}$$

Bai and Ng's 2002 [4]three-dimensional ICP is defined as follows:

$$\begin{aligned} IC_{P^1}(r) &= lnV\left(r, \hat{F}(r)\right) + r\left(\frac{N+T}{NT}\right) ln\left(\frac{NT}{N+T}\right) \\ IC_{P^2}(r) &= lnV\left(r, \hat{F}(r)\right) + r\left(\frac{N+T}{NT}\right) ln(min(N,T)) \\ IC_{P^3}(r) &= lnV\left(r, \hat{F}(r)\right) + r\left(\frac{ln(min(N,T))}{min(N,T)}\right) \end{aligned} \tag{9}$$

By defining $\hat{r}_i = argmin_r IC_{p^i}(r)$ for i = 1, 2, 3, or triple criteria, the estimated value *r* is placed as the least estimated value of these three criteria, and it is assumed that$\hat{r}_{BN} = min(\hat{r}_1, \hat{r}_2, \hat{r}_3)$. In each iteration, the ICp specification method will also be implemented. $\hat{r}_{BN} = min(\hat{r}_1, \hat{r}_2, \hat{r}_3)$is then used, which is a matrix containing the first estimated Eigen vector; $\hat{C}\hat{C}^T$and $\hat{\Lambda}(\hat{r}_{BN}) = \hat{C}^T \hat{F}(\hat{r}_{BN})/T$ are the matrices of estimated factors and loaded factors.

3.4 Determining Penalty Parameter

In the proposed P-PCA model as a combination of multi-factor estimation models and penalty function, the penalty parameter δ is of particular importance, with no specific rule for specifying it. In this study, the penalty parameter is defined as $\delta^{naive} = \bar{\hat{\sigma}}\sqrt{8lnT}$ and $\bar{\hat{\sigma}} = N^{-1}\sum_{i=1}^{N} \hat{\sigma}_i$, where $\hat{\sigma}_i$ is the non-detailed estimation of the parameter δ in the hybrid proposed model. This is the estimated standard deviation of the parameter penalty. The approximate estimates are one of the most commonly used methods in simulation. The idea is that the desired loss in the estimation of the J_{it} can only be achieved if the estimating threshold δ leads to a proper value Donaho and Johnston [67].

Assume that $\omega_{it} = J_{it} + e_{it}$ represents an unconventional component of the special error in the data matrix X_{it}, and e_{it} is also a certain standard deviation value with a standard normal distribution. In this case, it is assumed that J_{it} is estimated by zero or X_{it}. The optimal average value of the loss squares associated with the estimator for t = 1, 2, ..., T is equal to $Loss_i^{oracle} = \sum_{t=1}^{T} min(J_{it}^2, \sigma^2)$, while | Jit | > σ is unknown. Without such information, it has been proved that the average loss squares $\tilde{J}_{it}$ for t = 1, 2, ..., T can lead to losses$Loss_i^{oracle}$, while $\tilde{J}_{it} = ST(\omega_{it}, \sigma\sqrt{2lnT})$is used in the estimation of J_{it}. Assume that $\hat{L}_{it}$is the estimated value calculated by deducting X_{it} from the estimated and loaded conventional factors. Note that the J_{it} estimation with $\tilde{J}_{it} = ST(\hat{L}_{it}, \sigma\sqrt{2lnT})$ is equal to putting it with $\delta = \sigma\sqrt{8lnT}$in the proposed P-PCA model. If the conventional and loaded factors are accurately estimated, then $\hat{L}_{it} \approx \omega_{it}$ and $\hat{J}_{it} \approx \hat{J}_{it}$will be obtained. For t = 1, 2, ..., T, the mean squares of the loss $\hat{\mathrm{J}}_{it}$results in the estimation of the average loss squares $\hat{J}_{it}$and thus the loss $Loss_i^{oracle}$ reaches an appropriate value. Theoretical foundations and the proposed model extension show that the P-PCA model works well under different data conditions. The extension of computational complexity and details has been ignored. An approximate approach may not be the best option; however, it can be easily applied to indicators and further adjustments to achieve desired results.

3.5 Extending the Proposed Model

In the proposed P-PCA model, there are other examples of the X_{it} data application, which possess more complicated process of data generation. For example, you can see the matrix X_{it} so that:

$$X_{it} = U_{it}^T \beta_{U+} F_t^T \lambda_i + J_{it} + e_{it} = X'_{it} + J_{it} \tag{10}$$

where, U_{it} is a P * 1 matrix of the observed variables and β_U is also P * 1 vector of the coefficients of these variables. The data generation process for the matrix X^0_{it} (X_{it} without J_{it}) is included in Bai's model [64]. Assume that $U_{it} = 1$ and $\beta_U = \mu$, then we will have:

