Modeling and forecasting abnormal stock returns using the nonlinear Grey Bernoulli model

Bahar Doryab and Mahdi Salehi
Ferdowsi University of Mashhad, Mashhad, Iran

Received 28 June 2017
Revised 23 September 2017
Accepted 25 September 2017

Abstract

Purpose – This study aims to use gray models to predict abnormal stock returns.

Design/methodology/approach – Data are collected from listed companies in the Tehran Stock Exchange during 2005-2015. The analyses portray three models, namely, the gray model, the nonlinear gray Bernoulli model and the Nash nonlinear gray Bernoulli model.

Findings – Results show that the Nash nonlinear gray Bernoulli model can predict abnormal stock returns that are defined by conditions other than gray models which predict increases, and then after checking regression models, the Bernoulli regression model is defined, which gives higher accuracy and fewer errors than the other two models.

Originality/value – The stock market is one of the most important markets, which is influenced by several factors. Thus, accurate and reliable techniques are necessary to help investors and consumers find detailed and exact ways to predict the stock market.

Keywords Abnormal returns, Gray theory, Nash nonlinear gray Bernoulli model, Nonlinear gray Bernoulli model

Paper type Research paper

Introduction

Practically, there are various methods for measuring the abnormal returns[1]. These methods are usually different, depending on the expected return measure used. Expected or normal return is the return on equity after excluding the expected event. Expected returns in event studies are either estimated by using patterns such as pattern arbitrage in the market or easily measured as the average market return (Binder, 1998). There has been a substantial increase in the use of nonlinear models, as compared with the linear ones, in recent literature. Several studies show that the nonlinear models have higher estimation power than linear the ones and these are capable of modeling the behavior of the efficiency (Abbasi and Bagheri, 2011). The gray method is considered as one of the prediction methods that provides a relatively accurate prediction with less information. In other words, while other forecasting models such as neural models use a high volume of data, the gray model merely requires the data of previous years (Khajavi *et al.*, 2012). Accordingly, this study aims to identify the

Journal of Economics, Finance and Administrative Science
Vol. 23 No. 44, 2018
pp. 95-112
Emerald Publishing Limited
2077-1886
DOI 10.1108/JEFAS-06-2017-0075

behavior of stock returns in the current year in comparison to its behavior in previous year by applying the gray model.

Literature review

Stock returns have information contents and *per se* have encouraged most of the real and potential investors to perform financial analysis. For this purpose, a variety of models such as the single-factor model of CAPM have been used in the literature. Recent findings suggest that the CAPM has the ability to predict stock returns. This is evidenced by the fact that the information available to the public may cause some abnormal returns (Mashayekhi *et al.*, 2009). The information used to create efficient portfolios are the ratio of stock earnings to price, cash flow ratio to price, operating cash flow to price ratio, the ratio of book value to market value of equity and accrual and its components, which are known in the literature as financial statement problems for consumers. If abnormalities are defined on the basis of risk, then abnormal return is rather something more than a mere risk that a researcher fails to identify and measure. If so, the data obtained do not indicate deviations but deficiency from the market efficiency, which has an impact on quantifying asset price risks (Beneish and Vargus, 2002). However, if the abnormalities are based on profit, it is likely that investors suffer from mispricing. The latter is in apparent contradiction to the market efficiency hypothesis, primarily because investors overestimate or underestimate the future stock earnings. On the other hand, while the abnormal return stemming from risk is persistent, the abnormal return gained from investors' mispricing is less stable or more open to change (Chan *et al.*, 2006).

Abnormal return is the difference between the actual and the expected stock returns. For the most part, using monthly simulation results of event studies for measuring the abnormal returns is highly desirable (Brown and Warner, 1980). Raei and Chavoshi (2003) analyzed the predictability of the stock return in the Tehran Stock Exchange using the artificial neutral network and multi-factor model. They indicated that the stock return in the stock market is predictable and influenced by macroeconomic variables. Furthermore, the multi-index model is capable of predicting the stock return using the macroeconomic variables. However, the artificial neural network outperformed the other model and lessened the prediction error.

Using the artificial neural networks along with three stock indexes, the volume of shares traded and the daily stock price, Namazi and Kiamehr (2008) predicted that the time-series behavior of daily stock returns is not a random process. The authors demonstrate that the artificial neural networks are capable of predicting daily stock returns with relatively acceptable prediction error. Zinedine Zadeh (2011) examined the precision of the top 50 stock returns traded in the Tehran Stock Exchange by using the gray prediction and time-series models, and argued that, all other past and future data being equal, the prediction accuracy of the gray model is higher at 0.5 margin of error. Bozorgasl and Sahebqarani (2013) predicted the abnormal returns by using simulated portfolios based on standardized unexpected earnings variables and the earning announcement dates for each industry and for 100 firms during a period of five years. The results indicate that unexpected earnings lead to abnormal returns. Nevertheless, the relationship between these variables is weakened over a fiscal year. Using the hybrid gray model for a sample of telecommunications companies, Ping and Yang (2004) argued that the gray prediction model better predicts the performance of the companies operating in the telecommunications industry owing to the complex and adverse environment prevailing in the industry. Pullets and Wilson (2010) examined the relationship between correlation and average stock market returns and found that changes in the variance of stock market returns can be associated with the cumulative risk and abnormal stock returns. They also showed that the four-month abnormal return is predictable by using average correlation between the daily stock market returns. Using the simulation methods for abnormal returns and size effect based on statistics market model and the efficiency of the modified technique of market, Kothari and Wasley (1989) concluded that abnormal returns in large companies are more than those of the small companies. Bommer *et al.* (1991) were able to explain the increase of unusual returns by using the event study methodology under development variance and with a simple adjustment in artificially cross-sectional statistics and simulation of an event.

