Lim et al. Financial Innovation (2026) 12:4
https://doi.org/10.1186/s40854-025-00807-7

Financial Innovation

RESEARCH

Open Access

Dynamic factor analysis of price movements in the Philippine Stock Exchange

Brian Godwin Lim[1,2*], Dominic Dayta[1], Benedict Ryan Tiu[1], Renzo Roel Tan[1,2,3], Len Patrick Dominic Garces[4,5] and Kazushi Ikeda[1]

*Correspondence: brian.lim@naist.ac.jp

[1] Division of Information Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Nara 630-0192, Japan
[2] Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
[3] Department of Quantitative Methods and Information Technology, John Gokongwei School of Management, Ateneo de Manila University, Metro Manila 1108, Philippines
[4] School of Mathematical and Physical Sciences, Faculty of Science, University of Technology Sydney, Ultimo, NSW 2007, Australia
[5] Department of Mathematics, School of Science and Engineering, Ateneo de Manila University, Metro Manila 1108, Philippines

Abstract

The intricate dynamics of stock markets have led to extensive research on models that are able to effectively explain their inherent complexities. This study leverages the econometrics literature to explore the dynamic factor model as an interpretable model with sufficient predictive capabilities for capturing essential market phenomena. Although the model has been extensively applied for predictive purposes, this study focuses on analyzing the extracted loadings and common factors as an alternative framework for understanding stock price dynamics. The results reveal novel insights into traditional market theories when applied to the Philippine Stock Exchange using the Kalman method and maximum likelihood estimation, with subsequent validation against the capital asset pricing model. Notably, a one-factor model extracts a common factor representing systematic or market dynamics similar to the composite index, whereas a two-factor model extracts common factors representing market trends and volatility. Furthermore, an application of the model for nowcasting the growth rates of the Philippine gross domestic product highlights the potential of the extracted common factors as viable real-time market indicators, yielding over a 34% decrease in the out-of-sample prediction error. Overall, the results underscore the value of dynamic factor analysis in gaining a deeper understanding of market price movement dynamics.

Keywords: Dynamic factor analysis, Kalman filtering, Philippine Stock Exchange, State-space model, Stock price movement

Introduction

Researchers, practitioners, and investors have long been interested in financial markets due to the opportunities they offer to invest excess funds and generate positive returns. Among the primary objectives in the study of financial markets is the accurate prediction of stock price movements, as this enables one to outperform the market and achieve significant gains. Substantial efforts have been made to develop models that can effectively capture the complex dynamics of stock markets, providing tools for making informed decisions based on historical data (Chernov et al. 2003; Kothari and Zimmerman 1995; Long et al. 2019).

In recent years, machine learning models have gained immense popularity for feature selection and extraction (Kumari et al. 2023; Htun et al. 2023) as well as for predicting future price movements (Muhammad 2023; Li et al. 2023; Lawi et al. 2022; Shen and Shafiq 2020; Zhong and Enke 2019). Despite their success, the *black box* nature of these models hinders their interpretability. Traditional asset pricing models, such as the capital asset pricing model (CAPM), the Fama–French models, and the arbitrage pricing theory (APT) model, offer a simpler framework by characterizing the linear relationships between stock returns and some underlying factors (Giglio et al. 2021; Lintner 1965; Sharpe 1964; Fama and French 1993). However, these interpretable linear models often fail to capture the market's complexities. Thus, models must combine the interpretability of conventional linear asset pricing models with the predictive capabilities of machine learning models to provide a deeper understanding of stock price movement dynamics.

This study leverages the econometrics literature by analyzing stock market price movements through the lens of dynamic factor analysis. Although not commonly used in financial applications, the dynamic factor model (DFM) may be used to explain stock returns as the sum of a common component and an idiosyncratic component (Geweke 1977). The former is further decomposed as a linear combination of a set of predictive features, known as the common factors, extracted by the unsupervised model. Thus, similar to traditional linear asset pricing models, DFM offers an interpretable way to explain market phenomena while possessing the predictive capabilities of machine learning models. Moreover, although the model has been predominantly used for predictive purposes, this study focuses on analyzing the extracted loadings and common factors as an alternative approach to understanding the complex dynamics of price movements in the stock market.

Using the Kalman method and maximum likelihood estimation, the analysis of the Philippine Stock Exchange (PSE), validated against the CAPM, provides novel and alternative insights into classical market theories. The common factor in a one-factor model may be used to represent the systematic or market dynamics similar to the composite index, whereas the common factors in a two-factor model may be used to represent market trends and volatility. Moreover, an application of the model for nowcasting the growth rates of the Philippine gross domestic product (GDP) further demonstrates the utility of the extracted common factors as viable real-time market indicators, achieving a reduction of over 34% in the out-of-sample prediction error. These results highlight the unique perspective of dynamic factor analysis in understanding the dynamics of market price movements.

The study is further organized as follows. The next section describes related works with a particular focus on the CAPM, APT model, and principal component analysis (PCA). The following section then formally introduces the DFM, which addresses the limitations of the previous models. This section also covers the model fitting methodology, validation procedures, and implementation details. The subsequent section presents the results of the model when applied to the PSE, followed by a demonstration of the utility of the extracted common factors within a macroeconomic nowcasting application. The last section then concludes with a summary of the research and recommendations for future work.

Related works

Recent developments in the analysis of price movements in stock markets primarily focus on the application of machine learning models, which generally serve two main objectives: performing feature extraction for downstream analysis tasks and predicting future price movements. Toward the first objective, models such as random forests and autoencoders have been employed to extract features for explaining price movements (Kumari et al. 2023; Htun et al. 2023). For example, Gunduz (2021) applied variational autoencoders to extract features that successfully predicted the direction of stock price movements using long short-term memory and LightGBM models. Additionally, Shahvaroughi Farahani et al. (2021) used genetic algorithms to select representative features for the same purpose, using a simple neural network model. From another aspect, toward the second objective, Wang (2024) demonstrated the effectiveness of a neural network model in capturing the nonlinear relationships between firm-specific and macroeconomic factors with stock price returns. Similarly, Htun et al. (2024) explored the use of random forests, support vector machines, and long short-term memory models to predict the excess return of a stock relative to a composite index. Other studies employed deep learning models, such as Transformers and gated recurrent units, to capture the complex dynamics of stock price movements (Muhammad 2023; Li et al. 2023; Lawi et al. 2022; Shen and Shafiq 2020; Zhong and Enke 2019). Notably, the majority of these models rely only on historical stock price data as features. Nevertheless, some of these works incorporate additional features such as technical (Gunduz 2021; Shahvaroughi Farahani et al. 2021), fundamental (Wang 2024; Shen and Shafiq 2020), and macroeconomic indicators (Wang 2024) to boost predictive capability. Despite their superior performance, these machine learning models often lack the interpretability of classical linear models. Moreover, the features extracted through such models may possess limited explanatory value for analyses beyond price movement prediction.

The DFM, rooted in econometrics literature, addresses several limitations of machine learning models while remaining performant in predictive applications (Luciani et al. 2018; Hayashi et al. 2022; Chernis et al. 2020). Moreover, DFM also provides interpretable loadings and latent features that may be used for further analysis. In what follows, the CAPM and the APT models are first introduced as foundational models, offering insights into traditional approaches for analyzing stock price movements. Subsequently, PCA is presented as an alternative method for extracting data-driven factors within the APT model.

Capital asset pricing model CAPM is widely recognized in finance literature as a means of explaining stock price returns. It describes the linear relationship between the expected return of a given stock and its exposure to systematic or market risks (Lintner 1965; Sharpe 1964). Suppose $\mathbb{E}(R_i)$ is the expected return of stock i. The model assumes the following dynamics:

$$\mathbb{E}(R_i) - R_F = \beta_i[\mathbb{E}(R_M) - R_F], \tag{1}$$

where R_F is the risk-free rate of return, $\mathbb{E}(R_M)$ is the expected market return, and β_i is the CAPM beta of stock i that measures the sensitivity of the risk premium $\mathbb{E}(R_i) - R_F$ to the expected excess market return $\mathbb{E}(R_M) - R_F$.

Numerous studies have used CAPM to investigate the relationship between risk and return (Blume and Friend 1973; Perold 2004; Elbannan 2014; Rossi 2016). CAPM assumes that the expected return of a stock co-moves with the expected return of the market and that variations in the CAPM beta are sufficient to explain the cross-sectional differences in stock price returns.