$$X_{it} = \mu + F_t^T \lambda_i + J_{it} + e_{it} \tag{11}$$

In Equation 11, one extra assumption is that μ is constant over time and $\bar{F} = T^{-1} \sum_{t=1}^{T} F_t = 0$. In the X_{it} data, it can then be ignored and deleted. In this data, the parameter μ is not included; however, it contains the factor $(F_t - \bar{F})$of the error and mutation parameters in this equation. Simplified data is a form of exclusive factor structure with a special jump factor, in which the proposed P-PCA model is used. A more general strategy for estimating β_U can be considered when J_{it} can be used by simply modifying the proposed algorithm in the previous sections. Bai [68] showed that, if $X'_{it}$is achievable, a hybrid algorithm combining the least squares method and multi-factor estimation can be used to estimate βU, F_t, and k_i. The basic idea in this estimation is that F and ʌ can be estimated based on the regression estimation of $X'_{it} - F_t^T \lambda_t$on the basis of the U_{it} observation. When V´ = X´ - ΓU = FʌT + e, if the values of F and K are known, we can estimate β_U and vice versa. This estimation method can be defined in the form of a repetition of a hybrid algorithm combining the least squares method and multi-factor estimation and the simulation process continues until the estimations are carried out. If J_{it} is specified, instead of multi-factor estimation, we can exploit the hybrid model and distinguish them from other factors in the data matrix. For β_U, we can use the hybrid model for the matrix VV^T , while $V = X - \Gamma U = F\Lambda^T + J + e$ is a simplified form of the pure factor structure in the improved Eigen vector matrix. An algorithm can be defined with two repetitive loops, whose external loop is to estimate the coefficients of the variables or β_U and uses the least squares method to estimate U_{it}, and whose inner loop is also to estimate F, Λ, and J.

4 Results and Discussion

The proposed model of research was used to evaluate the stock returns based on the integration of the multi-factor model evaluation capital asset and penalty functions in Tehran Stock Exchange. After selecting the companies and forming the investment portfolios, factors affecting stock returns were identified based on an analysis of the knowledge field. With determining the most effective factors, the penalty function was defined based on the final factors and the simulation of multi-factor model and validation of the estimated model were then performed.

4.1 Formation of Investment Baskets

The spatial area of the study encompassed companies accepted in the Tehran Stock Exchange during a ten-year period ended on March 20, 2019. In order to homogenize the companies and to measure the study variables, the following inclusion conditions were applied to determine the selected companies: A) The company's fiscal year ended on March 20 and there was no change of fiscal year over the considered period; B) The company had been a member of the Exchange Stock Market prior to the period under study and its membership is not terminated; C) The company's data required to measure variables, especially daily stock price changes, are available; D) Their shares are traded during the study period and there is no more than three months of trade termination; and E) The company does not belong to financial and investment intermediary companies and banks. Accordingly, 118 companies were selected. In order to implement the simulation process, a combination of three companies in each industry was considered, and the stock prices and changes for each company were calculated monthly. Then the collected information for each stock portfolio was processed using MATLAB software.

4.2 Identification of Effective Factors

The factors affecting stock prices are as follows:
1. Liquidity of a company's stock. One of the important factors for the secondary investors is liquidity. Many financial investors tend to criticize their stock in the case of necessity; therefore, high liquidity of a company's shares is considered as an interesting criterion for the shareholders Morovati Sharifabadi and Golshan [69].
2. Number of shares available to the public: With an increase in free float shares, the possibility of

manipulating the stock price decreases and the stock liquidity enhances. Float shares are stocks that potentially allow trading and are not blocked by certain institutions or organizations Kashanipour et al. [70].

3. Composition of shareholders: Major shareholders, such as institutions, various organizations and investment companies, are usually judged by their performance in support of their stock prices in previous years, and, when the composition of a major shareholder of a company changes, micro-shareholders respond to it according to the new shareholder's performance Mashayekhi and Panahi [71].

Table 2: Summary of Factors Affecting Stock Returns or Valuation Roy and Shijin [36]

Type	Factors	Criteria (Metrics)
Financial Ratios	Liquidity ratios	Current ratio, quick ratio, current assets ratio, net working capital, liquidity ratios
	Activity ratios	Debt settlement period, current assets turnover, fixed asset turnover, asset turnover
	Asset ratios	Ownership ratio, interest coverage ratio, long-term debt to equity ratio, current debt to equity ratio
	Profitability ratios	Net profit to sales ratio, operating profit to sales ratio, gross profit to sales ratio, net p gross profit ratio, return on assets (ROA), return on equity, return on working capital rati asset return ratio
Stock valuation models	Solnik (1974a), CAMP	$R_{i,t} - R_{f,t} = \alpha_1 + \beta\,(R_{m,t} - R_{f,t}) + \varepsilon_{i,t}$ Beta or market risk premium, difference between the market returns rate and risk-free returns rate
	Gordon model	r_m or systematic risk = expected market returns through EPS, DPS, EPS forecast, EPS coverage, difference between real and expected EPS, EPS growth compared to last year, $P=\frac{DPS}{k-g}$ as DPS= EPS*DPR
	Campbell and Shiller model	P/E, P/S
	Walter model	Stock accumulated profit $P=\frac{DP\,DPS+(EPS-DPS)r/k}{k}$
	Three-factor Fama and French Mode (1993)	market risk premium, size, growth opportunities, $R_{i,t} - R_{f,t} = \alpha_1 + \beta_1\,(R_{m,t} - R_{f,t}) + \beta_2\,SMB_{i,t} + \beta_3\,HML_{i,t} + \varepsilon_{i,t}$
	Four-factor model of Carhart (1997)	market risk premium, size, growth and profit-making opportunities, $R_{i,t} - R_{f,t} = \alpha_1 + \beta_1\,(R_{m,t} - R_{f,t}) + \beta_2\,SMB_{i,t} + \beta_3\,HML_{i,t} + \beta_4\,WML_{i,t} + \varepsilon_{i,t}$
	Five-factor Fama and French Model	market risk premium, size, and growth, investment, and profit-making opportunities, $R_{i,t} - R_{f,t} = \alpha_1 + \beta_1\,(R_{m,t} - R_{f,t}) + \beta_2\,SMB_{i,t} + \beta_3\,HML_{i,t} + \beta_4\,RMW_{i,t} + \beta_5\,CMA_{i,t} + \varepsilon_{i,t}$
	Six-factor Roy and Shijin Model (20	market risk premium, size, and growth, investment, and profit-making opportunities, human resources $R_{i,t} - R_{f,t} = \alpha_1 + \beta_1\,(R_{m,t} - R_{f,t}) + \beta_2\,SMB_{i,t} + \beta_3\,HML_{i,t} + \beta_4\,RMW_{i,t} + \beta_5\,CMA_{i,t} + \beta_6\,LBR_{i,}$
Profit and Loss Report		Expected total income, income growth rate (total real income to the difference between total and expected total income), expected profit margin, growth rate of profit margins (real profit margin to the difference between real and expected profit margin), and efficiency (Percentage of exchange value to the company value in the previous period)