On the basis of the previously mentioned literature, it is observed that the trading halt is a prevalent issue in small and newly established stock markets. For instance, Tehran Stock exchange is a typical example of a market which experiences such interruptions at regular intervals during each fiscal year. This has prevented researchers from measuring abnormal stock returns and estimating market model parameters.

The gray theory is one of the leading methods in the mathematical analysis of systems equipped with tentative data. This method explores the correlation between the components of a system and reference series. The gray theory, which was first developed by Deng (1982), is a very effective method used for explaining the uncertain and incomplete information as well as discrete data. The name of the system has been chosen on the basis of the color of the research topic. In this system, the terms "black," "white" and "gray" stand for unspecified information, specified information and semi-specified information, respectively (Liu and Lin, 2006).

Dong *et al.* (2006) in their study entitled "a Grey decision-making for selecting the supplier" attempted to propose a new approach to explain the multi-variable decision-making issues under uncertainty conditions and incomplete information by using the concept of gray possible degree. Wang (2013) used a new method of optimization of nonlinear gray Bernoulli model to predict the main economic indicators of high-tech companies in China. The results indicate that the optimized model could provide companies with appropriate data and with a guidance for future growth and development.

Gray prediction models use differential equations to explain an unspecified system with low data. These prediction models are more suitable for static data smoothing. Each gray system is defined as a series of gray values, gray equations and gray matrixes, in which the values play the role of the cells of the system. Gray values can be defined as numbers with uncertain information. For example, the ranking criteria in decision-making can be expressed numerically in the form of linguistic variables. These numerical intervals may include uncertain information. Furthermore, it is stated that gray value refers to a number whose exact value is indefinite, but its intervals are known. In practice, the gray value is generally expressed as an interval and a series of numbers (Dong *et al.*, 2006).

Hesin and Chen (2014) examined the evolution of gray forecasting and its application in management and engineering, and provided some satisfactory and useful results. Their findings suggest three solutions to predict the stock market more accurately by using the gray theory:

(1) fixed stock price;

(2) finding the limitation of maximum and minimum stock price in each business day; and

(3) increasing the stock price inconsistently.

Campbell and Wiernik (2015) studied the modeling of the limited data for production prediction through the gray theory. They argued that one of the main problems of first production stage is data prediction. Producers need sufficient knowledge to reduce the production price. The authors used the gray theory because of their limited sample. Their sample included corporate managers and decision-makers to assist them in developing a modeling method. Their empirical results indicate that the errors of modeling can be reduced using the gray theory and it improves prediction results by using the limited data.

Hypotheses

H1. Nonlinear gray Bernoulli model can predict abnormal stock returns better than the gray model.

H2. Nash nonlinear gray Bernoulli model can predict abnormal stock returns better than the Bernoulli gray model.

H3. Nash nonlinear gray Bernoulli model can predict abnormal stock returns better than the gray model.

Research methodology

In the present study, we use panel data technique based on synthetic data sets. In other words, two sets of time-series and cross-section data are combined. Our hypotheses are tested using the Matlab_{2013}, Spss_{20}, Eviews_{7} and Minitab_{16} software. The statistical population of the study includes all listed companies in the Tehran Stock Exchange during 2005-2015.

On the basis of the type of research and industry conditions, 100 listed companies are selected from the top ten industries. In general, gray systems use three methods to estimate the accuracy of the model as follows:

(1) *Residual analysis*: This is conducted by comparing the model values and values that are gathered practically. In this respect, we conduct a point-by-point analysis.

(2) *Evaluation of randomness*: This is conducted by comparing the graph of model values with the graph of frequencies forming the model.

(3) *Residual distribution*: This is conducted by analyzing the statistical properties of the residual distribution (Liu and Lin, 2006).

The GM model is widely used in the gray theory systems, and accordingly is the basis of all other gray models. The original time-series $x^{(0)}$ with "n" observations is as follows:

$$x^{(0)} = \left\{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\right\}$$

According to the following equation, the prediction function model for the gray Bernoulli nonlinear model is defined as:

$$\hat{x}^{(0)}(i+1) = \left[\left(\hat{x}^{(0)}(1) - \frac{b}{a}\right)e^{-a(1-n)i} + \frac{b}{s}\right]^{1/(1-n)} \qquad n \neq 1 \ , k = 1, 2, 3, \ldots.$$

where "i" is a point in time, "a" is the improvement coefficient and "b" is the gray control coefficient.

The Nash nonlinear gray Bernoulli model is the augmented version of the gray Bernoulli nonlinear model. Therefore, we define the most suitable equation based on the gray Bernoulli nonlinear model as follows (Chun *et al.*, 2010):

$$\begin{array}{c} Min \ \ \varepsilon \left(\ avg\left(n, p|x^{(0)}\right)\right) \\ p \in [0,1] \quad , \quad n \in \mathrm{R}^2 \end{array}$$

From this stage, we explain our models based on research hypotheses. First, we measure abnormal returns in each model by using the gray models. Next, we develop a regression model for each hypothesis and define explanatory variables. Then, a detailed review of research regression models and variable measurement follows.