Arbitrage pricing theory model The APT model is another linear model widely used in finance literature to explain stock price returns. This model extends the CAPM since empirical evidence indicates the need for a multifactor model to explain stock price dynamics (Barucci and Fontana 2017). The APT model assumes that stock returns are explained by a linear combination of a finite number of risk factors and a random factor specific to each stock. Suppose R_i is the return of stock i. The model assumes the following dynamics:

$$R_i - R_F = \boldsymbol{\beta_i}^\top \boldsymbol{F} + Z_i, \tag{2}$$

where $\boldsymbol{F}$ is a vector of n risk factors, Z_i is a stock-specific random factor for stock i, and $\boldsymbol{\beta_i}$ measures the sensitivity of stock i to risk factors.

The APT model relaxes some assumptions of the CAPM and uses firm-specific or macroeconomic factors for $\boldsymbol{F}$ to explain the stock price returns. Firm-specific factors include the book-to-market ratio, dividend yield, and cash-flow-to-price ratio, whereas macroeconomic factors include expected inflation, the yield spread between long- and short-term interest rates, and the yield spread between corporate high- and low-grade bonds (Barucci and Fontana 2017). With the inclusion of different risk factors, the model can better explain stock price movements than CAPM (Reinganum 1981; Elshqirat 2019).

Despite these advantages, the question of which and how many factors to include remains unresolved, with empirical evidence indicating that models utilizing derived factors may sometimes outperform those based on traditional economic and financial indicators (French 2017; Reinganum 1981). This makes a compelling case for using factors derived from models such as PCA, which provides a systematic and data-driven approach to factor extraction, addressing the challenge of factor selection within the APT model.

Principal component analysis PCA is another linear model widely used in the literature. Unlike CAPM and APT models, which are used to explain stock price returns, PCA is primarily used for dimensionality reduction. This model compresses high-dimensional data into a lower-dimensional representation that preserves as much information and variability from the original data as possible (Jolliffe 2002; Pearson 1901). The compressed data, known as the principal components, are mutually uncorrelated linear combinations of the original variables (Jolliffe and Cadima 2016). Suppose that the data is represented as a $T \times S$ matrix $\boldsymbol{X} = [\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_S}]$ containing T observations of S variables. The model determines the linear combination

$$\boldsymbol{Xa} = \sum_{s=1}^{S} a_s \boldsymbol{x_s} \tag{3}$$

that maximizes the variance given by $\mathrm{Var}(\boldsymbol{Xa}) = \boldsymbol{a}^\top \hat{\boldsymbol{\Sigma}} \boldsymbol{a}$, where $\hat{\boldsymbol{\Sigma}}$ is the sample covariance matrix of $\boldsymbol{X}$. Thus, the problem is reduced to maximizing $\boldsymbol{a}^\top \hat{\boldsymbol{\Sigma}} \boldsymbol{a}$ subject to $\|\boldsymbol{a}\| = 1$, restricting $\boldsymbol{a}$ to be a unit vector. Using Lagrange multipliers, this is equivalent to maximizing

$$\boldsymbol{a}^\top \hat{\boldsymbol{\Sigma}} \boldsymbol{a} - \lambda\left(\boldsymbol{a}^\top \boldsymbol{a} - 1\right). \tag{4}$$

The above optimization problem results in the equation $\hat{\boldsymbol{\Sigma}} \boldsymbol{a} = \lambda \boldsymbol{a}$, indicating that $\boldsymbol{a}$ is a unit eigenvector and λ is the corresponding eigenvalue of the sample covariance matrix $\hat{\boldsymbol{\Sigma}}$. Moreover, given that

$$\mathrm{Var}(\boldsymbol{Xa}) = \boldsymbol{a}^\top \hat{\boldsymbol{\Sigma}} \boldsymbol{a} = \lambda \boldsymbol{a}^\top \boldsymbol{a} = \lambda, \tag{5}$$

λ must be the largest eigenvalue of $\hat{\boldsymbol{\Sigma}}$. The first principal component is therefore calculated as $\boldsymbol{Xa}_{(\boldsymbol{1})}$, where $\boldsymbol{a}_{(\boldsymbol{1})}$ is the unit eigenvector associated with the largest eigenvalue of $\hat{\boldsymbol{\Sigma}}$. The succeeding principal components may be similarly obtained by adding the constraint

$$\mathrm{Cov}\left(\boldsymbol{Xa}_{(i)}, \boldsymbol{Xa}_{(j)}\right) = \boldsymbol{a}_{(\boldsymbol{i})}^\top \hat{\boldsymbol{\Sigma}} \boldsymbol{a}_{(\boldsymbol{j})} = \lambda_{(j)} \boldsymbol{a}_{(\boldsymbol{i})}^\top \boldsymbol{a}_{(\boldsymbol{j})} = 0 \tag{6}$$

or equivalently, $\boldsymbol{a}_{(\boldsymbol{i})}^\top \boldsymbol{a}_{(\boldsymbol{j})} = 0$ for $j < i$. This results in $\boldsymbol{Xa}_{(\boldsymbol{i})}$ being the ith principal component, where $\boldsymbol{a}_{(\boldsymbol{i})}$ is the unit eigenvector associated with the ith largest eigenvalue of $\hat{\boldsymbol{\Sigma}}$ (Jolliffe 2002).

PCA is widely used in the literature and practical applications (Ghorbani and Chong 2020; Lim et al. 2024; Yu 2023; Xi et al. 2024). It learns the optimal linear compression of high-dimensional data into principal components without requiring additional data, making it an unsupervised feature extraction model. However, PCA inherently assumes that the T observations are independent, which does not hold when $\boldsymbol{X}$ is a time series data.

This study employs DFM to address the limitations of these models by integrating the strengths of linear and machine learning models, thereby effectively achieving a balance between interpretability and predictive performance. Although the model has been extensively applied for predictive purposes, this study focuses on the extracted loadings and latent features, along with their corresponding economic interpretation, as an alternative approach to understanding the dynamics of price movements in the stock market.

Dynamic factor model

DFM is another linear model that combines the features of PCA and APT models and may be regarded as an unsupervised time series extension of the latter (Geweke 1977). Similar to the APT model, DFM is a multifactor model that can be used to explain stock price returns. However, unlike most asset pricing factor models that rely on a predefined set of factors, DFM does not require such inputs. Instead, it directly estimates the factors from the observed data, offering valuable insights into the dynamics of price movements in a data-driven and unsupervised manner similar to PCA. Suppose R_{it} is the return of stock i at time t. The model assumes the following dynamics:

$$R_{it} = \boldsymbol{\beta_i}^\top \boldsymbol{F_t} + \sigma_i Z_{it}, \tag{7}$$

$$\boldsymbol{F_t} = \boldsymbol{\Lambda_1 F_{t-1}} + \boldsymbol{\Lambda_2 F_{t-2}} + \cdots + \boldsymbol{\Lambda_p F_{t-p}} + \boldsymbol{\varepsilon_t}, \tag{8}$$

$$Z_{it} = \psi_{i1} Z_{i(t-1)} + \psi_{i2} Z_{i(t-2)} + \cdots + \psi_{iq} Z_{i(t-q)} + \gamma_{it}, \tag{9}$$

where $\boldsymbol{F_t}$ is a vector of n common factors at time t, Z_{it} is the stock-specific factor of stock i at time t, $\boldsymbol{\beta_i}$ is the vector of loadings of stock i for the common factors, σ_i is the loading of stock i for the stock-specific factor, $\boldsymbol{\Lambda_j}$ is an $n \times n$ vector autoregressive coefficient matrix for $\boldsymbol{F_{t-j}}$, ψ_{ij} is an autoregressive coefficient for $Z_{i(t-j)}$, and $\boldsymbol{\varepsilon_t} \sim N(\mathbf{0}, \boldsymbol{I_n})$ and $\gamma_{it} \sim N(0, 1)$ are Gaussian noise processes. Furthermore, γ_{it} and $\gamma_{jt'}$ are independent for all $i \neq j$ and any t, t'. Thus, the model assumes that the return of stock i at time t is a combination of two components: common and idiosyncratic. A linear combination of the common factors $\boldsymbol{F_t}$ governs the common component, whereas Z_{it}, a stock-specific factor, governs the idiosyncratic component.