4. *Profitability* (EPS) and its stability: *Profitability is the most* important factor influencing a company's stock prices, and all other factors indirectly affected affect the stock price depending on its profitability or non- profitability. For example, when the CEO of a company changes, it is expected that the performance of the CEO and the profitability of the company change the stock value in the capital market. In other words, changing the manager changes the profitability expectations Makian and Mousavi [72].

5. Company's Development Plans: A company's development plans can represent the growth and

dynamism of the company. It should also be noted that the development plans that are to replace worn machinery are less important than the development plans aiming at the construction of a new production line Monajemi et al. [73].
6. Management of a company: The managers of the companies are also estimated and known based on their past performance in the stock exchange, and their transfer would change the stock prices of the companies Mehrara et al. [74].
7. Other factors: Other factors affecting the stock prices include factors such as the company's life, depreciation of machinery, age and reputation of the company, share price trend, lawsuits against the company and their frequency, information within the company, and rumors Kashanipour et al. [70].

Some researchers categorize factors affecting stock returns into two categories: (a) Basic variables including earnings per share and the price/profit ratio of each share; and (b) Technical variables such as inflation, industry status, alternative markets, major trades, age factors of investors, stock liquidity, and emotional variables Ebrahimi and Saeidi [6].
According to Roy and Shijin [36], the factors affecting the valuation of stock assets, which have been used in different capital asset valuation models or in different studies, are known as factors affecting stock returns and are summarized in Table 2.

4.3 Refining Effective Factors

According to Table 2, there more than 10 measures defined on the basis of the balance sheet items, profit and loss or financial ratios and as factors affecting the return on shares based on the analysis of the knowledge field and the application of a penalty function or any other method to estimate the multi-factor model was difficult and did not lead to reliable results due to the limited data volume. Accordingly, the refinement algorithms were used to reduce these factors or to refine the explanatory variables. Using the proposed model, the tolerance matrix and linear correlation between the effective factors and the dependent variable (share risk premium) were used. According to this model, each of the factors having a tolerance factor <10 with the dependent variable was selected as an explanatory variable; otherwise, it was excluded. The six factors used in Roy and Shijin's [36] model, with the least tolerance factor, were selected as the most effective factors (Table 3).

Table 3: Summary of results of tolerance factor analysis

Column	Factor	Definition and Measurement	Symbol	Variance	Tolerance Factor
1	Risk premium of market	Difference between the return rate of market and risk-free return rate	$R_{m,t} - R_{f,t}$	0.9825	4.3625
2	Size	Difference between the average returns of large and small enterprises	$SMB_{i,t}$	1.1251	2.2351
3	Growth opportunities	Difference between the average stock returns with high and low growth	$HML_{i,t}$	1.3526	1.6385
4	Profitability	Difference between the average stock returns with high and low momentum	$RMW_{i,t}$	0.9932	2.9524
5	Investment	Difference between the average stock returns with high and low investment	$CMA_{i,t}$	1.3625	2.3352
6	Human resources	Difference between the average stock returns with high and low human capital	$LBR_{i,t}$	1.0012	1.6385

The results obtained from the STATA statistical software in Table 3 show that the values of variances for the six factors are <2, and that the tolerance factor in all of these six cases is <5. The abovementioned factors are the most effective factors in evaluating stock returns.

4.4 Diagnostic Tests

In order to use various combinations of selected companies as different investment portfolios and to collect monthly data for a ten-year period ended on March 20, 2019, the model used in estimating stock returns was the panel data analysis, for which the diagnostic tests were considered as a prerequisite. The evaluation of linear independence of independent variables was performed as the first diagnostic test (Table 4):

Table 4: Summary of the Results of Evaluation of Linear Independence of Factors Affecting Stock Returns.

Variable	$R_{m, t} - R_{f, t}$	$SMB_{i, t}$	$HML_{i, t}$	$RMW_{i, t}$	$CMA_{i, t}$	$LBR_{i, t}$
$R_{m, t} - R_{f, t}$	1					
$SMB_{i, t}$	0.0931 0.0000	1				
$HML_{i, t}$	0.1024 0.0000	-0.1104 0.0000	1			
$RMW_{i, t}$	0.1304 0.0000	0.1604 0.0000	-0.1412 0.0000	1		
$CMA_{i, t}$	0.1634 0.0000	0.1932 0.0000	0.0987 0.0000	-0.0977 0.0000	1	
$LBR_{i, t}$	0.0331 0.0000	0.0361 0.0000	0.0745 0.0000	0.1192 0.0000	-0.0928 0.0000	1

The evaluation results of the linear independence between the explanatory variables indicated that the Pearson correlation coefficients summarized in Table 4 ranged from -0.1412 and 0.1932 and were inclined towards zero, representing a weak linear correlation. According to the common practices in social sciences, this value is negligible and it is possible to estimate the linear relationship between variables. Other diagnostic tests to estimate stock returns using the STATA software are shown in (Table 5).