Regression model of *H1*

$$\begin{aligned} GM_{i,t} * ARPE_{i,t} = \mathrm{a}_0 + \beta_1 NGBM_{i,t} * ARPE_{i,t} + \beta_2 XIM_{i,t} + \beta_3 PX_{i,t} \\ + \beta_4 Z(M)_{i,t} + \beta_5 IIP(K)_{i,t} + \beta_6 RMB(KX)_{i,t} + \ \varepsilon_{i,t} \end{aligned}$$

We include three variables in our model to explain the prediction power of each model in predicting abnormal stock returns, and also examine the accuracy of each model. These include the gray model power in predicting abnormal stock returns, the gray Bernoulli nonlinear model power in predicting abnormal stock returns and Nash nonlinear gray Bernoulli model power in predicting abnormal stock returns in measurement of research variables as follows:

- $GM_{i,t} \times ARPE_{i,t}$: The gray model power in predicting abnormal stock returns of firm i at time t.
- $NGBM_{i,t} \times ARPE_{i,t}$: The nonlinear gray Bernoulli model power in predicting abnormal stock returns of firm i at time t.
- $XIM_{i,t}$: Random prediction errors of firm i at time t.
- $PX_{i,t}$: Industry index prediction error of firm i at time t.
- $Z(M)_{i,t}$: Non-random mean square error of firm i at time t.
- $IIP(K)_{i,t}$: Return on risk volatilities of firm i at time t.
- $RMB(KX)_{i,t}$: Unforeseen risks arising from fluctuations in stock return of firm i at time t.

Regression model of *H2*

$$\begin{aligned} NGBM_{i,t} * ARPE_{i,t} = \mathrm{a}_0 + \beta_1 NNGBM_{i,t} * ARPE_{i,t} + \beta_2 XIM_{i,t} \\ + \beta_3 PX_{\mathrm{i,t}} + \beta_4 Z(M)_{i,t} + \beta_5 IIP(K)_{i,t} + \beta_6 RMB(KX)_{i,t} + \ \varepsilon_{i,t} \end{aligned}$$

The second regression model includes all explanatory variables used in the first model except the following variable:

$NNGBM_{i,t} \times ARPE_{i,t}$: The Nash nonlinear gray Bernoulli model power in predicting the abnormal stock returns of firm i at time t.

Regression model of *H3*

$$GM_{i,t} * ARPE_{i,t} = a_0 + \beta_1 NNGBM_{i,t} * ARPE_{i,t} + \beta_2 \ XIM_{i,t} + \beta_3 PX_{\mathrm{i,t}} + \beta_4 Z(M)_{i,t} + \beta_5 IIP(K)_{i,t} + \beta_6 RMB(KX)_{i,t} + \varepsilon_{i,t}$$

The third regression model includes all the variables used in the earlier models. In this model, the abnormal stock return is calculated as the difference between firm's actual return and its market return. We use the following equation, which includes price index and cash return, to calculate the abnormal stock return:

$$AR\ it = R\ it - R\ mt$$

$$R\ it = \frac{p1 - \ p0 + Dps + ((p1 - 1000) \times \alpha) \times (p1 \times \beta)}{P0}$$

$$R\ mt = \frac{I\ mt - \ I\ mo}{I\ mo}$$

where:

$ARit$ = abnormal return of stock i at time t;
Rit = stock return of firm i at time t;
$p1$ = stock price at the end of the fiscal year;
$p0$ = stock price at the beginning of the fiscal year;
Dps = gross dividends per share;
α = percentage increase in firm's equity from receivables and cash;
β = increase in firm's equity from retained earnings and reserves;
$R\,mt$ = cash return on equity and stock price index at time t;
$I\,mt$ = stock index at the beginning of time t; and
$I\,mo$ = stock index at the end of time t.

It is noteworthy that we use the final figure of the abnormal stock return of each listed company that is already provided in the TSE library and its corresponding software.

Various definitions of the concept of investment risk are given in the literature. However, investment risk is broadly defined as the risk of "investment return volatility." In other words, the more return on investment changes, the riskier the investment. The proxy used to measure the changes of return rate is the standard deviation, which is as follows:

$$\sigma = \sum_{i=1}^{n} \left(R_i - \overline{R}\right)^2 \ P_i$$

where:

σ = standard deviation (investment risk proxy);
R_i = return on assets in firm I;
$\overline{R}$ = average return on assets; and
P_{i} = probability of occurrence i

The stock risk is recognized by stock price or return volatility (Verchenco, 2002), primarily because stock return volatility can be a sign of uncertainty about stock future inflows. According to the volatility feedback theory, an increase in stock return volatility leads to the reduction of stock returns. This theory was initially introduced by Pindyck (1984). According to the theory of volatility feedback, the increase in stock return volatility gives rise to the expected returns on equity. In other words, investors are willing to take more risk in return for higher return. Increasing expected stock returns lead to increased future cash flows. Furthermore, the increase in the discount rate leads to lower present value of future cash flows, which subsequently reduces the stock price and returns. Following Bhattacharyya and Thomas's study (2006), we calculate the industry index, return on firm's risk volatility and the unpredictable risk arising from the stock return volatility of the Tehran Stock Exchange as follows:

$$Y_t = \frac{TI_{ti} - TI_{t0}}{TI_{t0}}$$

where:

Y_t = return on equity;

TI_{ti} = industry index of Tehran Stock Exchange at the end of the fiscal year; and

TI_{t0} = industry index of Tehran Stock Exchange at the beginning of the fiscal year.

We use the following equation to capture the return actual volatilities:

$$\sigma_m^2 = \frac{1}{N_j} \left(\sum_{j=1}^{N_j} y_{j,t} - y_t^m \right)^2.$$

where:

$y_{j,t}$ = stock return on day j and period t;

y_t^m = average stock returns during period t; and

N_j = number of working days in the stock market over a fiscal year.

