Study of Asian indexes by a newly derived dynamic model

Tsung-Jui Chiang-Lin, Yong-Shiuan Lee*, Tzong-Hann Shieh, Chien-Chang Yen, Shang-Yueh Tsai

1 Department of Finance, Minghsin University of Science and Technology, Hsinchu, Taiwan, 2 Department of Statistics, National Chengchi University, Taipei, Taiwan, 3 Department of Aerospace and Systems Engineering, Feng Chia University, Taichung, Taiwan, 4 Department of Mathematics, Fu-Jen Catholic University, New Taipei City, Taiwan, 5 Graduate Institute of Applied Physics, National Chengchi University, Taipei, Taiwan

* 99354501@nccu.edu.tw

Abstract

We take the stock prices as a dynamic system and characterize its movements by a newly derived dynamic model, called the new Price Reversion Model (nPRM), for which the solution is derived and carefully analyzed under different circumstances. We also develop a procedure of applying the nPRM to real daily closing prices of a stock index. This proposed procedure brings a different perspective to the study of stock prices based on thermodynamics, and the time varying coefficients in the nPRM offer economic meanings of the stock movements. More specifically, the average of smoothed historical data $A$ in the nPRM, analogous to the environment temperature in the Newton’s law of cooling, represent an implied equilibrium price. The heat transfer coefficient $\kappa$ is adapted to be either negative or positive, which illustrates the speed of convergence or divergence of stock prices, respectively. The empirical study of ten Asian stock indexes shows that the nPRM accurately characterizes and forecasts the market values.

Introduction

The stock index, a key variable reflecting the economic status in an area, is a popular research topic in macroeconomics. The market prices of an individual stock reflect the company’s existing value or future profit growth. A stock index, measured by the weighted mean prices of selected stocks in a stock market, represents the stock market’s performance and potential competitive ability. Investors who have different motives for participating in the market concern about the changes of a stock or a stock index’s market prices of different scales, intraday high-frequency data, daily closing prices, weekly or monthly summary, and etc. Accurately forecasting the stock movements could help the investors with decisions about investment strategy.

Whether the stock prices can be forecasted has been repeatedly discussed in finance [1, 2]. The efficient market hypothesis [3] forms a foundation of the theory that under what circumstances the stock prices is predictable. The efficient market hypothesis states that three types of
markets can be distinguished as the weak form, the semi-strong form, and the strong form depending on the level of market efficiency. The stock prices fully reflect all available market information in a strong efficient market, thus they are not predictable. The martingale depicts the stock movements in the strong efficient markets by the theory of random walks [4]. In contrast, many real markets are considered as the weak or the semi-strong efficient markets [2, 5]. Consequently, to forecast the prices in these markets is possible.

The analysis of the stock movements is based on either the stock prices or the stock returns. The choice between the stock prices and the returns mainly depends on the approach adopted. From the financial accounting perspective, fundamental analysis [6, 7] and technical analysis [8] are the two main intellectual traditions. The fundamental analysis evaluates the stocks by measuring their intrinsic values from overall economic and industrial conditions. The fundamental analysis assumes that the stock prices are linearly correlated to the predictors in the financial statement. As a result, the linear regression models are usually applied to forecast the stock prices. On the contrary, the technical analysis assumes that all known fundamentals are factored into prices. Hence, charting past fluctuations in the stock indexes allows the analyst to identify the patterns and trends. Technical analysis forecasts the market movements through popular techniques like chart analysis, cycle analysis and computerized technical trading systems [8].

Various statistical methods are prevalent in the history of financial academic researches. In econometric modeling, applying the autoregressive integrated moving average (ARIMA) models and vector autoregression (VAR) models is common. When using ARIMA and VAR models, the stationarity of the time series is required. Therefore, the stock returns instead of the stock prices are the primary research subject [9–11]. Financial ratios like the log dividend-price ratio [12, 13], the dividends [9], the price–earnings ratio [10, 13], the dividend yield [10, 14], the consumption-wealth ratio [15], the book-to-market ratio [10] serve as important predictors for asset returns in regression models, particularly VAR models. Although some good predictors in financial literature have restricted predictive ability for the equity premium due to model uncertainty and instability [11], combining individual forecasts produce out-of-sample gains [16]. Modeling stock returns reduces low-frequency information, the trend for example, from the perspective of signal processing. Therefore, forecasting the stock prices based on models constructed from detrended data may be less accurate in this sense.

The stock prices can be treated as signals. From the signal processing perspective, the forecasts can be constructed by decomposing the signals into features and putting them into artificial intelligence models. The artificial intelligence models applied to the stock prices in existing studies include support vector regression (SVR) [17], neural networks [18–22], tree-based models [23], etc. To applying the artificial intelligence models, acquiring features is an important step. The empirical mode decomposition (EMD) [17, 21], the ensemble empirical mode decomposition (EEMD) [24], and the complementary ensemble empirical mode decomposition (CEEMD) [22] are widely employed methods to decompose the stock prices, which are regarded as signals. The empirical results of US and UK stock indexes show that the artificial intelligence models are better than the ARIMA and GARCH models in terms of forecasting directional symmetry [21]. Furthermore, the CEEMD-PCA-LSTM model [22], which integrates the CEEMD, principal component analysis (PCA), and long short-term memory (LSTM) networks, outperforms naive recurrent neural networks (RNN) or LSTM networks in terms of forecasting accuracy for six selected stock indexes. Still, in a review study of the returns from financial perspective [25], the empirical evidences show that no single model out of statistical, artificial intelligence, frequency domain, and hybrid models under investigation could be applied uniformly to all markets. On the other hand, the direction prediction problem
is another task, which can be tackled by the tree-based models [23] to predict the stock price direction.