Table 5: Diagnostic tests in estimation of the linear multi-factor model of stock returns.

Premise evaluated	Diagnostic test	Test statistic	Level of significance	Norm	Judgement
Normality of the dependent variable	Kolmogorov	5.0125	0.2034	more than five percent	The distribution of the dependent variable is normal.
Linear independence of remnants	Durbin-Watson	1.8289	0.0041	Between 1.5-2.5 percent	The remnants have linear independence.
Equivalence of variances-Fisher	White	54.0127	0.1204	More than five percent	The consistency of variances is established.
Consistency of variances-Chi-square	White	17.4261	0.1740	More than five percent	The consistency of variances is established.
Assessment of the type of data analysis model	Chow	19.0251	0.0125	Less than five percent	Suitability of panel analysis model
Assessment of the type of data analysis model	Hausman	13.2814	0.2012	More than five percent	Suitability of random effects model

Findings of the diagnostic tests using the linear model in estimating the relationship between effective factors and stock returns (Table 5) showed that it was possible to adopt the combined linear model based on panel data analysis with random effects.

4.5 Estimation of Integrated Model

Considering the results for the refinement of the effective factors and the results of diagnostic tests (Table 5) based on the proposed model combining the penalty function and multi-factor model, the proposed model is defined and estimated, which is briefly described below. The hybrid P-PCA model was calculated and it was examined whether, with regard to cross-sectional changes, the expected stock returns can be explained by hidden components. Given the final refined factors of Roy and Shijin's models [33], they were used to estimate the expected returns:

$$R_{i,t} - R_{f,t} = \alpha_1 + \beta_1 (R_{m,t} - R_{f,t}) + \beta_2\, SMB_{i,t} + \beta_3\, HML_{i,t} + \beta_4\, RMW_{i,t} + \beta_5\, CMA_{i,t} + \beta_6\, LBR_{i,t} + \tag{12}$$

In this regard, the dependent variable, like the single-factor models and the previous three-factor models, is the stock risk premium (portfolio), shown as ($R_i - R_f$), which is obtained from the difference between the return on the stock or portfolio R_i and the risk-free return rate R_f. In the single-factor model, we mentioned how to calculate it. The first independent variable is like the single-factor model of market risk premium, indicated by the symbol (R_m-R_f), and is calculated based on the difference between the market return rate R_m and the risk-free return rate R_f . In this case, the risk-free return rate is usually calculated based on the long-term bank deposit interest rate. $SMB_{i\ t}$ is the size factor and $HML_{i\ t}$ is the growth opportunities factor, which is calculated as the three-factor model. $RMW_{i,\ t}$ is used as a profit factor based on the difference in the average returns of companies with high and medium momentum and those with low momentum, and is similar to Carhart's [32] four-factor model. $CMA_{i,\ t}$ is also used as an investment factor and is determined based on the difference between the average returns of companies with high and low profitability. In addition, $LBR_{i,\ t}$ is also used as a human capital factor, based on the difference between the average returns of companies with high and low human capital, and is similar to other factors mentioned in previous multi-factor models. In addition to the six effective factors, u_{it} is part of the expected revenues, which is not mentioned in the above six factors, and is the standard deviation of stock returns and net noise in the estimating the specified performance of the return that is probably due to hidden and unplanned factors in the estimation. It is assumed that u_{it} is controlled by hidden factors and defined in the following equation:

$$u_{it} = F_t^T \lambda_i + \omega_{it} = F_t^T \lambda_i + J_{it} + \varepsilon_{it} \tag{13}$$

Where, ω_{it} is the total jump J_{it}, and ε_{it} is the net noise or the final error term. Depending on the inclusion or exclusion of J_{it}, two new IVOL definitions can be presented as fluctuations in returns for stock portfolio *i*: This variable can be defined based on the standard deviation ω_{it} or standard deviation ε_{it}. To calculate the hidden structure u_{it}, one can first estimate the $R_{i,t} - R_{f,t}$ based on the performance data of the investment portfolios defined in the 10-year period in a monthly basis according to Equation 12 using the STATA software with Composite linear regression analysis, panel data analysis and random effects. The regression estimation results are summarized in Table 6.

Table 6: Estimation of Stock Returns Equation Relation Based on Six Factors

Description of variable	Mathematical symbol	Parameter	T statistics	Possibility
Y-intercept	α_0	0.0055	1.8547	0.0389
Risk premium market	$R_{m,t} - R_{f,t}$	0.0062	1.8854	0.0412
Size	$SMB_{i,t}$	0.0128	1.7542	0.0333
Growth opportunities	$HML_{i,t}$	0.0321	2.0021	0.0482
Profitability	$RMW_{i,t}$	0.0171	1.7854	0.0365
Investment	$CMA_{i,t}$	0.0046	1.7772	0.0341
Human capital	$LBR_{i,t}$	0.0003	1.9985	0.0462
Needs assessment	Coefficient of determination 0.7623		Modified coefficient of determination 0.7241	
Generalizability	Fisher statistics 61.1142		Fisher's level of significance 0.0045	