We use the following equation to study the effect of unpredictable risk arising from the return volatility on the stock return:

$$\sigma_u^2 = \left(\sigma_m^2 - \sigma_t^2 \right)$$

In the above equation, σ_t^2 is the return on fluctuations of the company's risk and σ_u^2 is the unpredictable risk arising from the return fluctuations. Finally, the effect of unpredictable risks arising from the volatility of returns on stock returns is estimated by using the following equation:

$$Y_t = f\left(\sigma_u^2\right)$$

If the assumptions of the classical linear regression model are met, we can use ordinary least squares to estimate the abovementioned equation.

Research findings and analysis

The nonlinear gray Bernoulli model power in predicting abnormal stock returns is positively correlated with the power of Nash nonlinear gray Bernoulli model in predicting abnormal stock returns, predicting random errors and predicting errors of the industry index.It is also negatively correlated with average squared non-random error and return on company's risk volatility. The Nash nonlinear gray Bernoulli model power in predicting abnormal stock returns is also positively correlated with random prediction errors and prediction errors of the industry index and negatively correlated with average squared non-random error and return on company's risk volatility. Random prediction errors display a positive and significant correlation with prediction errors of the industry index and return on company's risk volatility. Furthermore, there is a significant correlation between prediction errors of the industry index and average squared non-random error and return on company's risk volatility. The prediction error of industry index indicates a significantly negative correlation with average squared non-random error and return on company's risk volatility. Average squared non-random error is positively correlated with unpredictable risk arising from return volatility.

As there is a significant and positive correlation between the nonlinear gray Bernoulli model power in predicting abnormal stock returns and Nash nonlinear gray Bernoulli model in predicting abnormal stock returns, we included these variables in different models to deal with the multicollinearity concern. Regarding the reminding variables, multicollinearity is of no concern, because its correlations are not statistically significant (Table I).

Selecting the model

In general, a measure is required to assist us in evaluating and choosing the most appropriate model among other time-series models. Furthermore, there is always an error in all prediction models due to a degree of uncertainty prevailing in these models. The prediction error is usually caused by the irregularity of one or several time-series factors such as trends, fluctuations or periods that remained unconsidered. In general, the closer the (y_t) actual value to the ($\hat{y}_t$) predicted value, the higher is the accuracy of model predictions. The accuracy of a model is evaluated by the level of its prediction error:

$$e_{\mathrm{t}} = y_{\mathrm{t}} - \hat{y}_t$$

As prediction errors are accidental and either positive or negative, it is likely that these cancel out each other's impacts and consequently make the error calculation difficult. To solve this problem, we use the average squared errors equation (MSE) as follows:

$$MSE = \frac{\sum_{t=1}^{n} (y_{\mathrm{t}} - \hat{y}_t)^2}{n}$$

Hypothesis testing

H1 of the present study aims to compare the nonlinear gray Bernoulli model power in predicting abnormal stock returns with that of the gray model. Therefore, *H1* presented in null form is as follows:

$H1_0$. In comparison to the gray model, the nonlinear gray Bernoulli model is not powerful enough in terms of predicting abnormal stock returns.

Table I. Matrix of Pearson correlation coefficients between variables

Variable	N	K-S	Significance
Gray model power in predicting abnormal stock returns	1,000	0.699	0.735
Gray nonlinear Bernoulli model power in predicting abnormal stock returns	1,000	0.735	0.652

We examine *H1* by using panel data analysis. In this respect, we conduct Chow test to choose the appropriate model between panel data and pooled method. Next, we use Hausman test to determine the appropriate estimator between fix effects and random effects estimators. The results of these tests are presented in Table II.

According to the results of the Chow test and its p-value (0.0000), the null hypothesis of this test is rejected at the 95 per cent confidence level. This implies that the panel data model is more appropriate than pooled estimator. Likewise, the results of the Hausman test and its p-value (0.0003) do not provide supporting evidence for the null hypothesis of this test at 95 per cent confidence level, and consequently the fixed effects estimator is chosen as appropriate.

To test the normality of the error terms, various tests are available. One of these tests is the Jarque–Bera test, which is used in this research. Test results indicate that the residuals arising from the research model are normally distributed at 95 per cent of confidence level, because the probability value associated with this test (0.8254) is more than 0.05 margin of error. In addition to the aforementioned test, we conduct the Breusch–Pagan test to check the homogeneity of variance in a linear regression. As indicated in Table III, the results of this test suggest that the null hypothesis of the test is rejected owing to the sig value ($0.0001 < 0.05$). That is, homogeneity of variance in our linear regression is of concern. To address this problem, we use the generalized least squares (GLS) estimator.

We also conduct the Durbin–Watson test (DW) to check for residual correlations (autocorrelation). As the DW statistic (2.04) is between 1.5 and 2.5, it can be concluded that there is no autocorrelation between residuals. Finally, we conduct the Ramsey test to examine whether the model is a linear relationship and whether it is appropriately explained in terms of linearity. The null hypothesis of this test is rejected as the p-value (0.0860) is more than the 0.05 margin of error, implying that the linearity of our model is confirmed. A summary of above specification tests is presented in Table III.