This paper considers a physical-mathematical method, the approach of the differential equations and the dynamic systems. From thermodynamics, we adopt the Newton’s law of cooling to describe the mean reverting of the stock index prices. In recent years, modeling the stock index from the physical perspective by treating it as a dynamic system is emerging [26–29]. To the best of our knowledge, the first attempt to study the stock index, the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX), as well as the options through dynamic models [26], which are derived based on parabola approximation [30]. Another study [24] also considers the stock markets as complex dynamical systems, but their interactions are analyzed by methods of signal processing, the EMD and the EEMD along with detrended cross-correlation analysis (DCCA). Existing studies recognize the nonlinear behaviors of the stock movements [31–34]. However, the models fits [33, 34] are obtained via statistical method during a period of up trend or down trend.

Based on the related work [27], we consider to apply the Price reversion model (PRM), which has a form of the Newton’s law of cooling, to the stock markets. Yet, the empirical evidence in [27] shows that the PRM tends to underestimate the forecasts. Therefore, we identify the reasons of inaccurate forecasts by the PRM and propose to model the stock prices by a newly derived model, i.e. the new Price Reversion Model (nPRM). For the same dynamic model described using the differential equation in the PRM, we derive its solutions rigorously so that the solutions of the nPRM can better fit the realistic stock prices of the markets than the PRM. Two major contributions of the nPRM are as follows. Firstly, the nPRM seeks for the solutions that will not only achieve higher prediction accuracy but also provide useful economic meanings (refer to Figs 1 and 2 below). The coefficients in the nPRM, \( A \) and \( \kappa \), represent the implied equilibrium price, and the convergence speed to the implied equilibrium price, respectively. Secondly, for many predictive models involved in the time series analysis, it is necessary to transform the stock prices to other economic variables, usually the returns or the price increments. The nPRM can make the predictions directly for the stock prices, which implies that the model considers all daily information in the analysis. Furthermore, the conduction of the proposed procedure is simple and convenient.

**Methods**

We treat the stock index, denoted as \( S(t) \), as a dynamic system, which is a differentiable function with respect to time \( t \). We model the subject by the nPRM, represented by a differential equation as follows.

\[
\begin{align*}
\frac{dS(t)}{dt} &= f(S(t)) = \kappa [S(t) - A], \\
\frac{d^2S(t)}{dt^2} &= \kappa \frac{dS(t)}{dt}, \\
S(t_0) &= S_0, \quad S(t_1) = S_1,
\end{align*}
\]

(1)

where \( S(t) \) is the value of the subject at time \( t \), \( A \) and \( \kappa \) are known constants for \( t \in [t_0, t_1] \). We further denote \( v(t) = \frac{dS(t)}{dt} \) and \( a(t) = \frac{d^2S(t)}{dt^2} \), which represent the velocity and acceleration of the dynamic system. The system along with its velocity and acceleration at specified time points \( t_0 \) and \( t_1 \) are denoted as \( (S(t_0), v(t_0), a(t_0)) = (S_0, v_0, a_0) \) and \( (S(t_1), v(t_1), a(t_1)) = (S_1, v_1, a_1) \) respectively for simplicity.
The Eq (1) is adapted from the Newton’s law of cooling [35]. In the Newton’s law of cooling, $\kappa$ is always negative so that the temperature of the subject converges to the environment/surrounding temperature. Additionally, when the difference between the subject temperature and the environment temperature is large, the absolute value of $\kappa$ is large and the speed of convergence is high. When the subject temperature is closer to the environment temperature, the absolute value of $\kappa$ is small, which results in a slower convergence.

Consider the differential Eq (1), $\kappa$ and $A$ are constant during the period $t \in [t_0, t_1]$. Generally they are time related in a dynamic system for a longer time period. Furthermore, the time-related coefficients are often very dynamic in real-world data. For applications, the coefficients are assumed to be constant in a relatively small interval with given values between time $t_0$ and $t_1$.

For the consistence of this paper, let us first review the PRM in the following subsection.

**The Price Reversion Model (PRM)**

In the existing study [27], the PRM assumes that the external force equals zero at $A$. The second derivative is ignored in the Eq (1). Depending on the values of $S(t)$, $S_0$, and $A$, four cases are considered and the solutions are listed as follows.
Case A.
If \( S(t) > S_0 \) and \( (S(t) - A) \cdot (S_0 - A) > 0 \),
\[
S(t) = A + (S_0 - A) \cdot e^{x(t-t_0)}.
\] (2)

Case B.
If \( S(t) > S_0 \) and \( (S(t) - A) \cdot (S_0 - A) < 0 \),
\[
S(t) = A - (S_0 - A) \cdot e^{x(t-t_0)}.
\] (3)

Case C.
If \( S(t) < S_0 \) and \( (S(t) - A) \cdot (S_0 - A) > 0 \),
\[
S(t) = A + (S_0 - A) \cdot e^{\lambda(t-t_0)}.
\] (4)
Case D.

If \(S(t) < S_0\) and \((S(t) - A) \cdot (S_0 - A) < 0\),

\[
S(t) = A - (S_0 - A) \cdot e^{\kappa(t-t_0)}. \tag{5}
\]

However, essentially the situations represented by the Eqs (3) and (5) does not exist since the external force equals zero at \(A\). It motivates us to re-derive the nPRM in the next subsection.

**Derivation of solutions to the new Price Reversion Model (nPRM)**

In Eq (1), \(A\) is an equilibrium price of a stock index for a specific market, which is analogous to the environment temperature, and \(\kappa\) is a growth coefficient. The force on the stock price draws it near the equilibrium if \(\kappa < 0\) or pushes it away from the equilibrium if \(\kappa > 0\). The nPRM is a more generalized dynamic system than that described by the Newton’s law of cooling. Therefore, in the nPRM \(\kappa\) can be either positive or negative. When \(\kappa < 0\), the nPRM is similar to the thermodynamic system; when \(\kappa > 0\), the behavior of the dynamic system is a different story.