The results of regression estimation as described in Table 6 showed that the relationship between six factors and stock risk premium as a stock return measure was positive in all cases and because obtained values were <5% corresponding to T-Student, the relationship between the variables was significant at 95% and the estimated relationship could explain between 72.17% and 76.23% of the stock return variations based on the concerned six factors(namely market risk, size, growth opportunities, profitability, investment, and human capital) so that it has a relatively high explanatory power. After estimating the regression and substitution parameters in this equation, it is used as a nonparametric equation. With replacing the values of independent variables in the remaining portfolios, the estimated relationship is obtained as the difference between the expected risk and the real risk, and the proposed algorithm is followed in the estimation of hidden factors. This algorithm is used to investigate the effect of hidden factors on expected fluctuations in stock returns. The calculation $\widehat{\omega}_{i\tau}, \hat{J}_{i\tau}$ and $\widehat{\varepsilon}_{i\tau}$ are made using the P-PCA method with the remained regression $\hat{u}_{i\tau}$. The standard deviation of the sample $\widehat{\omega}_{i\tau}, \hat{u}_{i\tau}$ and $\widehat{\varepsilon}_{i\tau}$ is scaled by the number of days in the month *t* and is used for various calculations of *IVOL* for the asset *i* in month *t*. We used $IVOL^{u}$, $IVOL^{PCA}$, $IVOL^{P\text{-}PCA}$ to refer to these calculations.
To understand whether the IVOL risk is priced or not, it is more rational to use the IVOL conditional expectations instead of the actualized IVOLs. To calculate the conditional expectations IVOL for the asset *i* in month *t*, the following regression function should be first estimated:

$$IVOL_{it} = a_i + b_{i1} IVOL_{i,t-1} + b_{i2} IVOL_{i,t-1}^{6m} + b_{i3} IVOL_{i,t-1}^{24m} + \varepsilon_{it}^{IVOL} \tag{14}$$

Where, $IVOL_{i,t-1}^{6m} = 1/6 \sum_{k=0}^{5} IVOL_{i,t-1-k}$ and $IVOL_{i,t-1}^{24m}$=1/24 are the regression for the estimated $\sum_{k=0}^{23} IVOL_{i,t-1-k}$. In addition, the concerned time period was the period ended on March 20, 2019. Regression calculations were performed using MATLAB software. The coordinated amount of IVOL from the HAR regression is used for computing E_{t-1}[IVOLit]. The HAR regression is usually used in the dynamic modelling of the real asset yield variance and can achieve its consistent behaviour. These calculations based on the simulation algorithm were performed with 100 iterations, and the calculations were terminated in 100th iteration, where there were no further changes in the previous process.

To determine the relationship between the hidden components and the cross-sectional variation of the expected stock returns, Fama-McBeth (FM) regression and MATLAB software were adopted. Suppose that the stock earnings *i* is the linear function of the control variables V_{lit} and l = 1, ..., L, and the random error term ε_{it}^{cross} is then estimated as follows:

$$R_{it} = \gamma_{0t} + \sum_{l=1}^{L} \gamma_{lt} V_{lit} + \varepsilon_{it}^{cross} \tag{15}$$

where, $E_{t-1}[\varepsilon_{it}^{cross}] = 0$. Then expected earnings *i* is the shares of the expected linear function, and $E_{t-1}[R_{it}] = \gamma_0 + \sum_{l=1}^{L} \gamma_l E_{t-1}[V_{lit}]$and the main control variables, based on which the return fluctuations are defined, are $\lambda_i, IVOL_{it}$, and their expected values. For other beta variables of shares, the average market value algorithm from its structures in shares and the average ratio of book value to market (BM) algorithm from its structures in shares and gross income from the last seven to two months were adopted. Moreover, we had information about the specific mutation component in the FM regression so that its impact on expected cross-sectional earnings was to be assessed. $ABSJ_{it} = 1/\#\{\tau \in t\} \sum_{\tau \in t} |\hat{J}_{i\tau}|$ was used to measure the magnitude of the current specific mutation and $E_{t-1}[ABSJ_{it}]$was used to estimate $1/6 \sum_{k=0}^{5} ABSJ_{i,t-1-k}$. All calculations were defined based on the simulation algorithm run in the MATLAB software. For each t^{th} month, the FM regression is calculated with respect to the real and cross-sectional stock returns as a dependent variable and specific control variables as explanatory variables. To understand whether a certain control variable affects the expected cross-sectional stock return, T-Student test was used to calculate the coefficients $\hat{\gamma}_{lt}$ and l = 0, ..., L. When the difference between the real and expected returns approached zero, the simulation process was terminated.

5 Conclusions

In spite of the widespread use of multi-factor models in capital asset pricing, especially in evaluating the stock returns, these models have been influenced by turbulence and sudden shocks due to their dependence on the estimation of the common regressions based on the least squares and the data size so that they have no strong explanatory power in predicting stock returns. In other words, they cannot be used for capital decision making in emerging markets, including Iran's, which, on the other hand, have no proper functioning and, on the other hand, are accompanied with wide fluctuations because of turbulent political situations. In this research, with combining a penalty function and a multi-factor model, the factors affecting the stock returns were first identified based on the field of knowledge, and then, using the factor analysis, the six factors of market risk, size, growth opportunities, profitability, investment, and capital Human being, having a tolerance factor <5, were considered as the most effective ones in predicting the stock returns. The monthly data in the 10-year period ended on March 20, 2019 included in the stock portfolios defined by selected companies along with the linear regression was used to estimate the cross-sectional returns. Then with replacing the parameters and values of the explanatory variables, the difference between the real returns and the expected returns, the remnants of the model were determined. The predicted cross-sectional returns were then estimated using simulation algorithm, penalty function, and monthly FM regressions. The use of the hybrid algorithm of penalty and multi-factor functions, compared to the exclusive use of multi-factor models, brings a higher accuracy in estimating stock returns.