On the basis of the results of the Chow and Hausman tests as well as the specification tests of the classical regression model explained earlier, we estimate the first model by using panel data analysis and fixed effects estimator. The estimation results are presented in Table IV. The estimated model is as follows:

$$GM_{i,t} * ARPE_{i,t} = 0.8800 + 0.2966\ NGBM_{i,t} * ARPE_{i,t} + 0.1587\ XIM_{i,t} + 0.0426\ PX_{i,t} \\ + 0.0062\ Z(M)_{i,t} - 0.0498\ IIP(K)_{i,t} + 0.0492\ RMB(KX)_{i,t} + \varepsilon_{i,t}$$

Test	N	Statistic type	Statistic value	df	Significance
Chow	1,000	F	8.6999	99.394	0.0000
Hausman	1,000	χ^2	256,536	6	0.0003

Table II. Results of Chow and Hausman tests for Model (1)

Ramsey		Durbin–Watson	Breusch–Pagan		Jarque–Bera	
F	p-value	D	F	p-value	χ^2	p-value
5.1642	0.0860	2.04	4.7284	0.0001	1.9455	0.8254

Table III. The results of specification tests in regression model (1)

Table IV.
Estimation results of the first model using fixed effects estimator

Variable	2010-2014 Significance	t	Coefficient	2005-2009 Significance	t	Coefficient	Result
Dependent variable: gray model power in predicting abnormal stock returns (n = 1,000)							
Constant	0.0000	24.437	0.8800	0.0000	0.454	0.4211	Positive, significant
$NGBM_{i,t} \times ARPE_{i,t}$	0.0000	5.9576	0.2966	0.0344	2.123	0.3245	Positive, significant
$XIM_{i,t}$	0.0000	8.1498	0.1587	0.0146	2.452	0.1631	Positive, significant
$PX_{i,t}$	0.0000	−4.5084	−0.0426	0.0002	3.717	0.1177	Significant, negative
$Z(M)_{i,t}$	0.6435	0.4632	0.0062	0.5999	0.524	0.0218	Insignificant
$\prod P(K)_{i,t}$	0.0000	−6.0195	−0.0498	0.0192	−2.351	−0.0664	Significant, negative
$RMB(KX)_{i,t}$	0.3318	0.9589	0.0492	0.0011	3.281	0.2297	Insignificant
R^2							0.7565
F-statistic significance							11.6587 (0.0000)

Based on the F-statistic of the model and its significance value ($0.00 < 0.05$), the overall model is significant at 95 per cent confidence level. The R^2 squared coefficient of the overall model (0.7565) also indicates that 75.65 per cent of the gray model power in predicting abnormal stock returns is explained by the explanatory variables included in the model. With respect to explanatory variables shown in Table IV, it can be observed that there is a significant relationship between the gray model and the nonlinear gray Bernoulli model in predicting abnormal stock returns. In other words, the significance value of the nonlinear gray Bernoulli model power in predicting abnormal stock returns (0.000) is less than the margin of error of 5 per cent, suggesting that this variable is statistically significant. Furthermore, the coefficient on this variable (0.2966) is positive and indicates that the nonlinear gray Bernoulli model power in predicting abnormal stock returns is positively associated with abnormal stock returns prediction. More specifically, one unit increase in abnormal stock return prediction leads to 0.2966 increase in the nonlinear gray Bernoulli model power in predicting abnormal stock returns.

H2 of the present study attempts to compare the Nash nonlinear gray Bernoulli model power in predicting abnormal stock returns with that of the nonlinear gray Bernoulli model. Therefore, *H2* is presented in null form as follows:

$H2_0$. In comparison to the nonlinear gray Bernoulli model, the Nash nonlinear gray Bernoulli model is not powerful enough in terms of predicting abnormal stock returns.

We test *H2* by using the panel data analysis. In this respect, we conduct the Chow test to choose the appropriate model between panel data and pooled method. Next, we use Hausman test to determine the appropriate estimator between fixed effects and random effects estimators. The results of these tests are presented in Table V.

According to the results of the Chow test and its p-value (0.0000), the null hypothesis of this test is rejected at the 95 per cent confidence level. This implies that the panel data model is more appropriate than pooled estimator. Likewise, the results of the Hausman test and its p-value (0.0222) do not provide supporting evidence for the null hypothesis of this test at 95 per cent confidence level and consequently the fixed effects estimator is chosen as appropriate.

To test the normality of the error terms, we use the Jarque–Bera test. Test results indicate that the residuals arising from the second model are normally distributed at 95 per cent of confidence level, because the probability value associated with this test (0.3265) is more than 0.05 margin of error. In addition to the aforementioned tests, we conduct the Breusch–Pagan test to check the homogeneity of variance in a linear regression. As indicated in Table VI, the results of this test suggest that the null hypothesis of the test is rejected owing to the sig

Test	Statistic type	df	Statistic value	Significance
Chow	F	99.394	3.5644	0.0000
Hausman	χ^2	6	0.0222	0.0222

Table V. Results of Chow and Hausman tests for Model (2)

Ramsey		Durbin–Watson	Breusch–Pagan		Jarque–Bera	
F	p-value	D	F	p-value	χ^2	p-value
12.1968	0.2154	2.38	15.8448	0.0000	1.4454	0.3265

Table VI. The results of specification tests in regression model (2) value (0.0001 < 0.05). That is, homogeneity of variance in our linear regression is of concern. To address this problem, we use the GLS estimator.

We also conduct the Durbin–Watson test (DW) so as to check for residual correlations (autocorrelation). As the Durbin Watson statistic (2.38) is between 1.5 and 2.5, it can be concluded that there is no autocorrelation between residuals. Finally, we conduct the Ramsey test to examine whether the model is a linear relationship and whether it is appropriately explained in terms of linearity. The null hypothesis of this test is rejected as the p-value (0.2154) is more than the 0.05 margin of error, implying that the linearity of our model is confirmed. A summary of above specification tests are presented in Table VI.