If \(\kappa > 0\) in (1), \(A\) is a repeller. Fig 1 is the phase plot of nPRM when \(\kappa > 0\). The first zone demonstrates the situation that the force pushes the price \(S(t)\) away from \(A\) if the initial price \(S_0\) is smaller than \(A\). In the second zone, the force repels the price \(S(t)\) from the equilibrium price \(A\) if the initial price is larger than \(A\). On the contrary, if \(\kappa < 0\) in (1), \(A\) is an attractor. Fig 2 is the phase plot of nPRM when \(\kappa < 0\). When the initial price is higher than \(A\), the force draws the price to the equilibrium price \(A\) as the figure demonstrates in the third zone. In the fourth zone, if the initial price is lower than the equilibrium \(A\), the price converges to \(A\).

We solve the Eq (1) of the nPRM piecewisely as follows. We divide \(S(t)\) into intervals by the critical point, \(A\), where \(\frac{dS}{dt}|_{S=A} = 0\). As a result, \(S(t)\) is strictly monotonic in each of the intervals. From Eq (1), after the change of variables we have

\[
\int_{S(t_0)}^{S(t)} \frac{dr}{r-A} = \int_{S_0}^{S(t)} \kappa dr.
\]

Therefore,

\[
\int_{S(t_0)}^{S(t)} \frac{dr}{r-A} = \ln|r-A|_{S(t)}^{S(t_0)} = \ln[S(t) - A] - \ln(S_0 - A) = \kappa(t-t_0).
\]

Then, we have

\[
S(t) = A + (S_0 - A) \cdot e^{\kappa(t-t_0)}, \quad t \in [t_0, t_1]. \tag{6}
\]

**Case 1.**

Given \(S_0 < A\) and \(\frac{dS(t)}{dt} < 0\) for \(t \in [t_0, t_1]\) as in the first zone in Fig 1, we obtain \(\kappa > 0\). Therefore, the theoretical value \(S(t)\) is strictly decreasing in \(t\), and \(S(t) < A\) for \(t \in [t_0, t_1]\). The solution blows up since

\[
\lim_{t \to \infty} S(t) = \lim_{t \to \infty} [A + (S_0 - A) \cdot e^{\kappa(t-t_0)}] = -\infty.
\]
Case 2.
Given $S_0 > A$ and $\frac{dS(t)}{dt} > 0$ for $t \in [t_0, t_1]$ as in the second zone in Fig 1, we obtain $\kappa > 0$. Therefore, the theoretical value $S(t)$ is strictly increasing in $t$, and $S(t) > A$ for $t \in [t_0, t_1]$. The solution blows up since
\[
\lim_{t \to \infty} S(t) = \lim_{t \to \infty} \left[ A + (S_0 - A) \cdot e^{\kappa(t-t_0)} \right] = \infty.
\]

Case 3.
Given $S_0 < A$ and $\frac{dS(t)}{dt} > 0$ for $t \in [t_0, t_1]$ as in the third zone in Fig 2, we obtain $\kappa < 0$. Therefore, the theoretical value $S(t)$ is strictly increasing in $t$, and $S(t) < A$ for $t \in [t_0, t_1]$. The solution converges to $A$ since
\[
\lim_{t \to \infty} S(t) = \lim_{t \to \infty} \left[ A + (S_0 - A) \cdot e^{\kappa(t-t_0)} \right] = A.
\]

Case 4.
Given $S_0 > A$ and $\frac{dS(t)}{dt} < 0$ for $t \in [t_0, t_1]$ as in the fourth zone in Fig 2, we obtain $\kappa < 0$. Therefore, the theoretical value $S(t)$ is strictly decreasing in $t$, and $S(t) > A$ for $t \in [t_0, t_1]$. The solution converges to $A$ since
\[
\lim_{t \to \infty} S(t) = \lim_{t \to \infty} \left[ A + (S_0 - A) \cdot e^{\kappa(t-t_0)} \right] = A.
\]

A comparison of the cases on different conditions between PRM and nPRM is given by Table 1. As mentioned before, cases B and D in the PRM do not exist due to zero external force at $A$. By examining the phase plots of Figs 1 and 2, the case A of the PRM can be further divided into cases 2 and 4 of the nPRM. The case C of the PRM can be further divided into cases 1 and 3 of the nPRM. We derive solutions for the cases 3 and 4 of the nPRM, which cannot be obtained identically from the solutions of the PRM. As a result, the theoretical values obtained via the PRM are not so accurate as those from the nPRM.

### Implementation details of the nPRM
The differential Eq (1) characterizes a dynamic system, whose movement is continuous and at least second-order differentiable. However, the real data of daily stock prices consist of two parts, the trajectory of the dynamic system and the noise. The trajectory of the dynamic system

| Model | Case | $S_0 - A$ | $S(t) - A$ | $S(t) - S_0$ | $\frac{dS(t)}{dt}$ | $\kappa$ | Solution | Phase plot |
|-------|------|-----------|-------------|---------------|-------------------|--------|----------|------------|
| PRM   | A    | +         | +           | +             | +                 | +      | Eq (2)   | second zone in Fig 1 or fourth zone in Fig 2 |
|       | B    | -         | +           | +             | +                 | +      | Eq (3)   | no existence                      |
|       | C    | -         | -           | -             | -                 | +      | Eq (4)   | first zone in Fig 1 or third zone in Fig 2 |
|       | D    | -         | -           | -             | -                 | -      | Eq (5)   | no existence                      |
| nPRM  | 1    | +         | -           | -             | -                 | +      | Eq (6)   | first zone in Fig 1                |
|       | 2    | +         | +           | +             | +                 | +      | Eq (6)   | second zone in Fig 1                |
|       | 3    | -         | +           | +             | -                 | -      | Eq (6)   | third zone in Fig 2                 |
|       | 4    | +         | +           | -             | -                 | -      | Eq (6)   | fourth zone in Fig 2                |

https://doi.org/10.1371/journal.pone.0266600.t001
is the functional part while the noise is random. Modeling the noisy data, the stock index for instance, by a differential equation is not appropriate. The reasonable approach is to model the functional movement that can be characterized by the differential equation after filtering out the random noise. We establish a procedure of 4 main steps for applying the nPRM to daily stock prices as follows.