References

[1] Ng, L., *Tests of the CAPM with Time-Varying Covariances: A Multivariate GARCH Approach*, The Journal of Finance, 2009, **46**(4), P.1507-1521, Doi:10.1111/j.1540-6261. 1991.tb04628.x.

[2] Keller, C., Kunzel, P., Souto, M., *Measuring Sovereign Risk in Turkey: An Application of the Contingent Claims Approach*, IMF Working Paper, 2007, WP/07/233, International Monetary Fund, Doi:10.5089/9781451 867978.001

[3] Chou, R. Y., Yen, T. J., Yen, Y. M., *Risk evaluations with robust approximate factor models*, Journal of Banking and Finance, 2017, **82**, P.244-264, Doi:10.1016/j.jbankfin.2016.05.008.

[4] Bai, J., Ng, S., *Determining the number of factors in approximate factor models*, Econometrica, 2002, **70**, P.191–22, Doi:10.1111/1468-0262.00273.

[5] Alaleh, N. Tamimi, M., Nematpour Dezfuli, A. M., *Explanation of Returns Changes in CAPM*, TFPM, FFPM in Tehran Stock Exchange, 2013, P. 115-128, Doi:10.2139/ssrn.2252016, (in Persian).

[6] Ebrahimi, M. and Saeidi, A., *The effect of accounting variables and company features on stock prices in Tehran stock exchange*, Accounting and Auditing, 2010, **17**(62), P.1-16, (in Persian).

[7] Markowitz, H. M., *Portfolio Selection*, Journal of Finance, 1952, **7**(1), P.77-91, Doi:10.1111/j.1540-6261. 1952.tb01525.x.

[8] Markowitz, H. M., *Portfolio Selection: Efficient Diversification of Investment*, John Wiley and Sons, Inc., 605 Third Avenue, New York, US, 1959, Doi:10.2307/3006625.

[9] Sharpe, W.F., *Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*, Journal of Finance, 1964, **19**(3), P.425-442, Doi:10.1111/j.1540-6261.1964.tb02865.x.

[10] Lintner, J., *The valuation of risk assets and selection of risky investments in stock portfolios and capital budgets*, Review of Economics and Statistics, 1965, **47**, P.13-37, Doi:10.2307/1924119.

[11] Mossin, J., *Equilibrium in a Capital Asset Market*, Econometrica, 1966, **34**(4),P.768-83, Doi:10.2307/1910 098.

[12] Black, F., *Capital Market Equilibrium with Restricted Borrowing*, Journal of Business, 1972, **45**(3), P.44-54, Doi:10.1086/295472.

[13] Mayers, D., *Nonmarketable Assets and Capital Market Equilibrium under Uncertainty*, Studies in the Theory of Capital Markets, 1972, **3**, P.223-248, Praeger Publishers, N.Y., (Cited in Jagannathan, R. and Z. Wang. 1996, Doi:10.1086/295528.

[14] Breeden, D. T., *An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities*, Journal of Financial Economics, 1979, **7**(3),P. 265-96, Doi:10.1016/0304-405x(79)90016-3.

[15] Solnik, B. H., *An Equilibrium Model of the International Capital Market*, Journal of Economic Theory, 1974, **8**(4), P.500-24, Doi:10.1016/0022-0531(74)90024-6.

[16] Adler, M., Dumas, B., *International Portfolio Choices and Corporation Finance: A Synthesis*, The Journal

of Finance, 1983, **38**(3),P. 925-984, Doi:10.1111/j.1540-6261.1983.tb02511.x.

[17] Ross, S., *Arbitrage Theory of Capital Asset Pricing*, Journal of Economic Theory, 1976, **13**(3), 341360, Doi:10.1016/0022-0531(76)90046-6.

[18] Fama, E. F., French, K. R., *Common Risk Factors in the Returns on Stocks and Bonds*. Journal of Financial Economics, 1993, **33**(1), P.3-56, Doi:10.1016/0304-405x(93)90023-5.

[19] Hogan, W., Warren, J., *Toward the Development of an Equilibrium Capital Market Model Based on Semivariance*, Journal of Financial and Quantitative Analysis, 1974, **9**(1),P. 1-11, Doi:10.2307/2329964.

[20] Bawa, V., Lindenberg, E., *Capital Market Equilibrium in a mean, Lower Partial Moment Framework*, Journal of Financial Economics, 1977, **5**(2), P.189-200, Doi:10.1016/0304-405x(77)90017-4.

[21] Harlow, W.V., Rao, R. K. S., *Asset Pricing in a Generalized Mean-Lower Partial Moment Framework: Theory and Evidence*, Journal of Financial and Quantitative Analysis, 1989, **24**(3), P.285-311, Doi:10.2307/23 30813.

[22] Rubinstein, M. E., *The Fundamental Theorem of Parameter-Preference Security Valuation*, Journal of Financial and Quantitative Analysis, 1973, **8**(1),P. 61-69, Doi:10.2307/2329748.

[23] Kraus, A., Litzenberger, R.H., *Skewness preference and the valuation of risk assets*, Journal of Finance, 1976, **31**(4),P.1085-1100, Doi:10.2307/2326275.