In addition, the model assumes that stock price returns follow a Gaussian distribution similar to other literature (Kendall 1953; Osborne 1959; Black and Scholes 1973; Officer 1972; Roll and Ross 1980; Phelan 1997; Marathe and Ryan 2005; Hull et al. 2016; Li 2023). Under the stationarity assumption of the vector autoregressive and autoregressive processes and the model specifications, the common factors $\boldsymbol{F_t} \sim N(\mathbf{0}, \boldsymbol{\Sigma_F})$ and stock-specific factors $Z_{it} \sim N(0, \sigma_{Zi})$ follow a Gaussian distribution for all i and t, where $\boldsymbol{\Sigma_F}$ is some $n \times n$ covariance matrix and $\sigma_{Zi} > 0$. Hence, since $\boldsymbol{F_t}$ and Z_{it} are independent by construction, R_{it} follows a Gaussian distribution with parameters

$$\mathbb{E}(R_{it}) = \mathbb{E}\left(\boldsymbol{\beta_i}^\top \boldsymbol{F_t}\right) + \mathbb{E}(\sigma_i Z_{it}) = 0 \tag{10}$$

and

$$\mathrm{Var}(R_{it}) = \mathrm{Var}\left(\boldsymbol{\beta_i}^\top \boldsymbol{F_t}\right) + \mathrm{Var}(\sigma_i Z_{it}) = \boldsymbol{\beta_i}^\top \boldsymbol{\Sigma_F} \boldsymbol{\beta_i} + \sigma_i^2 \sigma_{Zi}^2. \tag{11}$$

Notably, Eq. 11 highlights the model's assumption of constant variance across time. Therefore, conditional heteroskedasticities, such as volatility clustering in periods of high uncertainty, cannot be accounted for. Nonetheless, DFM remains a versatile and powerful tool with broad applicability across various domains. For instance, it has been effectively used in nowcasting economic indicators, demonstrating efficiency and accuracy in providing timely insights crucial for policymakers, financial analysts, and other stakeholders involved in decision-making processes (Luciani et al. 2018; Hayashi et al. 2022; Chernis et al. 2020). Beyond nowcasting, the DFM is also extensively applied in business cycle, inflation dynamics, and structural analysis (Stock and Watson 1999; Boivin and Giannoni 2006). In contrast to these predictive applications, this study focuses on analyzing the extracted loadings, common factors, and their evolution processes along with their economic interpretations as an alternative framework. It aims to understand the complex dynamics of stock price movements, complementing established market theories.

Model fitting

One approach to fitting the DFM is to formulate the model as a state-space model and apply the Kalman method and maximum likelihood estimation. The mathematical details of the fitting methodology are presented as follows:

Suppose $\{\boldsymbol{Y_t}\}$ is an observed time series process. A linear Gaussian state-space model assumes the following dynamics:

$$\boldsymbol{Y_t} = \boldsymbol{M}\boldsymbol{X_t} + \boldsymbol{\epsilon_t}, \tag{12}$$

$$\boldsymbol{X_t} = \boldsymbol{T}\boldsymbol{X_{t-1}} + \boldsymbol{\eta_t}, \tag{13}$$

where $\{\boldsymbol{X_t}\}$ is an unobserved latent factor process, $\boldsymbol{M}$ is the measurement loading matrix, $\boldsymbol{T}$ is the transition loading matrix, and $\boldsymbol{\epsilon_t} \sim N(\mathbf{0}, \boldsymbol{\Sigma_\epsilon})$, $\boldsymbol{\eta_t} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma_\eta}\right)$ are Gaussian noise processes.

Denote the following for convenience:

$$\begin{aligned}
\boldsymbol{R_t} &:= \begin{bmatrix} R_{1t} \\ R_{2t} \\ \vdots \\ R_{St} \end{bmatrix}, \\
\boldsymbol{\beta} &:= \begin{bmatrix} \boldsymbol{\beta_1}^\top \\ \boldsymbol{\beta_2}^\top \\ \vdots \\ \boldsymbol{\beta_S}^\top \end{bmatrix}, \\
\boldsymbol{\sigma} &:= \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_S \end{bmatrix}, \\
\boldsymbol{\tilde{Z}_t} &:= \begin{bmatrix} \tilde{Z}_{1t} \\ \tilde{Z}_{2t} \\ \vdots \\ \tilde{Z}_{St} \end{bmatrix}, \\
\boldsymbol{\Psi_j} &:= \begin{bmatrix} \psi_{1j} & 0 & \cdots & 0 \\ 0 & \psi_{2j} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_{Sj} \end{bmatrix},
\end{aligned} \tag{14}$$

where $\boldsymbol{R_t}$ is the vector of S stock returns at time t, $\boldsymbol{\beta}$ is the combined loadings matrix of all S stocks for the common factors, $\boldsymbol{\sigma}$ is the diagonal loadings matrix of all S stocks for the stock-specific factors, $\boldsymbol{\tilde{Z}_t}$ is the vector of $\tilde{Z}_{it} := \sigma_i Z_{it}$, and $\boldsymbol{\Psi_j}$ is the diagonal matrix containing the jth AR coefficients of every $Z_{i(t-j)}$. The DFM can be formulated as a state-space model as follows:

$$\underbrace{R_t}_{Y_t} = \underbrace{\begin{bmatrix}\beta & 0 & \cdots & 0 & I_s & 0 & \cdots & 0\end{bmatrix}}_{M} \underbrace{\begin{bmatrix} F_t \\ F_{t-1} \\ \vdots \\ F_{t-p+1} \\ \tilde{Z}_t \\ \tilde{Z}_{t-1} \\ \vdots \\ \tilde{Z}_{t-q+1} \end{bmatrix}}_{X_t} + \epsilon_t, \tag{15}$$

$$\underbrace{\begin{bmatrix} F_t \\ \vdots \\ F_{t-p+2} \\ F_{t-p+1} \\ \tilde{Z}_t \\ \vdots \\ \tilde{Z}_{t-q+2} \\ \tilde{Z}_{t-q+1} \end{bmatrix}}_{X_t} = \underbrace{\begin{bmatrix} \Lambda_1 & \cdots & \Lambda_{p-1} & \Lambda_p & 0 & \cdots & 0 & 0 \\ I_n & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & I_n & 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & \Psi_1 & \cdots & \Psi_{q-1} & \Psi_q \\ 0 & \cdots & 0 & 0 & I_s & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & I_s & 0 \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} F_{t-1} \\ \vdots \\ F_{t-p+1} \\ F_{t-p} \\ \tilde{Z}_{t-1} \\ \vdots \\ \tilde{Z}_{t-q+1} \\ \tilde{Z}_{t-q} \end{bmatrix}}_{X_{t-1}} + \eta_t, \tag{16}$$

with $\Sigma_\epsilon = 0$ and $\Sigma_\eta = \mathrm{diag}(I_n, 0, \ldots, 0, \sigma^2, 0, \ldots, 0)$.

For a fixed set of parameters M, T, Σ_ϵ, and Σ_η, the Kalman filter may be used to estimate the state of the model X_t, and consequently the common factors F_t and stock-specific factors Z_{it}. For convenience, denote

$$Y_{1:t} := \{Y_1, Y_2, \ldots, Y_t\}, \tag{17}$$

$$\mu_{t|t'} := \mathbb{E}(X_t \mid Y_{1:t'}), \tag{18}$$

$$\Sigma_{t|t'} := \mathrm{Cov}(X_t \mid Y_{1:t'}). \tag{19}$$

The Kalman filter prediction step predicts the current state of the system as follows:

$$X_t \mid Y_{1:t-1} \sim N\left(\mu_{t|t-1}, \Sigma_{t|t-1}\right), \tag{20}$$

where

$$\mu_{t|t-1} = T\mu_{t-1|t-1}, \tag{21}$$

$$\Sigma_{t|t-1} = T\Sigma_{t-1|t-1}T^\top + \Sigma_\eta. \tag{22}$$

The Kalman filter update step then combines knowledge about the predicted state X_t with the new observation Y_t to produce an updated estimate of the current state of the system as $X_t \mid Y_{1:t} \sim N\left(\mu_{t|t}, \Sigma_{t|t}\right)$, where

$$K_t := \Sigma_{t|t-1}M^\top \left(M\Sigma_{t|t-1}M^\top + \Sigma_\epsilon\right)^{-1}, \tag{23}$$

$$\boldsymbol{\mu}_{t|t} = \boldsymbol{\mu}_{t|t-1} + \boldsymbol{K}_t\left(\boldsymbol{Y}_t - \boldsymbol{M}\boldsymbol{\mu}_{t|t-1}\right), \tag{24}$$

$$\boldsymbol{\Sigma}_{t|t} = \boldsymbol{\Sigma}_{t|t-1} - \boldsymbol{K}_t\boldsymbol{M}\boldsymbol{\Sigma}_{t|t-1}. \tag{25}$$

Hence, the Kalman filter uses past and current observations to estimate the current state of the system. The optimal parameters can then be obtained via maximum likelihood estimation.