**Step 1—Smoothing.** Moving average is a simple and commonly used filtering method, and hence we implement this method on stock data. We filter out the noise of the data, and define the smoothed stock prices as

$$\tilde{S}(t) = \frac{\sum_{i=0}^{M-1} S(t - i)}{M}.$$  

(7)

Afterwards, we apply model (1) to the smoothed data and obtain the theoretical values. In order to obtain accurate calculation of $\kappa$, the adequate smoothing period $M$ has to be specified. For this purpose, we estimate the smoothing period $M$ in the training set, which we will further explain later.

This smoothing step is not involved in the procedure of PRM [27]. Nevertheless, directly modeling the market prices of a stock index suffers from severe fluctuations in coefficient values resulted from random part of the price movements, the Brownian motion. Characterizing the smoothed stock prices by the differential Eq (1) is more reasonable so that this step is crucial in our proposed method nPRM.

**Step 2—Calculating $A$.** To solve (1) we have to know the fixed value of $A$ for a given time period $[t_0, t_1]$ in advance. However, in reality the equilibrium price $A$ is unknown so that the estimation of $A$ is required. $A$ is calculated using smoothed historical data before the period $[t_0, t_1]$. More specifically, we calculate $A$ as

$$A = \frac{\sum_{i=0}^{N-1} \tilde{S}(t_0 - i)}{N},$$  

(8)

that is the average of stock prices during a period earlier than $t_0$. The length of the time interval $[t_0 - N + 1, t_0]$ should be carefully chosen. We also determine the time period $N$ based on the training set.

**Step 3—Calculating and controlling $\kappa$ value.** During the solving of model (1), the value of $\kappa$ is first calculated from

$$\frac{d\tilde{S}(t)}{dt} = \kappa[\tilde{S}_0 - A],$$  

(9)

by substituting $\tilde{S}(t)$ for $S(t)$ in (1). The value of $\kappa$ is constant during a short period $[t_0, t_1]$. However, it changes over time, and thus we further denote it as $\kappa_t, t = t_0, \ldots, t_\tau$ for different time intervals. When $\kappa_t > 0$ and has large value, the theoretical value at time $t$ will explode. Nevertheless, this situation does not correspond with the changes of stock prices in the real market. To avoid generating unreasonable theoretical values, we restrict the values of $\kappa_t$ lower than an upper bound.

A gradient control method [36] is proposed to control the values of the gradient in a reasonable range. The $\kappa_t$ is a one-dimensional gradient in the nPRM, and we apply this method to $\kappa_t$ and restrict its value as

$$\kappa_t' = \begin{cases} 
\kappa_{t,Q3}, & \text{if } \kappa_t > \kappa_{t,Q3}, \\
\kappa_t, & \text{otherwise},
\end{cases}$$  

(10)
where $\kappa_{i,Q3}$ is the third quantile from previous values of $\kappa_i$, for $i = t - j, \cdots, t - 1$ and $j \geq 100$ in the series we obtained. The pool of $\kappa_i$'s comprises at least 100 values from historical data so that we can obtain stable values of $\kappa_i$.

**Step 4—Making $\tau$-step forecasts.** The theoretical values of the stock prices during time $[t_0, t_1]$ are obtained by the solution (6), in which $S_i$ are replaced by $\hat{S}_i$. Assuming the coefficients unchanged for a longer period $[t_0, t_\tau]$, we can obtain theoretical values before time $t_\tau$. For discrete time representation, we obtain theoretical daily closing prices at $t = t_0, t_1, \cdots, t_\tau$ as the $\tau$-step forecasts.

**Model evaluation**

The forecasting error is measured by three deviation-type measurements, the Mean Absolute Percentage Error (MAPE), the Mean Absolute Error (MAE), and the Root Mean Square Error (RMSE), along with three trend-type measurements, directional Symmetry (DS), correct up-trend (CP) and correct down-trend (CD). The definitions are as follows.

$$MAPE = \frac{1}{T} \sum_{i=t_1}^{t_\tau} |\hat{S}(i) - S(i)| \cdot 100,$$

$$MAE = \frac{1}{T} \sum_{i=t_1}^{t_\tau} |\hat{S}(i) - S(i)|,$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=t_1}^{t_\tau} (\hat{S}(i) - S(i))^2},$$

$$DS = \frac{1}{T} \sum_{i=t_1}^{t_\tau} a_i \cdot 100, \quad a_i = \begin{cases} 1, & \text{if } (\hat{S}(i) - \hat{S}(i - 1))(S(i) - S(i - 1)) \geq 0, \\ 0, & \text{otherwise}, \end{cases}$$

$$CP = \frac{1}{T_1} \sum_{i=t_1}^{t_\tau} b_i \cdot 100, \quad \text{where}$$

$$b_i = \begin{cases} 1, & \text{if } (\hat{S}(i) - \hat{S}(i - 1)) > 0 \text{ and } (\hat{S}(i) - \hat{S}(i - 1))(S(i) - S(i - 1)) \geq 0, \\ 0, & \text{otherwise}, \end{cases}$$

$$CD = \frac{1}{T_2} \sum_{i=t_1}^{t_\tau} c_i \cdot 100, \quad \text{where}$$

$$c_i = \begin{cases} 1, & \text{if } (\hat{S}(i) - \hat{S}(i - 1)) < 0 \text{ and } (\hat{S}(i) - \hat{S}(i - 1))(S(i) - S(i - 1)) \geq 0, \\ 0, & \text{otherwise}, \end{cases}$$

where $\hat{S}(i)$ is the theoretical value of the model at time $i$, $S(i)$ is the market value of the data at time $i$, $T$ is the length of the forecasts, $T_1$ is the number of data points belonging to up trend, and $T_2$ is the number of data points belonging to down trend.
A smaller value of the MAPE indicates that the model better describes the movements of the stock prices. The MAE and the RMSE show similar tendency with the MAPE. DS measures the frequency of correct directional forecasts. CP and CD measure the frequency of correct directional forecasts when real market values rise and fall, respectively. The larger values of DS, CP, and CD suggest that the model obtains more correct directional forecasts in general, for uptrend, and for downtrend, respectively.