[24] Fang, H., Lai, T., *Co-Kurtosis and Capital Asset Pricing*, The Financial Review, 1997, **32**(2), P.293-307, Doi:10.1111/j.1540-6288.1997.tb00426.x.

[25] Dittmar. R. R., *Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross-Section of Equity Returns*, Working Paper, University of North Carolina at Chapel Hill, 2002, Doi:10.1111/1540-6261.0 0425.

[26] Merton, R. C., *An Intertemporal Capital Assets Pricing Model*, Econometrica, 1973, **41**(5),P.867-87, Doi:10.2307/1913811.

[27] Lucas, R., *Asset Prices in an Exchange Economy*, Econometrica, 1978, **46**(6), P.1429-1445, Doi:10.2307/1 913837.

[28] Brock, W. A., *Asset Prices in a Production Economy*, Social Science Working Paper 275, California Institute of Technology, 1979.

[29] Cochrane, J. H., *Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations*, The Journal of Finance, 1991, **46**(1), P.209-37, Doi:10.2307/2328694.

[30] Acharya, V.V., Pedersen, L.H., *Asset Pricing with Liquidity Risk*. Journal of Financial Economics, 2005, **77** (2), P.375–410, Doi:10.1016/j.jfineco.2004.06.007.

[31] Jagannathan, R., Wang, Z., *The Conditional CAPM and the Cross-Section of Expected Returns*, Journal of Finance, 1996, **51**(1), P.3-53, Doi:10.1111/j.1540-6261.1996.tb05201.x.

[32] Carhart, M. M., *On persistence in mutual fund performance*, The Journal of Finance, 1997, **52**(1), P.57-82, Doi:10.2307/2329556.

[33] Hu, G. X., J. Pan, and J.Wang, *Noise as information for illiquidity*, Working Paper, 2012, Doi:10.2139 /ssrn.1689067.

[34] Fama, E. F., K. R. French, *A Five-Factor Asset Pricing Model*, Working Paper, 2013, Available at SSRN 2287202 (2013).

[35] Racicot, F.-E., and Rentz, W. F., *Testing Fama-French's new five-factor asset pricing model: Evidence from robust instruments*, Applied Economics Letters, 2017, **23**(6), P.1-5. Doi:10.1080/13504851.2015.1080798

[36] Roy, R., Shijin, S., *Dissecting anomalies and dynamic human capital: The global evidence*, Borsa Istanbul Review, 2018, **18**(1), P.1-32, Doi:10.1016/j.bir.2017.08.005.

[37] Shiller, R. J., *Human Behavior and the Efficiency of the Financial System*, John B., Taylor and Michael Woodford, Editors, Handbook of Macroeconomics, 1998, Doi:10.1016/s1574-0048(99)10033-8.

[38] Mehra, R., Prescott, E.C., T*he Equity Premium: A Puzzle*, Journal of Monetary Economics, 1985, **15**(2), P.145-165, Doi:10.1016/0304-3932(85)90061-3.

[39] Fama, F., French, K.R., *Permanent and Temporary Components of Stock Prices*, The Journal of Political Economy, 1988, **96**(2), P. 246-273, Doi:10.1086/261535.

[40] Shiller, R. J., *From Efficient Market Theory to Behavioral Finance*, The Journal of Economic Perspectives, 2003, **17**(1), P. 83-104, Doi:10.1257/089533003321164967.

[41] Daniel, K. D., Hirshleifer, D., Subrahmanyam, A., *Investor Psychology and Security Market Under- and Overreactions*, Journal of Finance, 1998, **53**(6), P.1839-1885, Doi:10.1111/0022-1082.00077.

[42] De Bondt, W. F. M., Thaler, R., *Does the Stock Market Overreact?*, The Journal of Finance, 1985, **40**(3), P.793-805, Doi:10.1111/j.1540-6261.1985.tb05004.x.

[43] Statman, M., *Lottery Players/Stock Traders*, Financial Analysts Journal, 2002, **58**(1), P.14-21, Doi:10.2469 /faj.v58.n1.2506.

[44] Barberies, N., Thaler R., *A Survey OF Behavioral Finance*, In G.M., 2003, Doi:10.3386/w9222.

[45] Tversky, A., Kahneman, D., *Judgement under Uncertainity: Heuristics and Biases*, Science, 1974, **185** (4157), P.1124-1131, Doi:10.1126/science.185.4157.1124.

[46] Benartzi, S., Thaler, R.H., *Naive Diversification Strategies in Defined Contribution Savings Plan*, American Economic Review, 2001, **91**(1), P.79-98, Doi:10.1257/aer.91.1.79.

[47] Saunders, E. M., *Stock Prices and Wall Street Weather*, American Economic Review, 1993, **83**(5), P.1337-1345.

[48] Fama, E.F., French, K.R., *The Cross-Section of Expected Returns*, Journal of Finance, 1992, **47**(2), P.427-465, Doi:10.2307/2329112.

[49] Fama, E. F., French, K. R., *Common Risk Factors in the Returns on Stocks and Bonds*, Journal of Financial Economics, 1993, **33**(1), P.3-56, Doi:10.1016/0304-405x(93)90023-5.

[50] Jegadeesh, N., Titman, S., *Returns to buying winners and selling losers: implications for stock market efficiency*, Journal of Finance, 1993, **48**, P. 65-91, Doi:10.1111/j.1540-6261.1993.tb04702.x.

[51] Black, F., Scholes M., *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, 1973, **81**, P. 637–654, Doi:10.1086/260062.