Additionally, Kalman smoothing may be used at time $t' > t$ to combine knowledge about all observations until time t' to produce an updated estimate of the state of the system at time t as $\boldsymbol{X}_t \mid \boldsymbol{Y}_{1:t'} \sim N\left(\boldsymbol{\mu}_{t|t'}, \boldsymbol{\Sigma}_{t|t'}\right)$, where

$$\boldsymbol{J}_t := \boldsymbol{\Sigma}_{t|t}\boldsymbol{T}^\top \boldsymbol{\Sigma}_{t+1|t}{}^{-1}, \tag{26}$$

$$\boldsymbol{\mu}_{t|t'} = \boldsymbol{\mu}_{t|t} + \boldsymbol{J}_t\left(\boldsymbol{\mu}_{t+1|t'} - \boldsymbol{\mu}_{t+1|t}\right), \tag{27}$$

$$\boldsymbol{\Sigma}_{t|t'} = \boldsymbol{\Sigma}_{t|t} + \boldsymbol{J}_t\left(\boldsymbol{\Sigma}_{t+1|t'} - \boldsymbol{\Sigma}_{t+1|t}\right)\boldsymbol{J}_t{}^\top. \tag{28}$$

Unlike the Kalman filter, Kalman smoothing uses past, current, and future observations to estimate the current state of the system.

Model validation

The theoretical and empirical validity of the DFM relies on specifying the correct number of common factors n. Although previous research often set n based on prior knowledge and existing studies, Bai and Ng (2002) provided three information criteria as statistical measures to consistently estimate n from a given dataset. These criteria, extensively used in literature (Bai 2003; Stock and Watson 2002, 2016; Giglio et al. 2022), are expressed as follows:

$$\mathrm{IC}_1(n) = \ln V(n) + n\left(\frac{S+T}{ST}\right)\ln\left(\frac{ST}{S+T}\right), \tag{29}$$

$$\mathrm{IC}_2(n) = \ln V(n) + n\left(\frac{S+T}{ST}\right)\ln\min\{S, T\}, \tag{30}$$

$$\mathrm{IC}_3(n) = \ln V(n) + n\left(\frac{\ln\min\{S, T\}}{\min\{S, T\}}\right), \tag{31}$$

where $V(n)$ is the mean of the squared residuals when PCA is used to estimate n common factors. In contrast, Onatski (2009) proposed a hypothesis testing procedure for determining the number of common factors n. More recently, Molero-González et al. (2023) provided an alternative method based on random matrix theory (RMT).

Additionally, following the approach in Molero-González et al. (2023), additional analyses can be performed to assess the model's alignment with established market theories. Examining the relationship between common factors $\boldsymbol{F}_t$ and a market's composite index may reveal the model's ability to capture systematic market movements and distinguish idiosyncratic components of stock price dynamics. Moreover, analyzing the correlation between factor loadings $\boldsymbol{\beta_i}$ and the CAPM beta, as defined in Eq. 1, offers insights into how effectively the model reflects a conventional measure of systematic risk exposure. By selecting an appropriate number of factors n with corresponding loadings and common factors that align with market theories, the empirical validity of the DFM can be effectively demonstrated.

Model implementation

The DFM fitting and validation procedures are implemented in the `DynamicFactorAnalysis` Python package. Other common data science libraries were also used in the implementation.

Results

To reiterate, this study analyzes the extracted loadings $\boldsymbol{\beta_i}$ and common factors $\boldsymbol{F_t}$ to offer an alternative perspective on the dynamics of stock price movements, distinguishing itself from recent developments in literature, particularly that of Molero-González et al. (2023), which did not provide a subsequent analysis into $\boldsymbol{\beta_i}$ and $\boldsymbol{F_t}$ after using RMT to determine the number of common factors underlying stock price dynamics. Whereas Molero-González et al. (2023) primarily focused on factor dimensionality, the results of this study specifically analyzed and interpreted $\boldsymbol{\beta_i}$ and $\boldsymbol{F_t}$ in relation to the broader framework of known econometric and market facts.

To this end, the PSE is considered owing to its distinct economic landscape and investor behavior. This approach broadens insights into stock price dynamics in a unique economic context while demonstrating the robustness of the DFM in providing insights aligned with established market theories. The historical stock price data used in this model are obtained using the Python library `fastquant`, which wraps the data request process using the Phisix API.[1] Additionally, data for the PSE index (PSEi) is obtained from https://stooq.com.

The period from January 1, 2015, to December 31, 2020 is considered. Following the data cleaning procedure of existing works (Neszveda 2025; Feder-Sempach et al. 2024; Feng 2019), stocks with over 1% missing observations were excluded from the analysis to ensure data integrity. This results in a dataset comprising 72 stocks, exceeding the top 30 in terms of market capitalization included in the PSEi. Given the objective of the study, the aforementioned procedure would result in the exclusion of generally low-volume stocks, which are likely to make a relatively smaller contribution to overall market dynamics. Consequently, the remaining high-volume stocks would carry greater weight in the analysis, guaranteeing generalizability despite the data cleaning procedure.

Notably, although the COVID-19 pandemic introduces significant volatility to the data, demonstrating the ability of the model to provide robust insights into the market even in the presence of extreme events is an important aspect of the present study. Thus, the inclusion of the pandemic period provides a rigorous test case to evaluate the

[1] https://github.com/phisix-org/phisix.

robustness of the model under such unprecedented conditions, in line with Molero-González et al. (2023).

The percentage return R_{it} is considered to ensure the stationarity of the data. If S_{it} is the closing price of stock i at time t, the percentage return is obtained as follows:

$$R_{it} = \frac{S_{it} - S_{i(t-1)}}{S_{i(t-1)}}. \tag{32}$$

Information criteria in Eqs. 29, 30, and 31 across different values of n are then calculated and presented in Fig. 1. The information criteria are lower for models with $n = 1$ or $n = 2$ factors than for a white noise model containing zero factors. This result validates the choice of a factor model as it suggests that the inclusion of at least one factor considerably improves the fit with the data. Although an analysis based entirely on the information criteria would indicate that a model with $n = 1$ common factor fits the data best, a model with $n = 2$ common factors also exhibits a comparably close fit, making it a viable alternative for consideration.

One-factor model

First, the model with $n = 1$ common factor following an AR(3) process and the stock-specific factors following AR(5) processes is considered, where p and q are chosen based on the Bayesian information criterion (BIC). As a validation, the common factor F_t is compared with the return of PSEi, as shown in Fig. 2. Composite indices generally contain a diversified portfolio of stocks in a particular market. This diversification removes stock-related movements and risks, leaving only systematic movements and risks. Consequently, the PSEi is commonly used as a proxy for systematic market movements. The (Kalman-smoothed) common factor F_t and the PSEi returns exhibit a correlation of 0.9283, establishing the former as a viable indicator of systematic market movements. Nevertheless, in contrast to the PSEi, which only covers the top 30 stocks with the largest market capitalization, the common factor F_t captures the systematic movements across the broader market, reflecting the overall