Results

For application to real data, we split the data into a training set and a testing set, consisting of the earlier 80% and the latest 20% of the data, respectively. We use the training set to specify adequate smoothing period and the time interval for estimation of the equilibrium, that is, the appropriate choices of $M$ and $N$ in (7) and (8), respectively.

Applied data

The empirical study was conducted based on information from real daily closing price data of ten stock indexes from 2009 to 2019. The indexes include Taiwan TAIEX, Japan Nikkei 225, Korea KOSPI, Hong Kong Hang Seng Index, the Philippines PSEi Index, Thailand SET Index, India S&P BSE SENSEX, Singapore Straits Times Index (STI), Indonesia JKSE, and Malaysia KLCI. The data are collected from Refinitiv Datastream. Each stock index is viewed as a dynamic system and characterized by the nPRM separately.

We implemented the nPRM with a few choices of $M$ and $N$ to obtain $\tau$-step forecasts for the training dataset of each stock index. The values of $M$ range from 2 to 10 while the choice of $N$ includes 30, 60, 90, 120, and 180. We set the number of forecasting steps $\tau = 5$. We assess all combinations of $M$ and $N$ by the criterion of the grand mean of MAPEs across 1-step to 5-step forecasts in the training set. Afterwards, we select the combination of the best performance as the coefficient setting for prediction in the testing set. Table 2 shows that the best combination of $M$ and $N$ with the smallest grand mean of MAPEs in the training set for all the ten indexes. In the following, the forecasts are conducted with $M$ and $N$ specified for each stock index in Table 2.

We conduct all the analyses on a computer with an Intel(R) Core(TM) i7–8565U Processor at 1.80GHz with 24.0 GB RAM running Windows 10. All computations are performed using R version 4.0.3 [37] and the R package Metrics.

Empirical results

The main reason of underestimation of the PRM in the existing study [27] is the difficulty in determining the sign of the growth coefficient, $\kappa$, during the integration. In the proposed nPRM procedure, we derive the solution to the differential equation of the nPRM and propose adequate economic explanations of $\kappa$. Additionally, when $\kappa > 0$ we control its value by a gradient control method [36], and the nPRM is therefore more suited to real markets.

Table 2. The best combination of $M$ and $N$ in the training set.

| Stock Index       | $M$ | $N$ | $\tau$-step forecasts | $M$ | $N$ |
|-------------------|-----|-----|------------------------|-----|-----|
| Hang Seng         | 4   | 180 | SET                    | 6   | 180 |
| Nikkei 225        | 5   | 180 | BSE SENSEX             | 5   | 180 |
| KOSPI             | 4   | 180 | STI                    | 4   | 180 |
| TAIEX             | 4   | 180 | JKSE                   | 8   | 180 |
| PSEi              | 7   | 180 | KLCI                   | 4   | 180 |

https://doi.org/10.1371/journal.pone.0266600.t002
Fig 3 displays the 1-step forecasts of the nPRM for Hong Kong Hang Seng Index along with the market values. Fig 3 suggests that the nPRM possesses accurate forecasting ability. The patterns of the forecasts from 2-step to 5-step by the nPRM are all similar to Fig 3, and thus we omit the demonstration here. Figs 4–12 illustrate the 1-step forecasts of the nPRM for other Asian stock indexes individually, which also indicate accurate forecasts by the nPRM.

Table 3 displays the MAPEs of applying the investigated methods to all of the ten Asian Indexes from 1-step forecasts to 5-step forecasts. Since the performances of three deviation-type errors (MAPE, MAE, and RMSE) are similar, we simply report and discuss the results of the MAPEs. From Table 3, the martingale (see details in S1 Appendix) has the smallest errors followed by those from the nPRM, and the curve fitting (see details in S2 Appendix) has the worst performance.

Table 4 displays the directional symmetry (SD) of the nPRM and the curve fitting from 1-step forecasts to 5-step forecasts for ten Asian stock indexes. The trend-type errors cannot be applied to the forecasts of the martingale since the martingale forecasts the price of the next day by the present price. Therefore, we only compare the nPRM and the curve fitting here. Table 4 shows that nPRM and the curve fitting have similar percentage of correct trend forecasts. The differences between up-trend forecasts and down-trend forecasts of the nPRM and the curve fitting are also not significant (see details in S1 and S2 Tables).
Apart from the martingale, we can investigate the level of market efficiency by sample entropy [38, 39]. Sample entropy measures complexity. When we have a time series of length \( n \) as \( \{S(1), S(2), \ldots, S(n)\} \), a template vector of length \( m \) is defined as \( S_m(i) = \{S(i), S(i + 1), \ldots, S(i + m - 1)\} \) and the distance function \( d[S_m(i), S_m(j)], i \neq j \) is defined as the Chebyshev distance.
distance. Then, we define the sample entropy of the series, \( SampEn \), as

\[
SampEn = -\log\frac{B^{m+1}}{B^{m}},
\]  

(11)
where $B^m$ is number of template vector pairs having $d[S_m(i), S_m(j)] \leq r$, and $r$ is the tolerance. The value of sample entropy is nonnegative. The value 0 of sample entropy indicates that the series is perfectly regular without noise while a high value reflects randomness and unpredictability. If we examine the market efficiency from the perspective of sample entropy, a high value of sample entropy indicates a market with high level of efficiency. To show how large the possible value of sample entropy from a highly random series can be, we simulate a series of length 180 from a normal distribution with zero mean and unit variance for 1000 times. The average value of sample entropy is 2.21.