[52] Fama, E. F., French, K. R., *Incremental variables and the investment opportunity set*, Journal of Financial Economics, 2015, **117**(3), P.470-488, Doi:10.1016/j.jfineco.2015.05.001.

[53] Chiah, M., Chai, D., Zhong, A., Li, S., *A better model? An empirical investigation of the Fama-French five-factor model in Australia*, International Review of Finance, 2016, **16**(4), P.595-638, Doi:10.1111/ irfi.1209 9.

[54] Kubota, K., Takehara, H., *Does the Fama and French five-factor model work well in Japan?*, International Review of Finance, 2018, **18**(1), P.137-146, Doi:10.1111/irfi.12126.

[55] Campbell, J. Y., *Understanding risk and return*, Journal of Political Economy, 1996, **104**(2), P.298-345, Doi:10.1086/262026.

[56] Kim, D., Kim, T. S., Min, B. K., *Future labor income growth and the cross-section of equity returns*. Journal of Banking and Finance, 2011, **35**(1), P.67-81, Doi:10.1016/j.jbankfin.2010.07.014.

[57] Kuehn, L.-A., Simutin, M., Wang, J. J., *A labor capital asset pricing model*, The Journal of Finance, 2017, **72**(5), P.2131-2178, Doi:10. 1111/jofi.12504.

[58] Racicot, F.-E., and Rentz, W. F., *Testing FamaeFrench's new five-factor asset pricing model: Evidence from robust instruments*, Applied Economics Letters, 2017, **23**(6), P.1-5, Doi:10.1080/13504851.2015.1080798

[59] Khani Farahani, A., Sheshmani, M., Mohades, A., *Studying the Expected Returns Based on Carhart Model Com-pared to CAPM Model and Implicit Capital Cost Model Based on Cash and Capital Flow of Growth and Value Stocks*, Advances in Mathematical Finance and Applications, 2017, **2**(4), P. 61-79. Doi:10.22034/AMFA. 2017.536267.

[60] Matevž, S., Lončarski, I., *Multi-factor asset pricing models: Factor construction choices and the revisit of pricing factors*, Journal of International Financial Markets, Institutions and Money, 2018, **55**, P.65-80, Doi:10. 1016/j.intfin.2018.02.006.

[61] Cao, V. N., Gray, P., Zhong, A., *Investment-related anomalies in Australia: Evidence and explanations*, International Review of Financial Analysis, 2019, **61**(C) P.97-109. Doi:10.1016/j.irfa.2018.10.007.

[62] Ramzi Radchobeh, Z., Rezazadeh, J., Kazemi, H., *Ambiguity Theory and Asset Pricing: Empirical Evidence from Tehran Stock Exchange*, Advances in Mathematical Finance and Applications, **3**(4), P.101-114, 2018, Doi: 10.22034/AMFA.2019.579558.1143.

[63] Ando, T., Bai, J., *Asset pricing with a general multifactor structure*, Journal of Financial Econometrics, 2014, **13**(3), P.556-604, Doi:10.1093/jjfinec/nbu026.

[64] Takács, G., *Gas Lift Manual*, Pennwell Books, 2005.

[65] Francis, B., *Low-Dimensional Geometry: from Euclidean Surfaces to Hyperbolic Knots*, AMS Bookstore, 2009, Doi:10.1090/stml/049.

[66] Stock, J.H., Watson, M.W., *Forecasting using principal components from a large number of predictors*, Journal of the American Statistical Association, 2002, **97**, P.1167–1179, Doi:10.1198/016214502388618960.

[67] Donoho, D.L., Jonhstone, I.M., *Ideal spatial adaptation by wavelet shrinkage*, Biometrika, 1994, **81**, P. 425–455, Doi:10.1093/biomet/81.3.425.

[68] Bai, J., *Panel data models with interactive fixed effects*, Econometrica, 2009, **77**, P.1229–1279, Doi:10. 398 2/ecta6135.

[69] Morovati Sharifabadi, A., Golshan, M., *Prediction of stock returns using artificial neural networks and multivariate regression (Case study: Companies accepted in Tehran Stock Exchange)*, Second National Conference on Accounting, Financial Management and Investment, Gorgan, Scientific and Professional Association of Golestan Managers and Accountants, 2013, (in Persian).

[70] Kashanipour, M., Salehnejad, H., Rezaei, A., Yousefi Manesh, D., *The effect of variations of the stock price fluctuation rate on stock returns and trade volume of companies accepted in Tehran Stock Exchange*, Journal of Accounting Knowledge, 2013, **2**(12), P.155-174, (in Persian).

[71] Mashayekhi, B., Panahi, D., *The Investigation of Relationship between Financial Ratios and Stock Return: Income Smoother Firms Versus Non-Income Smoother ones*, Financial Research Journal, 200۸, **9**(۳), (in Persian).

[72] Makian, N., Mousavi, F. S., *Prediction of stock prices in Pars Oil Products Company using Neural Networks and Regression Method*, Economic Modeling, 2012, **6**(18), P.105-121, (in Persian).

[73] Monajemi, A. H., Abzari, M. and Rayati, A. R., *Prediction of stock prices in the stock market using fuzzy neural network and genetic algorithms and its comparison with Artificial Neural Network*, Quantitative Economics Quarterly, 2014, **6**(3), 26-41, (in Persian).

[74] Mehrara, A., Atf, Z., Asgari, Z., *Data mining techniques in stock price prediction*, National Conference on Accounting, Finance, and Investment Management, Gorgan, Golestan University of Applied Sciences, 2012, (in Persian).