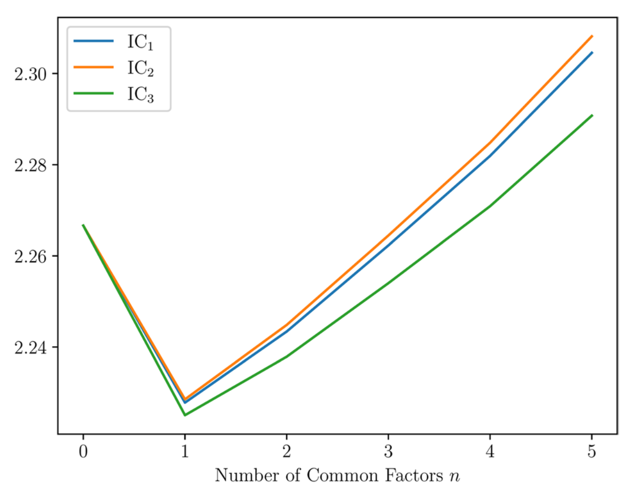

Fig. 1 Information criteria for different numbers of common factors n

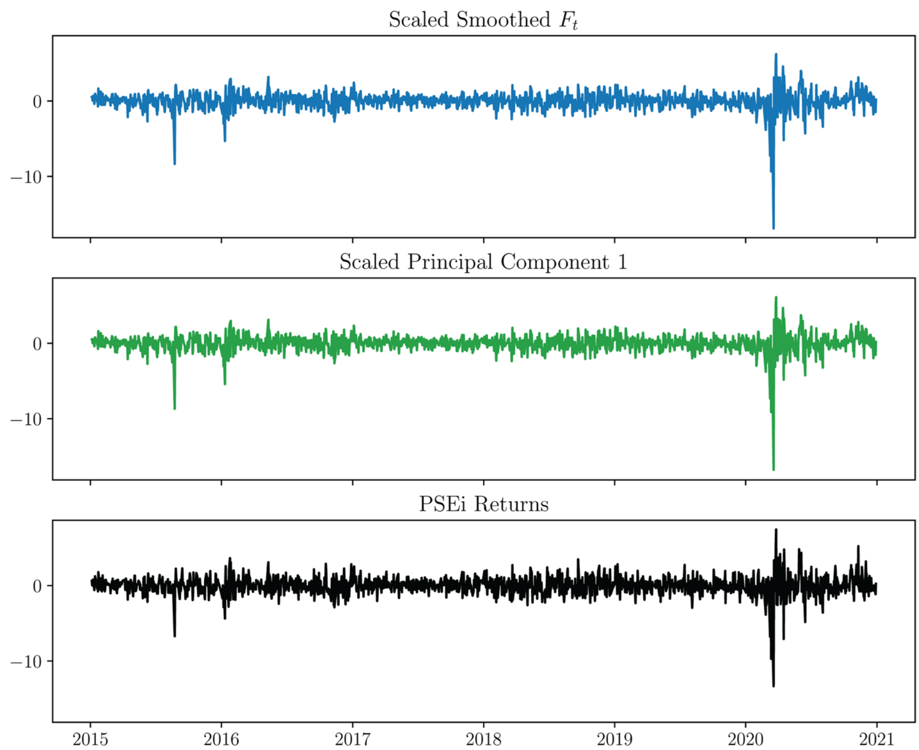

Fig. 2 The common factor F_t for DFM ($n = 1, p = 3, q = 5$), the first principal component, and the PSEi return from 2015 to 2020

stock price movements. This accounts for the discrepancy between the common factor and the PSEi and explains why their correlation coefficient falls short of a perfect association.

Nonetheless, the common factor F_t remains versatile in capturing market conditions at various time points. Notably, it accurately reflects significant market events, such as the sharp downturn on August 24, 2015, which resulted from the global financial market sell-offs owing to concerns about China's economy. It also reflects the high economic volatility during the first two quarters of 2020, when the Philippine economy underwent a lockdown because of the COVID-19 pandemic. The common factor remains informative in reflecting the challenges during the subsequent recovery process. Other similar observations, such as volatility clustering during times of high economic uncertainty, can also be noted. This finding further supports that the common factor F_t represents systematic market movements.

The common factor F_t is estimated to evolve according to

$$F_t = 0.1256F_{t-1} + 0.0225F_{t-2} + 0.1380F_{t-3} + \varepsilon_t, \tag{33}$$

where only the parameter $\Lambda_2 = 0.0225$ is not statistically significant. This indicates that systematic market shocks are expected to persist for at least three trading days. Such persistence may reveal important aspects of market behavior. For example, policy-related shocks may influence the market over an extended period, indicating the need for caution when announcing or implementing policies that are expected to impact the market (Li et al. 2010; Chatziantoniou et al. 2013).

Table 1 Summary of correlation figures for DFM ($n = 1, p = 3, q = 5$)

Series 1	Series 2	Correlation
CAPM β_i	β_i	0.8348
F_t	PSEi	0.9283
F_t	PC1	0.9975
PC1	PSEi	0.9144

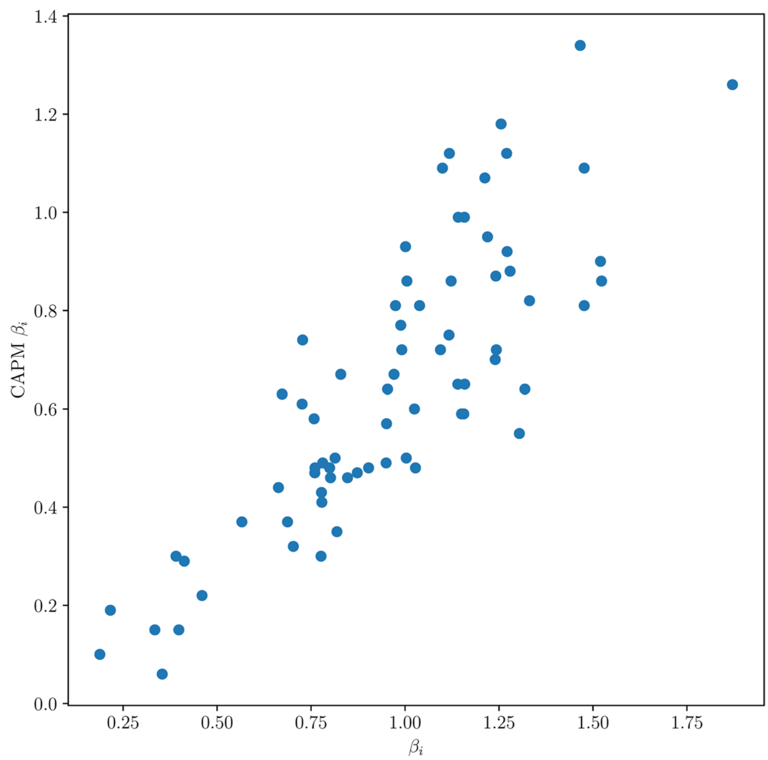

Fig. 3 Scatterplot of the loading β_i for DFM ($n = 1, p = 3, q = 5$) and the CAPM beta

Moreover, PCA is investigated for comparison. Table 1 presents a summary of the correlation figures. The (Kalman-smoothed) common factor F_t exhibits a correlation of 0.9975 with the first principal component, as illustrated in Fig. 2. This principal component also has a strong correlation of 0.9144 with the PSEi returns. Although these results may indicate that DFM performs similarly to PCA in explaining systematic stock price movements, it must be noted that the two models are fundamentally distinct. PCA derives principal components as a linear combination of stock price returns, whereas DFM explains stock price returns as a linear combination of some underlying common factors. The latter is also a time series model that accounts for the dynamics of stock price returns across time. Moreover, the first principal component will remain constant when considering $n > 1$ factors, whereas the common factors $\boldsymbol{F_t}$ will adapt depending on the model specifications. Hence, DFM can better capture the dynamics of stock price movements than PCA.

Table 2 Loadings for DFM ($n = 1, p = 3, q = 5$)

Stock	β_i	σ_i
2GO	0.8470	3.9217
ABA	0.9707	3.2377
AC	1.2704	1.4373
AEV	1.0996	1.8298
AGI	1.2715	1.7219
ALI	1.8711	1.6140
AP	0.7276	1.5130
BDO	1.2555	1.4833
BEL	0.8189	1.6531
BLOOM	1.5229	2.7239
BPI	1.0013	1.4661
BRN	1.2399	3.3060
CEB	1.3312	2.4174
CHIB	0.4127	0.9347
CNPF	0.6633	1.7277
COSCO	0.7777	1.4535
CPG	0.8730	2.2062
DD	0.7787	2.4391
DMC	1.2795	1.8911
DNL	1.1504	2.2307
EEI	0.8018	2.1847
EW	1.0251	1.8732
FGEN	0.8139	1.8883
FLI	1.0034	1.5730
FNI	1.2429	6.0448
FPH	0.7029	1.2352
GERI	1.0278	2.1231
GLO	0.7262	1.8556
GMA7	0.3908	1.3900
GTCAP	1.2195	1.8579
HOUSE	0.3541	1.8103
ICT	1.1416	1.8282
IMI	0.9508	2.4919
JFC	1.1580	1.7656
JGS	1.4660	1.8076
LC	0.1881^{ns}	3.0327
LPZ	0.7766	2.0411
LTG	0.9536	2.1000
MAXS	1.1587	2.0482
MBT	1.1226	1.5110
MEG	1.5200	1.7571
MER	0.6730	1.5231
MPI	1.2415	1.9402
MWC	0.7810	2.2912
MWIDE	0.9886	2.3709
NI	0.3982	2.5281
NIKL	1.3189	3.6951
PCOR	0.9497	2.1039
PGOLD	0.7605	1.6123

Table 2 (continued)

Stock	β_i	σ_i
PLC	1.1563	2.4863
PNB	0.7994	1.5993
PNX	0.4600	1.9619
PX	0.5659	2.4845
PXP	1.3042	5.2152
RLC	1.4767	1.9403
RRHI	0.7599	1.7353
SCC	0.9913	2.8519
SECB	1.0943	1.7228
SLI	0.3344	2.0121
SM	1.1180	1.6795
SMC	0.7581	1.7769
SMPH	1.2124	1.5934
SSI	1.4767	2.7815
SSP	0.9031	3.6809
STI	0.6874	2.6140
TECH	0.9746	3.7055
TEL	0.8290	1.9001
UBP	0.2162	1.0130
URC	1.0050	1.7758
VITA	1.1404	3.6114
VLL	1.1170	2.1362
WEB	1.0387	4.8616

The loading β_i is compared against its corresponding CAPM beta for further model validation, which is widely accepted as a measure of exposure to market risks. Figure 3 presents the scatterplot between the loading β_i and the CAPM beta.[2] The figure indicates a strong positive correlation between the two measures, with a correlation of 0.8348, validating the relationship of β_i and exposure to systematic movements. Table 2 subsequently presents the summary of the loading β_i and σ_i for the 72 stocks included in the analysis. At a significance level of $\alpha = 0.05$, the loading β_i for the common factor F_t is statistically significant for all but one stock.[3] This result highlights that the common factor F_t explains price movements in the market, capturing the systematic or market dynamics similar to the PSEi. More mature and developed stocks, such as AC, GTCAP, JFC, SM, and URC, have relatively balanced β_i and σ_i values. Conversely, less mature or more volatile stocks, such as ABA, BRN, FNI, PXP, and SSP, have relatively higher σ_i values, indicating more significant contribution from idiosyncratic movements owing to stock-specific shocks. Hence, the ratio between β_i and σ_i may also indicate relative sensitivity to market-specific or stock-specific volatilities.