Table 5 shows the mean value of sample entropy calculated from all series used in estimation of $A$ and $\kappa$ to make forecasts dynamically in the testing set for ten indexes in our empirical study. The mean values range between 0.5952 (Korea KOSPI) and 0.7833 (the Philippines PSEi), which represent similar complexity among these markets. We suspect that the forecasting errors should be larger for markets with higher level of efficiency since it is harder to predict the series with random noise. By the order of the mean sample entropy values and the MAPEs for the ten stock indexes, the nPRM shows a tendency toward better forecasts for series with lower market efficiency. That is, the forecasting ability of the nPRM usually increases as the mean sample entropy decreases. However, there is an exception of Taiwan TAIEX. The mean sample entropy obtained from the TAIEX series is the second highest, so that we won’t expect a smaller prediction error among the ten indexes. However, the MAPEs
of the nPRM for 1-step to 5-step forecasts are relatively low. The changes of TAIEX are similar to the movement which is characterized by the nPRM among the ten markets.

We summarize the performances of the methods as follows. Considering the deviation-type forecasting errors, the martingale seems to perform the best among the methods. According to the empirical evidence [2], the test based on the martingale can be used to examine the market efficiency of the stock market. Despite the absence of the test in this study, the low errors of the martingale for these ten Asian stock indexes imply that the markets tend to be efficient during the period of the data. The martingale suggests that no information we can obtain from the past to forecast the future stock prices. However, our proposed method nPRM still captures the trend of the stock movements quite well, and provides meaningful coefficients, $A$ and $\kappa$, as well under the circumstances. If applying the nPRM to the markets that are less efficient, the model coefficients may acquire more information with economic meanings and obtain more accurate forecasts than the martingale.

The average computational times of all the investigated methods are shown in Table 6. The martingale uses the least computational time while the curve fitting spends the most. The

| Table 3. MAPE (%) for the testing set. |
|----------------------------------|---|---|---|---|---|
|                                  | $\tau = 1$ | $\tau = 2$ | $\tau = 3$ | $\tau = 4$ | $\tau = 5$ |
| Nikkei 225                       |             |             |             |             |             |
| nPRM                             | 1.04        | 1.27        | 1.53        | 1.78        | 2.05        |
| Martingale                       | 0.72        | 1.05        | 1.28        | 1.52        | 1.74        |
| Curve fitting                    | 3.05        | 3.32        | 3.60        | 3.87        | 4.15        |
| Hang Seng                        |             |             |             |             |             |
| nPRM                             | 1.12        | 1.44        | 1.77        | 2.11        | 2.43        |
| Martingale                       | 0.81        | 1.21        | 1.53        | 1.79        | 2.00        |
| Curve fitting                    | 2.33        | 2.52        | 2.71        | 2.89        | 3.06        |
| TAIEX                            |             |             |             |             |             |
| nPRM                             | 0.76        | 1.00        | 1.21        | 1.43        | 1.65        |
| Martingale                       | 0.57        | 0.85        | 1.03        | 1.22        | 1.38        |
| Curve fitting                    | 1.96        | 2.14        | 2.32        | 2.49        | 2.67        |
| KOPSI                            |             |             |             |             |             |
| nPRM                             | 0.86        | 1.10        | 1.38        | 1.62        | 1.84        |
| Martingale                       | 0.61        | 0.92        | 1.17        | 1.39        | 1.56        |
| Curve fitting                    | 2.54        | 2.76        | 2.98        | 3.20        | 3.42        |
| PSEi                             |             |             |             |             |             |
| nPRM                             | 1.46        | 1.71        | 1.93        | 2.16        | 2.35        |
| Martingale                       | 0.75        | 1.02        | 1.22        | 1.42        | 1.58        |
| Curve fitting                    | 3.09        | 3.36        | 3.65        | 3.93        | 4.22        |
| SET                              |             |             |             |             |             |
| nPRM                             | 1.18        | 1.47        | 1.70        | 1.98        | 2.20        |
| Martingale                       | 0.46        | 0.69        | 0.86        | 1.01        | 1.13        |
| Curve fitting                    | 2.08        | 2.25        | 2.43        | 2.60        | 2.78        |
| BSE SENSEX                       |             |             |             |             |             |
| nPRM                             | 1.31        | 1.62        | 1.88        | 2.20        | 2.42        |
| Martingale                       | 0.53        | 0.78        | 0.99        | 1.17        | 1.33        |
| Curve fitting                    | 2.29        | 2.51        | 2.72        | 2.94        | 3.17        |
| STI                              |             |             |             |             |             |
| nPRM                             | 1.19        | 1.52        | 1.77        | 2.09        | 2.32        |
| Martingale                       | 0.54        | 0.76        | 0.91        | 1.06        | 1.19        |
| Curve fitting                    | 1.76        | 1.91        | 2.05        | 2.18        | 2.31        |
| JKSE                             |             |             |             |             |             |
| nPRM                             | 1.27        | 1.42        | 1.57        | 1.73        | 1.84        |
| Martingale                       | 0.60        | 0.85        | 1.02        | 1.14        | 1.27        |
| Curve fitting                    | 2.39        | 2.62        | 2.85        | 3.08        | 3.31        |
| KLCI                             |             |             |             |             |             |
| nPRM                             | 0.98        | 1.30        | 1.58        | 1.89        | 2.07        |
| Martingale                       | 0.49        | 0.72        | 0.90        | 1.05        | 1.18        |
| Curve fitting                    | 1.76        | 1.92        | 2.07        | 2.22        | 2.38        |

https://doi.org/10.1371/journal.pone.0266600.t003
nPRM is computationally efficient in view of the relationship between the forecasting accuracy and the computational time.