On the basis of the results of the Chow and Hausman tests, as well as the specification tests of the classical regression model explained earlier, we estimate the second model by using panel data analysis and fixed effects estimator. The estimation results are presented in Table VII. The estimated model is as follows:

$$NGBM_{i,t} * ARPE_{i,t} = 0.0176 + 0.0791\ NNGBM_{i,t} * ARPE_{i,t} + 0.1001\, XIM_{i,t} - 0.0019 PX_{i,t} + 0.0011\ \mathrm{Z}(M)_{i,t} + 0.0008\, IIP(K)_{i,t} + 0.0881\, RMB(KX)_{i,t} + \varepsilon_{i,t}$$

Based on the F-statistic of the model and its significance value (0.00 < 0.05), the overall model is significant at 95 per cent confidence level. The R^2 coefficient of the overall model (0.7082) also indicates that 70.82 per cent of the nonlinear gray Bernoulli model power in predicting abnormal stock returns is explained by the explanatory variables included in the model. With respect to explanatory variables shown in Table VII, it can be observed that there is a significant relationship between the Nash nonlinear gray Bernoulli model and the nonlinear gray Bernoulli model in predicting abnormal stock returns. In other words, the significance value of the Nash nonlinear gray Bernoulli model power in predicting abnormal stock returns (0.000) is less than the margin of error of 5 per cent, suggesting that this variable is statistically significant. Furthermore, the coefficient on this variable (0.0791) is positive and indicates that the Nash nonlinear gray Bernoulli model power in predicting abnormal stock returns is positively associated with the nonlinear gray Bernoulli model power in predicting abnormal stock returns. More specifically, one unit increase in prediction power of the Nash nonlinear gray Bernoulli model leads to 0.0791 increase in the nonlinear gray Bernoulli model power in predicting abnormal stock returns.

H3 of the present study attempts to compare the Nash nonlinear gray Bernoulli model power in predicting abnormal stock returns with that of the gray model. Therefore, this hypothesis is presented in null form as follows:

$H3_O$. In comparison to the gray model, the Nash nonlinear gray Bernoulli model is not powerful enough in terms of predicting abnormal stock returns.

We test *H3* by using panel data analysis. In this respect, we conduct the Chow test to choose the appropriate model between panel data and pooled method. Next, we use the Hausman test to determine the appropriate estimator between fix effects and random effects estimators. The results of these tests are presented in Table VIII.

According to the results of the Chow test and its p-value (0.0000), the null hypothesis of this test is rejected at the 95 per cent confidence level. This implies that the panel data model is more appropriate than pooled estimator. Likewise, the results of the Hausman test and its p-value (0.000 < 0.05) do not provide supporting evidence for the null hypothesis of this test at 95 per cent confidence level and consequently the fixed effects estimator is chosen as appropriate.

Table VII. Estimation results of the second model using fixed effects estimator

Variable	2010-2014 p-value	t	Coefficient	2005-2009 p-value	t	Coefficient	Result
Dependent variable: the nonlinear grey Bernoulli model power in predicting abnormal stock returns (n = 1,000)							
Constant	0.3817	−0.8757	−0.0176	0.0541	1.9320	0.0305	Insignificant
$NNGBM_{i,t} \times ARPE_{i,t}$	0.0000	6.4433	0.0791	0.0000	6.7518	0.0746	Positive, significant
$XIM_{i,t}$	0.0000	10.5108	0.1001	0.0000	20.2255	0.2555	Positive, significant
$PX_{i,t}$	0.7113	−0.3703	−0.0019	0.0296	−2.1824	−0.0114	Insignificant
$Z(M)_{i,t}$	0.8583	0.1786	0.0011	0.0011	−3.2791	−0.0231	Insignificant
$\prod P(K)_{i,t}$	0.8500	0.1892	0.0008	0.1751	1.3584	0.0054	Insignificant
$RMB(KX)_{i,t}$	0.0000	4.8279	0.0881	0.0017	−3.1542	−0.0405	Positive, significant
R^2							0.7082
F-statistic significance							9.1078 (0.0000)

To test the normality of the error terms, we use the Jarque–Bera test. Test results indicate that the residuals arising from the second model are normally distributed at 95 per cent of confidence level, because the probability value associated with this test (0.5122) is more than 0.05 margin of error. In addition to the aforementioned tests, we conduct the Breusch–Pagan test to check the homogeneity of variance in a linear regression. As indicated in Table IX, the results of this test suggest that the null hypothesis of the test is rejected owing to the sig value ($0.000 < 0.05$). That is, homogeneity of variance in our linear regression is of concern. To address this problem, we use the GLS estimator.

We also conduct the DW test so as to check for residual correlations (autocorrelation). As the DW statistic (2.34) is between 1.5 and 2.5, it can be concluded that there is no autocorrelation between the residuals. Finally, we conduct the Ramsey test to examine whether the model is a linear relationship and whether it is appropriately explained in terms of linearity. The null hypothesis of this test is rejected as the p-value (0.5864) is more than the 0.5 margin of error, implying that the linearity of our model is confirmed. A summary of above specification tests is presented in Table IX.