[2] https://www.barrons.com/market-data/stocks.

[3] The loading β_i for stock LC is not statistically significant due to a significant stock-specific shock—the sizable closure order against its operations (https://denr.gov.ph/news-events/lopez-orders-closure-of-23-metallic-mines/, Catajan 2021)—which lasted from 2017 to 2020. This case is expected to have a pronounced impact on its financial outlook, thereby overshadowing general market conditions.

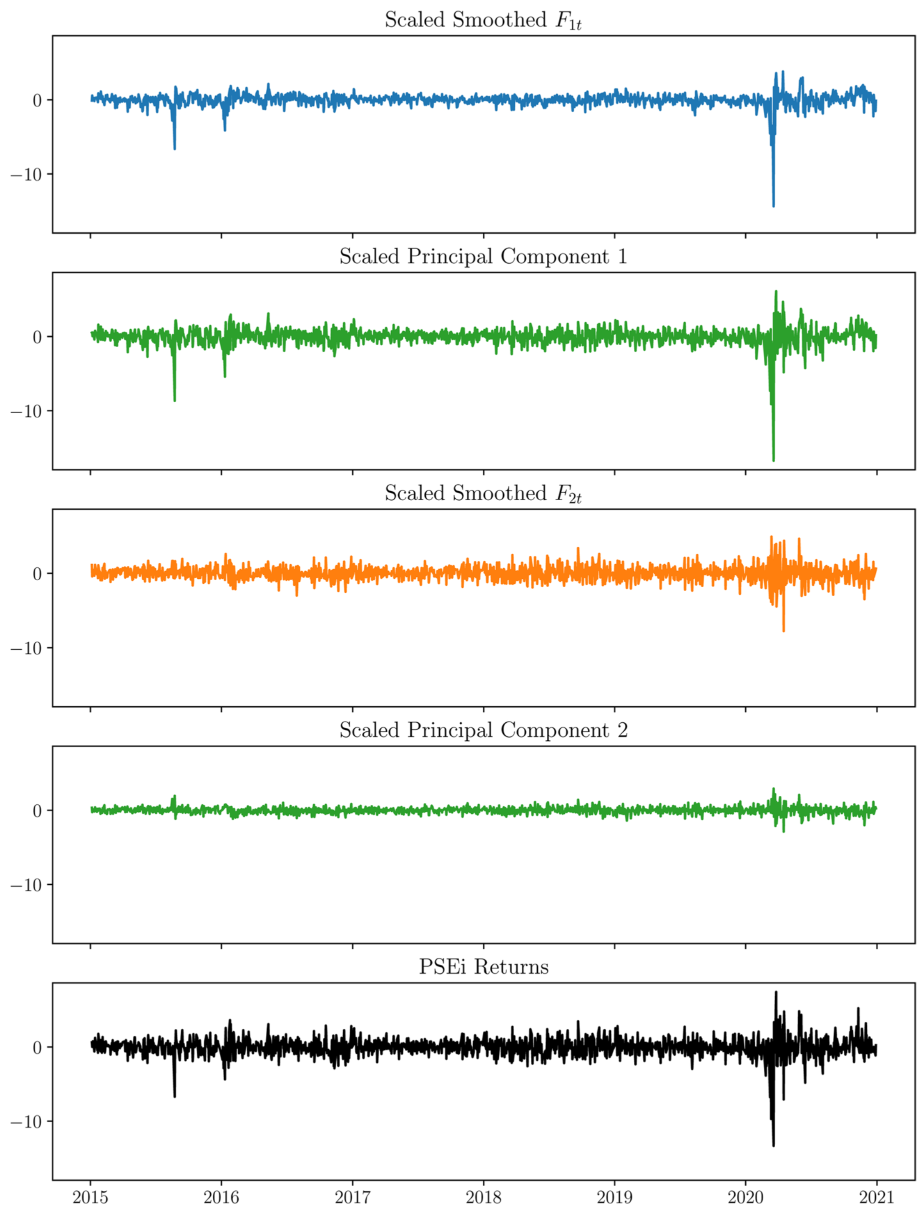

Fig. 4 The common factors $\boldsymbol{F_t}$ for DFM ($n = 2, p = 2, q = 5$), the first two principal components, and the PSEi return from 2015 to 2020

Table 3 Summary of correlation figures for DFM ($n = 2, p = 2, q = 5$)

Series 1	Series 2	Correlation
CAPM β_i	β_{1i}	0.5132
F_{1t}	PSEi	0.6727
F_{1t}	PC1	0.8982
PC1	PSEi	0.9144
CAPM β_i	β_{2i}	0.8528
F_{2t}	PSEi	0.7704
F_{2t}	PC2	0.8229
PC2	PSEi	0.3498

Two-factor model

Next, the model with $n = 2$ common factors following a VAR(2) process and stock-specific factors following AR(5) processes is considered, where p and q are chosen based on the BIC. The (Kalman-smoothed) common factors $\boldsymbol{F_t}$ are also compared against the PSEi returns and principal components in Fig. 4. The correlation figures are presented in Table 3. The first common factor F_{1t} has correlations of 0.6727 and 0.8982 with the PSEi returns and the first principal component, respectively. The second common factor F_{2t} has correlations of 0.7704 and 0.8229 with the PSEi returns and the second principal component, respectively. The first and second principal components have correlations of 0.9144 and 0.3498 with the PSEi returns, respectively. The common factors $\boldsymbol{F_t}$ of the DFM adapt based on the model specification, whereas PCA remains static regardless of the number of factors considered. This result may be observed from the common factor F_{1t} in Fig. 4 that deviates from the common factor F_t in Fig. 2 while still maintaining the overall trend. This highlights the distinctive advantage of DFM over PCA in determining systematic movements in the PSE. PCA requires that the principal components maximize variance while remaining mutually orthogonal. Conversely, DFM does not explicitly impose such restrictions, thereby offering greater flexibility.

Exploring common factors in greater detail, it is worth noting that F_{1t} and F_{2t} have a correlation of 0.0798, which is nearly orthogonal. A linear regression of F_t on F_{1t} and F_{2t} also produced an R^2 of 0.998. These results indicate that the model further decomposed F_t into two nearly uncorrelated signals, namely, F_{1t} and F_{2t}. A visual inspection of Fig. 4 reveals that F_{1t} represents the broader market trend, whereas F_{2t} represents market uncertainties independent of the general market direction. This further allows for the following interpretations: β_{1i} as exposure to market trends, β_{2i} as exposure to market volatility, and σ_i as exposure to stock-specific volatility. This provides a new perspective on portfolio risk management beyond the traditional CAPM framework, as investors can now account for two dimensions of market movements.

The common factors $\boldsymbol{F_t} = [F_{1t}, F_{2t}]^\top$ are then estimated to evolve according to

$$\boldsymbol{F_t} = \begin{bmatrix} 0.4894 & 0.1680 \\ -0.0262 & -0.2526 \end{bmatrix} \boldsymbol{F_{t-1}} + \begin{bmatrix} 0.5094 & -0.1726 \\ -0.1627 & -0.0931 \end{bmatrix} \boldsymbol{F_{t-2}} + \boldsymbol{\varepsilon_t}, \tag{34}$$

where only the coefficients −0.0262 and −0.0931 are not statistically significant. Similarly, the results also indicate the persistence of systematic shocks in the market (Gil-Alana et al. 2023).