**Discussion**

In this study, we propose a novel procedure adapted from the thermodynamic system to model the stock prices by a newly derived model, the new Price Reversion Model (nPRM), which is a modification of PRM. From the viewpoint of signal processing, the proposed procedure of nPRM models the stock prices with all available information. Therefore, the coefficients in the nPRM preserve and reveal more information about the stock movements economically. They also may be treated as important features for forecasting by advanced models such as deep learning models.
We evaluate the forecasting accuracy of the nPRM along with the martingale, and the cubic curve fitting by ten Asian stock index examples. By investigating the daily closing prices, the nPRM accurately characterizes and forecasts the large-scale motions in these ten stock markets. Although the forecasting accuracy declines over time, which is the same phenomenon for all the methods, the deviation-type errors of the nPRM are the smallest except for the martingale. However, the martingale can not be employed to forecast the trend of the stock even though it offers an eligible prediction about the prices if the market is efficient.

The trend-type errors of the nPRM and the curve fitting are close. Furthermore, the trend-type errors of the two methods do not significantly deviate from 0.5 from 1-step to 5-step forecasts. This indicates that the nPRM and the curve fitting does not possess good trend forecasting ability. In order to improve the trend forecasting in the future, we will develop a more sophisticated method to calculate the $\kappa$ values without the gradient control.

**Supporting information**

S1 Appendix. The martingale.  
(PDF)

S2 Appendix. The curve fitting.  
(PDF)

S1 Table. Correct up-trend (CP) (%) for the testing set.  
(PDF)

S2 Table. Correct down-trend (CD) (%) for the testing set.  
(PDF)

**Acknowledgments**

This research would not have been possible without the conceptual and theoretical development from Prof. Meng-Rong Li, the retired professor from Department of Mathematical Sciences, National Chengchi University. The authors are inspired by his works and have unreserved support from him.

**Author Contributions**

**Data curation:** Tsung-Jui Chiang-Lin.

**Formal analysis:** Tsung-Jui Chiang-Lin, Yong-Shiuan Lee.

**Funding acquisition:** Shang-Yueh Tsai.

**Investigation:** Tsung-Jui Chiang-Lin, Yong-Shiuan Lee, Tzong-Hann Shieh.

**Methodology:** Tsung-Jui Chiang-Lin, Yong-Shiuan Lee, Tzong-Hann Shieh, Chien-Chang Yen, Shang-Yueh Tsai.

**Project administration:** Tzong-Hann Shieh.

**Resources:** Tsung-Jui Chiang-Lin, Yong-Shiuan Lee.

**Supervision:** Tzong-Hann Shieh, Shang-Yueh Tsai.

**Validation:** Tsung-Jui Chiang-Lin, Yong-Shiuan Lee, Tzong-Hann Shieh, Chien-Chang Yen.

**Visualization:** Tsung-Jui Chiang-Lin, Yong-Shiuan Lee.

**Writing – original draft:** Tsung-Jui Chiang-Lin, Yong-Shiuan Lee.
Writing – review & editing: Tzong-Hann Shieh, Chien-Chang Yen, Shang-Yueh Tsai.

References

1. Cowles A 3rd. Can stock market forecasters forecast? Econometrica: Journal of the Econometric Society. 1933; p. 309–324. https://doi.org/10.2307/1907042
2. Lim KP, Luo W. The weak-form efficiency of Asian stock markets: New evidence from generalized spectral martingale test. Applied Economics Letters. 2012; 19(10):905–908. https://doi.org/10.1080/13504851.2011.607123
3. Malkiel BG, Fama EF. Efficient capital markets: A review of theory and empirical work. The Journal of Finance. 1970; 25(2):383–417. https://doi.org/10.1111/j.1540-6261.1970.tb00518.x
4. Fama EF. The behavior of stock-market prices. The Journal of Business. 1965; 38(1):34–105. https://doi.org/10.1086/294743
5. Cochrane JH. Presidential address: Discount rates. The Journal of Finance. 2011; 66(4):1047–1108. https://doi.org/10.1111/j.1540-6261.2011.01671.x
6. Abbaranel JS, Bushee BJ. Fundamental analysis, future earnings, and stock prices. Journal of accounting research. 1997; 35(1):1–24. https://doi.org/10.2307/2491464
7. Abbaranel JS, Bushee BJ. Abnormal returns to a fundamental analysis strategy. Accounting Review. 1998; p. 19–45.
8. Park CH, Irwin SH. What do we know about the profitability of technical analysis? Journal of Economic surveys. 2007; 21(4):786–826. https://doi.org/10.1111/j.1467-6419.2007.00519.x
9. Campbell JY, Shiller RJ. Stock prices, earnings, and expected dividends. the Journal of Finance. 1988; 43(3):661–676. https://doi.org/10.1111/j.1540-6261.1988.tb04598.x
10. Lewellen J. Predicting returns with financial ratios. Journal of Financial Economics. 2004; 74(2):209–235. https://doi.org/10.1016/j.jfineco.2002.11.002
11. Welch I, Goyal A. A comprehensive look at the empirical performance of equity prediction. The Review of Financial Studies. 2008; 21(4):1455–1508. https://doi.org/10.1093/rfs/hhm014
12. Campbell JY, Shiller RJ. The dividend-price ratio and expectations of future dividends and discount factors. The Review of Financial Studies. 1998; 1(3):195–228. https://doi.org/10.1093/rfs/1.3.195
13. Campbell JY, Shiller RJ. Valuation ratios and the long-run stock market outlook: An update. Advances in Behavioral Finance. 2005; 2:173–201.
14. Cochrane JH. The dog that did not bark: A defense of return predictability. The Review of Financial Studies. 2008; 21(4):1533–1575. https://doi.org/10.1093/rfs/hhm046
15. Lettau M, Ludvigson S. Consumption, aggregate wealth, and expected stock returns. the Journal of Finance. 2001; 56(3):815–849. https://doi.org/10.1111/0022-1082.00347
16. Rapach DE, Strauss JK, Zhou G. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. The Review of Financial Studies. 2010; 23(2):821–862. https://doi.org/10.1093/rfs/hhp063
17. Cheng CH, Wei LY. A novel time-series model based on empirical mode decomposition for forecasting TAIEX. Economic Modelling. 2014; 36:136–141. https://doi.org/10.1016/j.econmod.2013.09.033
18. Atsalakis GS, Valavanis KM. Surveying stock market forecasting techniques—Part II: Soft computing methods. Expert Systems with Applications. 2009; 36(3):5932–5941. https://doi.org/10.1016/j.eswa.2008.07.006
19. Guresen E, Kayakutlu G, Daim TU. Using artificial neural network models in stock market index prediction. Expert Systems with Applications. 2011; 38(8):10389–10397. https://doi.org/10.1016/j.eswa.2011.02.068
20. Wang J, Wang J. Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks. Neurocomputing. 2015; 156:68–78. https://doi.org/10.1016/j.neucom.2014.12.084
21. Zhang C, Pan H. A novel hybrid model based on EMD-BPNN for forecasting US and UK stock indices. In: 2015 IEEE International Conference on Progress in Informatics and Computing (PIC). IEEE; 2015. p. 113–117.
22. Yan B, Aasma M, Zhang Y. A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM. Expert Systems with Applications. 2020; 159:113609. https://doi.org/10.1016/j.eswa.2020.113609
23. Basak S, Kar S, Saha S, Khaidem L, Dey SR. Predicting the direction of stock market prices using tree-based classifiers. The North American Journal of Economics and Finance. 2019; 47:552–567. https://doi.org/10.1016/j.najef.2018.06.013