On the basis of the results of the Chow and Hausman tests as well as the specification tests of the classical regression model explained earlier, we estimate the second model by using panel data analysis and fixed effects estimator. The estimation results are presented in Table X. The estimated model is as follows:

$$
\begin{aligned}
GM_{i,t} * ARPE_{i,t} = 1.0420 + 1.9731\, NNGBM_{i,t} *\ ARPE_{i,t} + 0.0276\ XIM_{i,t} \\
+ 0.0656\ PX_{i,t} - 0.1751\ Z(M)_{i,t} + 0.0136\ IIP(K)_{i,t} \\
+ 0.1685\ RMB(KX)_{i,t} + \varepsilon_{i,t}
\end{aligned}
$$

On the basis of the F-statistic of the model and its significance value ($0.00 < 0.05$), the overall model is significant at 95 per cent confidence level. The R^2 coefficient of the overall model (0.3261) also indicates that 32.61 per cent of the gray model power in predicting abnormal stock returns is explained by the explanatory variables included in the model. With respect to explanatory variables shown in Table X, it can be observed that there is a significant relationship between the Nash nonlinear gray Bernoulli model and the gray model in predicting abnormal stock returns. In other words, the significance value of the Nash nonlinear gray Bernoulli model power in predicting abnormal stock returns (0.0230) is less than the margin of error of 5 per cent,

Table VIII.
Results of Chow and Hausman tests for Model (3)

Test	Statistic type	*df*	Statistic value	Significance
Chow	F	99.394	3.3978	0.0000
Hausman	χ^2	6	11,660.2176	0.0000

Table IX.
The results of specification tests in regression model (3)

Ramsey		Durbin–Watson	Breusch–Pagan		Jarque–Bera	
F	p-value	D	F	p-value	χ^2	p-value
0.5342	0.5864	2.34	66.0667	0.0000	1.6481	0.5122

Table X. Estimation results of the third model using fixed effects estimator

Variable	2010-2014 p-value	t	Coefficient	2005-2009 p-value	t	Coefficient	Result
Dependent variable: the nonlinear gray Bernoulli model power in predicting abnormal stock returns (n = 1,000)							
Constant	0.0210	1.5531	1.0420	0.0001	3.9484	0.2847	Insignificant
$NNGBM_{i,t} \times ARPE_{i,t}$	0.0230	2.2803	1.9731	0.0000	14.5474	0.7738	Positive, significant
$XIM_{i,t}$	0.3109	1.0142	0.0276	0.0000	−7.9850	−0.4374	Positive, significant
$PX_{i,t}$	0.1781	1.3485	0.0656	0.0801	1.7548	0.0443	Insignificant
$Z(M)_{i,t}$	0.0097	−2.595	−0.1751	0.5497	0.5986	0.0190	Insignificant
$\prod P(K)_{i,t}$	0.6645	0.4340	0.0136	0.0156	−2.4279	−0.0540	Insignificant
$RMB(KX)_{i,t}$	0.0017	3.1615	0.1685	0.0119	2.5256	0.1593	Positive, significant
R^2							0.3261
F-statistic significance							2.2605 (0.0000)

suggesting that this variable is statistically significant and *H3* is confirmed. Furthermore, the coefficient on this variable (1.9731) is positive and indicates that the Nash nonlinear gray Bernoulli model power in predicting abnormal stock returns is positively associated with gray model power in predicting abnormal stock returns. More specifically, one unit increase in prediction power of the Nash nonlinear gray Bernoulli model leads to 1.9731 increase in the nonlinear gray Bernoulli model power in predicting abnormal stock returns.

Having tested research hypotheses and found the most appropriate regression models, we use the Akaike information criterion to choose the best model. According to AIC values reported in Table XI, it can be concluded that the nonlinear gray Bernoulli model is the most suitable model in predicting abnormal stock returns.

Concluding remarks

Having captured the prediction error for each model along with the prediction error of 50 more active companies and industry index, we moved on to examine the prediction power of the models. According to the results of the study, we predict abnormal stock returns by using three models, the gray model, the nonlinear gray Bernoulli model and the Nash nonlinear gray Bernoulli model, among which the nonlinear gray Bernoulli model is found to be the most powerful. In this paper, we presented models for predicting abnormal stock returns with the lowest prediction errors. This was consistent with the primary goal of this research, i.e. providing accurate predictions by using minimum information. Our findings are consistent with the results of Kazemi *et al.* (2010), Wang and Hsu (2008), Huang and Jane (2009) and Mohammadi and Zeinodin Zade (2011) in terms of prediction accuracy of gray models as compared with other models. These findings are also consistent with those of Chun *et al.* (2010) and Wang (2013) in terms of the Nash nonlinear gray Bernoulli model's superiority over the gray and the nonlinear gray Bernoulli models (Table XII).

Table XI. Comparison of regression models using AIC

Regression model	AIC
1	1.3014
2	1.3014
3	1.3014

Table XII. Summary of research findings

No.	Hypotheses	Results and findings
H1	Nonlinear gray Bernoulli model can predict abnormal stock returns better than the gray model	Hypothesis supported: Chun *et al.* (2010) and Wang (2013)
H2	Nash nonlinear gray Bernoulli model can predict abnormal stock returns better than the Bernoulli gray	Hypothesis supported: Chun *et al.* (2010) and Wang (2013)
H3	Nash nonlinear gray Bernoulli model can predict abnormal stock returns better than the gray model	Hypothesis supported: Chun *et al.* (2010) and Wang (2013)

Note

1. Outlier's observations that are located farther from the other data and their values relative to other values in the data set are larger or smaller. Outliers can cause adverse effects such as increasing error variance, reducing test power, disrupting normal dispatch data and estimation of the parameters. Therefore, it is necessary that researchers take this into consideration.

References

Abbasi, E. and Bagheri, S. (2011), "Predict stock returns using nonlinear models and the role of turnover in improving the performance of these models", *Journal of Financial Research*, Vol. 13 No. 32, pp. 91-108.

Beneish, M.D. and Vargus, M.E. (2002), "Insider trading, earnings quality and accrual mispricing", *The Accounting Review*, Vol. 77 No. 4, pp. 755-791.