For validation, the loadings β_{1i} and β_{2i} are also compared against the CAPM beta, as presented in Fig. 5. The CAPM beta exhibited correlations of 0.5132 and 0.8528 with β_{1i} and β_{2i}, respectively. The lower correlation with β_{1i} indicates that the first common factor F_{1t} captures market dynamics not considered by the CAPM. Meanwhile, the higher correlation between β_{2i} and the CAPM beta aligns with the interpretation of the second common factor F_{2t} capturing market volatility, a key aspect that the CAPM also emphasizes. Despite these differences, the common factors F_{1t} and F_{2t} significantly contribute to the explanation of systematic stock price movements, highlighting the ability of the model to capture market trends and additional specific sources of volatility. Table 4 presents the summary of the loadings β_{1i}, β_{2i}, and σ_i. The β_{1i}'s are statistically significant for all stocks, whereas half of the β_{2i}'s are statistically significant at $\alpha = 0.05$. This indicates

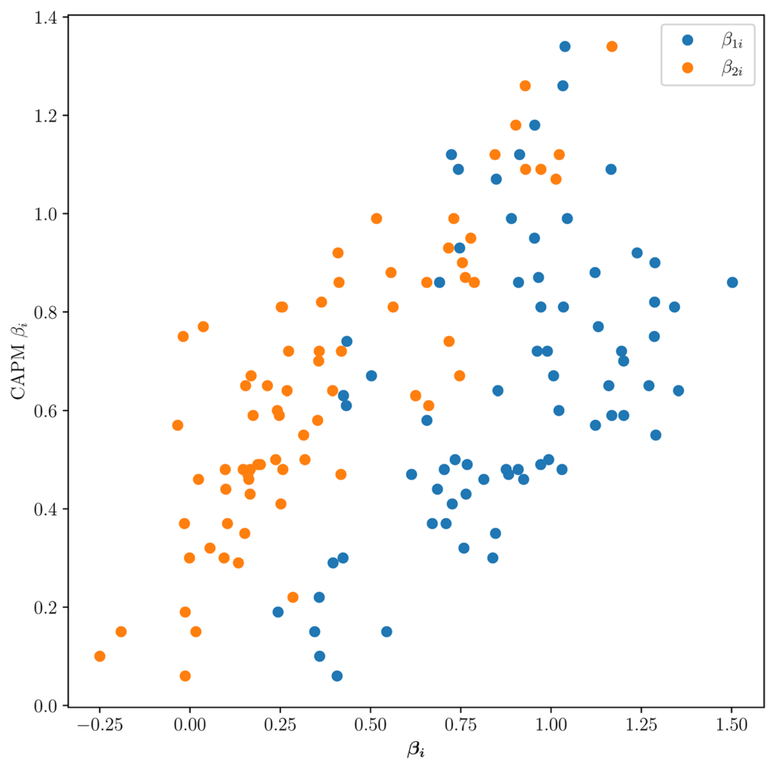

Fig. 5 Scatterplot of the loadings $\boldsymbol{\beta_i}$ for DFM ($n = 2, p = 2, q = 5$) and the CAPM beta

that the two-factor model captures additional variance in the data, highlighting how the common factors obtained are viable indicators of systematic market movements.

In summary, the results from DFM ($n = 1$, $p = 3$, $q = 5$) and DFM ($n = 2$, $p = 2$, $q = 5$) align with established theories and provide a new and alternative understanding of price movement dynamics. The relationship between the common factors $\boldsymbol{F_t}$ and the PSEi returns provides unique insights into how these models quantify systematic market movements and idiosyncratic movements. Notably, DFM ($n = 2$, $p = 2$, $q = 5$) offers a more nuanced characterization of market dynamics, which complements established portfolio risk management theories by decomposing market trends and market volatility. Furthermore, the correlation between the factor loadings $\boldsymbol{\beta_i}$ and the CAPM beta highlights that the models closely capture the conventional measure of systematic risk exposure. Overall, the results demonstrate that dynamic factor analysis can provide novel insights into classical market theories.

Nowcasting GDP application

The above results indicate that the common factors $\boldsymbol{F_t}$ are viable real-time market indicators that can be effectively extracted from real-time stock price returns data. These factors can be extended to various economic and financial applications. This subsection illustrates the utility of the common factors $\boldsymbol{F_t}$ of DFM ($n = 2, p = 2, q = 5$) within an economic context.

A major limitation faced by economic leaders is the substantial delay in releasing key economic indicators. For instance, policymakers rely on the quarterly GDP

Table 4 Loadings for DFM ($n = 2, p = 2, q = 5$)

Stock	β_{1i}	β_{2i}	σ_i
2GO	0.9241	0.0236[ns]	3.8987
ABA	1.0070	0.1693[ns]	3.2195
AC	0.9129	0.8448	1.3988
AEV	0.7433	0.9720	1.7430
AGI	1.2384	0.4100	1.7036
ALI	1.0328	0.9285	1.5250
AP	0.4345	0.7176	1.4191
BDO	0.9548	0.9023	1.4394
BEL	0.8462	0.1517	1.6286
BLOOM	1.5024	0.4127	2.6979
BPI	0.7470	0.7162	1.4348
BRN	1.2011	0.3564	3.2954
CEB	1.2868	0.3642	2.3957
CHIB	0.3963	0.1341	0.9321
CNPF	0.6854	0.0993[ns]	1.7111
COSCO	0.7646	0.1666	1.4362
CPG	0.8825	0.1603[ns]	2.1836
DD	0.7264	0.2518	2.4382
DMC	1.1218	0.5568	1.8945
DNL	1.2014	0.1748[ns]	2.1834
EEI	0.8138	0.1628[ns]	2.1700
EW	1.0218	0.2420	1.8502
FGEN	0.7343	0.3185	1.8883
FLI	0.9935	0.2374	1.5466
FNI	1.1952	0.3580[ns]	6.0388
FPH	0.7587	0.0554[ns]	1.1902
GERI	1.0304	0.1669[ns]	2.0919
GLO	0.4331	0.6612	1.8105
GMA7	0.4239	− 0.0013[ns]	1.3740
GTCAP	0.9541	0.7777	1.8355
HOUSE	0.4075	− 0.0132[ns]	1.7982
ICT	0.8903	0.7307	1.8127
IMI	1.1230	− 0.0340[ns]	2.4140
JFC	1.0451	0.5168	1.7643
JGS	1.0386	1.1688	1.7109
LC	0.3592	− 0.2497[ns]	3.0048
LPZ	0.8387	0.0945[ns]	2.0117
LTG	0.8529	0.3950	2.0992
MAXS	1.2711	0.1540[ns]	1.9789
MBT	0.9099	0.6560	1.5007
MEG	1.2877	0.7545	1.7593
MER	0.4251	0.6250	1.4782
MPI	0.9654	0.7626	1.9291
MWC	0.7678	0.1878[ns]	2.2833
MWIDE	1.1309	0.0369[ns]	2.3085
NI	0.5448	− 0.1908[ns]	2.4880
NIKL	1.3530	0.2689[ns]	3.6684
PCOR	0.9715	0.1949[ns]	2.0778

Table 4 (continued)

Stock	β_{1i}	β_{2i}	σ_i
PGOLD	0.7043	0.2572	1.6095
PLC	1.1682	0.2474	2.4584
PNB	0.8758	0.0977^{ns}	1.5576
PNX	0.3580	0.2852	1.9598
PX	0.6711	− 0.0153^{ns}	2.4571
PXP	1.2902	0.3148^{ns}	5.2036
RLC	1.1662	0.9296	1.9170
RRHI	0.6137	0.4181	1.7350
SCC	0.9616	0.2731^{ns}	2.8422
SECB	0.9898	0.4194	1.7240
SLI	0.3454	0.0164^{ns}	2.0057
SM	0.7241	1.0227	1.5732
SMC	0.6562	0.3536	1.7768
SMPH	0.8483	1.0137	1.5017
SSI	1.3420	0.5626	2.7782
SSP	0.9091	0.1473^{ns}	3.6650
STI	0.7093	0.1038^{ns}	2.6003
TECH	0.9719	0.2527^{ns}	3.6977
TEL	0.5027	0.7467	1.8468
UBP	0.2440	− 0.0131^{ns}	1.0038
URC	0.6913	0.7876	1.7318
VITA	1.1600	0.2145^{ns}	3.5918
VLL	1.2861	− 0.0188^{ns}	2.0218
WEB	1.0343	0.2561^{ns}	4.8515

growth rates as the primary indicator of economic performance, as it measures the total monetary value of all goods and services produced in a country over the specific period. However, this figure is typically released several weeks following the end of a quarter. Fiscal and monetary policies must be enacted based on incomplete information during this period. Many institutions have developed nowcasting models to predict these economic indicators to aid economic decisions. Hence, inspired by previous works (Luciani et al. 2018; Hayashi et al. 2022; Babii et al. 2022; Ashwin et al. 2021; Giannone et al. 2005), this subsection considers the problem of nowcasting the GDP growth rates.