24. Xu M, Shang P, Lin A. Cross-correlation analysis of stock markets using EMD and EEMD. Physica A: Statistical Mechanics and its Applications. 2016; 442:82–90. https://doi.org/10.1016/j.physa.2015.08.063

25. Mallikarjuna M, Rao RP. Evaluation of forecasting methods from selected stock market returns. Financial Innovation. 2019; 5(40):1–16. https://doi.org/10.1186/s40854-019-0157-x.

26. Chiang-Lin TJ, Li MR, Lee YS. TAIEX index option model by using nonlinear differential equation. Mathematical and Computational Applications. 2014; 19(1):78–92. https://doi.org/10.3390/mca19010078

27. Chen NP, Li MR, Chiang-Lin TJ, Lee YS, Miao DWC. Applications of linear ordinary differential equations and dynamic system to economics—An example of Taiwan stock index TAIEX. International Journal of Dynamical Systems and Differential Equations. 2017; 7(2):95–111. https://doi.org/10.1504/IJDSDE.2017.10006751

28. Li MR, Chiang-Lin TJ, Lee YS. Applications of parameterized nonlinear ordinary differential equations and dynamic systems—An example of the Taiwan stock index. International Journal of Differential Equations. 2018; p. Article ID 4762485. https://doi.org/10.1155/2018/4762485

29. Li MR, Miao DWC, Chiang-Lin TJ, Lee YS. Modelling DAX by applying parabola approximation method. International Journal of Computing Science and Mathematics. 2019; 10(6):566–579. https://doi.org/10.1504/IJCSM.2019.10025668

30. Li MR, Shieh TH, Yue CJ, Lee P, Li YT. Parabola method in ordinary differential equation. Taiwanese Journal of Mathematics. 2011; 15(4):1841–1857. https://doi.org/10.1055/tjm/1500406383

31. Abhyankar A, Copeland LS, Wong W. Nonlinear dynamics in real-time equity market indices: Evidence from the United Kingdom. The Economic Journal. 1995; 105(431):864–880. https://doi.org/10.12307/2235155

32. Abhyankar A, Copeland LS, Wong W. Uncovering nonlinear structure in real-time stock-market indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100. Journal of Business and Economic Statistics. 1997; 15(1):1–14. https://doi.org/10.1080/07350015.1997.10524681

33. Gkranas A, Rendoumis VL, Polatoglou HM. Athens and Lisbon stock markets—A thermodynamic approach. WSEAS Transactions on Business and Economics. 2004; 1(1):95–100.

34. Zarikas V, Christopoulos AG, Rendoumis VL. A thermodynamic description of the time evolution of a stock market index. European Journal of Economics, Finance and Administrative Sciences. 2009; 16:73–83.

35. Besson U. The history of the cooling law: When the search for simplicity can be an obstacle. Science & Education. 2012; 21(8):1085–1110. https://doi.org/10.1007/s11191-010-9324-1

36. Shieh TH, Li MR, Li YT. Study of new accurate supplementary energy modeling with multi-components based on newly developed high resolution scheme. International Communications in Heat and Mass Transfer. 2019; 107:121–140. https://doi.org/10.1016/j.icheatmasstransfer.2019.04.003

37. R Core Team. R: A Language and Environment for Statistical Computing; 2020. Available from: https://www.R-project.org/.

38. Yentes JM, Hunt N, Schmid KK, Kaipust JP, McGrath D, Stergiou N. The appropriate use of approximate entropy and sample entropy with short data sets. Annals of biomedical engineering. 2013; 41(2):349–365. https://doi.org/10.1007/s10439-012-0668-3 PMID: 23064819

39. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. American Journal of Physiology-Heart and Circulatory Physiology. 2000; 278:H2039–H2049. https://doi.org/10.1152/ajpheart.2000.278.6.H2039 PMID: 10843903