Bhattacharyya, N. and Thomas, R.G. (2006), "Why people choose negative expected return assets - An empirical examination of a utility theoretic explanation", Working paper, Federal Reserve Bank of St. Louis, available at: journals.plos.org/plosone/article?id=10.1371/journal.pone.0071966 (accessed 20 March 2015).

Binder, J. (1998), "The event study methodology since 1969", *Review of Quantitative Finance and Accounting*, Vol. 11 No. 2, pp. 111-137.

Bommer, E., Musumeci, J. and Poulsen, A. (1991), "Event-study methodology conditions of event induced variance", *Journal of Financial Economics*, Vol. 30 No. 2, pp. 253-272.

Bozorgasl, M. and Sahebqarani, A. (2013), "Abnormal returns based on the model-based forecasts and industry momentum profits in Tehran stock exchange", *Experimental Studies Quarterly of Financial Accounting*, Vol. 38, pp. 53-67.

Brown, S.J. and Warner, J.B. (1980), "Measuring security price performance", *Journal of Financial Economics*, Vol. 8 No. 3, pp. 205-258.

Campbell, J.P. and Wiernik, B.W. (2015), "The modeling and assessment of work performance", *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 2 No. 1, pp. 47-74.

Chan, K., Chan, L.K.C., Jegadeesh, N. and Lakonishok, J. (2006), "Earnings quality and stock returns", *The Journal of Business*, Vol. 79 No. 3, pp. 1041-1082.

Chun, C., Hsin, P. and Wu, C. (2010), "Forecasting Taiwan's major stock indices by the Nash nonlinear grey Bernoulli model", *Expert Systems with Applications*, Vol. 37 No. 12, pp. 7557-7562.

Deng, J. (1982), "Grey control system", *Journal of Huazhong University of Science and Technology*, Vol. 3, pp. 9-18.

Dong, G., Yamaguchi, D. and Nagai, M. (2006), "A grey-based decision making approach to the supplier selection problem", *Mathematical and Computer Modeling*, Vol. 46, pp. 573-581.

Hesin, P.H. and Chen, C.I. (2014), "The evolution of grey forecasting and its application on stock price prediction", *J Stock Forex Trad*, Vol. 3, p. 124.

Huang, K. and Jane, C. (2009), "A hybrid model for stock market forecasting and portfolio selection based on ARX, grey system and RS theories", *Expert Systems with Application*, Vol. 36 No. 3, pp. 5387-5392.

Kazemi, A., Moddarres, M., Mehregan, M.R. and Foruqi-Asl, A. (2010), "The forecast of transport sector energy demand by the Grey Markov Chain Model", *Conference on Energy Efficiency*.

Khajavi, S., Najafi, Z. and Zinedine Zadeh, S. (2012), "Applying the grey models and the development of a simple way to forecast free cash flow of listed companies in Tehran stock exchange", *Journal of Financial and Management Engineering Exchange*, Vol. 13, pp. 21-34.

Kothari, S.P. and Wasley, C.E. (1989), "Measuring security price performance in size-clustered samples", *The Accounting Review*, Vol. 64 No. 2, pp. 228-249.

Liu, S. and Lin, Y. (2006), *Grey Information*, Springer, London.

Mashayekhi, B., Fadaee Nejad, M. and Kalateh Rahmani, R. (2009), "Explanatory power of accruals (abnormal) stock returns in relation to behavior: Effects of corporate profits and the systematic risk (β)", *Journal of Accounting and Auditing*, Vol. 4, pp. 1-20.

Mohammadi, A. and Zeinodin Zade, S. (2011), "Appling grey forecasting method to forecast the portfolio's rate of return in stock market of Iran", *Australian Journal of Business and Management Research*, Vol. 1 No. 7, pp. 1-16.

Namazi, M. and Kiamehr, M. (2008), "Forecasting daily stock returns of listed companies in Tehran stock exchange using artificial neural networks", *Journal of Financial Research*, Vol. 6 No. 3, pp. 60-72.

Pindyck, R. (1984), "Risk, inflation, and the stock market", *American Accounting Review*, Vol. 74 No. 3, pp. 335-351.

Ping, Y. and Yang, H. (2004), "Using hybrid grey model to achieve revenue assurance of telecommunication companies", *Journal of Grey System*, Vol. 7, pp. 39-50.

Pullets, J.M. and Wilson, M. (2010), "Average correlation and stock market returns", *Journal of Financial Economics*, Vol. 96 No. 3, pp. 364-380.

Raei, R. and Chavoshi, K. (2003), "Forecasting of stock return in Tehran stock exchange: artificial neural networks and multi factors models", *J. Financial Study*, Vol. 15 No. 5, pp. 97-120.

Verchenco, O. (2002), "Determinants of stock market volatility dynamics", Working Paper, HEC University of Lausann.

Wang, C.-H. and Hsu, L. (2008), "Using genetic algorithms grey theory to forecast high technology industrial output", *Application Mathematics and Computation*, Vol. 195 No. 1, pp. 256-263.

Wang, Z. (2013), "An optimized Nash nonlinear grey Bernoulli model for forecasting the main economic indices of high technology enterprises in China", *Computers & Industrial Engineering*, Vol. 64 No. 0, pp. 780-787.

Zinedine Zadeh, S. (2011), "Predicting stock returns in Tehran Stock Exchange by using the Grey prediction model", MA thesis, Faculty of Economics, Management and Social Sciences, Shiraz University.

Corresponding author
Mahdi Salehi can be contacted at: mehdi.salehi@um.ac.ir