As a benchmark, the AR(1) model is considered. The in-sample period for model training is from January 1, 2015 to December 31, 2020. The period from January 1, 2021 to December 31, 2022 is then treated as the out-of-sample period for model evaluation. Figure 6 presents the nowcasts of the model, whereas Table 5 presents a summary of the root mean squared error (RMSE) figures. The model obtained an RMSE of 3.8533 and 12.3095 during the in- and out-of-sample periods, respectively. Notably, the out-of-sample RMSE is expected to be greater owing to the more significant fluctuations in GDP growth rates caused by the COVID-19 pandemic.

To demonstrate the utility of the DFM, the common factors $\boldsymbol{F_t}$ are first transformed into monthly indicators to adjust for the difference in frequency with the

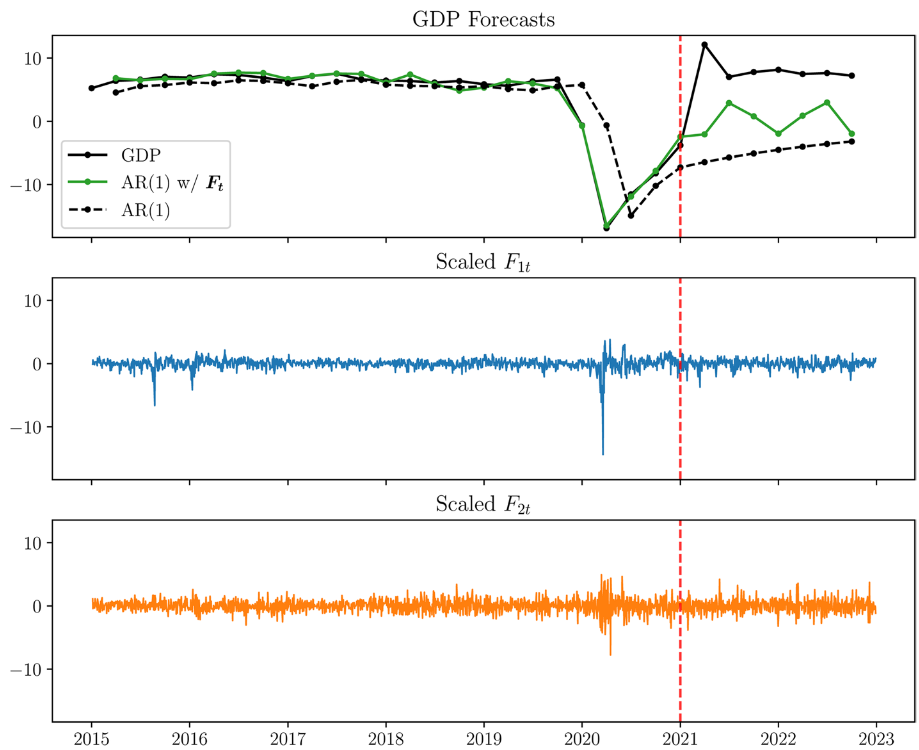

Fig. 6 Philippine GDP growth rate nowcasts of AR(1) and AR(1) with F_t

Table 5 Philippine GDP growth rate nowcasts RMSE values of AR(1) and AR(1) with F_t

Model	In-sample	Out-of-sample
AR(1)	3.8533	12.3095
AR(1) w/ F_t	0.6122	8.0667

GDP releases. The common factors $\boldsymbol{F_t}$ are the Kalman-smoothed and Kalman-filtered common factors during the in-sample and out-of-sample periods, respectively. The mean and standard deviation of the common factors F_{1t} and F_{2t} are calculated. These monthly indicators are then integrated into the AR(1) model via ordinary least squares regression to explain the GDP growth rate for the quarter. For example, the common factors $\boldsymbol{F_t}$ for January, February, and March are used to explain the GDP growth rate for the first quarter. Figure 6 presents the nowcasts of the model, whereas Table 5 presents a summary of the RMSE figures. The model obtained RMSE values of 0.6122 and 8.0667 during the in-sample and out-of-sample periods, respectively. The inclusion of the derived monthly indicators, constructed from the common factors, substantially improved the in-sample RMSE by 84.11% and the out-of-sample RMSE by 34.47%, demonstrating the utility of the model. This result further supports how the common factors $\boldsymbol{F_t}$ may be used as real-time market indicators.

Conclusion

In summary, this study explores DFM to analyze price movements in the PSE, integrating the predictive capabilities of machine learning models with the interpretability of traditional linear asset pricing models, thereby effectively bridging econometric theory and financial practice. Specifically, it focuses on the extracted loadings and common factors to provide alternative perspectives for understanding the dynamics of price movements. The results of a validation analysis with the CAPM reveal novel insights into market phenomena. Similar to the composite index, the one-factor model closely captures systematic or market dynamics, whereas the two-factor model further decomposes it into market trends and volatility, providing novel perspectives beyond conventional portfolio risk management theories. Additionally, an application on nowcasting GDP growth rates demonstrates the viability of the common factors as real-time market indicators in economic and financial applications by providing substantial performance improvements. These results demonstrate the value of dynamic factor analysis in providing a deeper understanding of price movements in the market.

Future studies may build on current findings by relaxing the DFM assumptions in Gaussian stock price returns and the cross-sectional independence in the errors Z_{it}. Specifically, exploring deviations from normality in the form of heavy tails may be useful, as observed in empirical results (Peiró 1994; Li 2023). To this end, one may incorporate a generalized autoregressive conditional heteroskedasticity (GARCH) process for the common factors $\boldsymbol{F_t}$ to allow for time-varying conditional variances, leading to more accurate stock return distribution modeling. Although this would typically introduce additional layers of complexity to the model fitting methodology, they may be accommodated through approximate inferential methods using advances in Bayesian variational inference algorithms (Dayta et al. 2024). Additionally, replacing the Kalman filter for model fitting with variational inference may allow for higher orders of flexibility in terms of the factor distribution and other additional effects, such as GARCH terms and market or industry-level dynamics. These model improvements would better capture the intricate dynamics of stock price movements while maintaining predictive performance and interpretability.

Abbreviations

APT	Arbitrage pricing theory
AR	Autoregressive process
BIC	Bayesian information criterion
CAPM	Capital asset pricing model
DFM	Dynamic factor model
GARCH	Generalized autoregressive conditional heteroskedasticity
GDP	Gross domestic product
PCA	Principal component analysis
PSE	Philippine Stock Exchange
PSEi	Philippine Stock Exchange Index
RMSE	Root mean squared error
RMT	Random matrix theory
VAR	Vector autoregressive process

Author contributions

BGL, DD, and BRT contributed to the overall conceptualization, code development, data gathering, statistical analyses, and manuscript writing. RRT, LPDG, and KI contributed to the overall conceptualization, manuscript writing, and manuscript finalization. All authors have read and approved the final version of the manuscript.

Funding
The Japan Society for the Promotion of Science supported this study through Grants-in-Aid for Scientific Research Program (KAKENHI 18K19821).

Data availability
The datasets generated and analyzed during the current study are available in the `DynamicFactorAnalysis` repository accessible at https://github.com/briangodwinlim/DynamicFactorAnalysis.

Declarations

Conflict of interest
The authors have no conflict of interest to declare.

Received: 9 January 2024 Accepted: 15 October 2025
Published online: 08 January 2026

References

Ashwin J, Kalamara E, Saiz L (2021) Nowcasting Euro area GDP with news sentiment: a tale of two crises. J Appl Econ 39:887–905
Babii A, Ghysels E, Striaukas J (2022) Machine learning time series regressions with an application to nowcasting. J Bus Econ Stat 40:1094–1106
Bai J (2003) Inferential theory for factor models of large dimensions. Econometrica 71:135–171
Bai J, Ng S (2002) Determining the number of factors in approximate factor models. Econometrica 70:191–221
Barucci E, Fontana C (2017) Factor asset pricing models: CAPM and APT, 201–253. Springer, London
Black F, Scholes M (1973) The pricing of options and corporate liabilities. J Polit Econ 81:637–654
Blume ME, Friend I (1973) A new look at the capital asset pricing model. J Financ 28:19–33
Boivin J, Giannoni MP (2006) Has monetary policy become more effective? Rev Econ Stat 88:445–462
Catajan ME. Philippine mines continue unhampered 4 years after Gina's shutdown order. https://www.sunstar.com.ph/baguio/feature/philippine-mines-continue-unhampered-4-years-after-ginas-shutdown-order
Chatziantoniou I, Duffy D, Filis G (2013) Stock market response to monetary and fiscal policy shocks: multi-country evidence. Econ Model 30:754–769
Chernis T, Cheung C, Velasco G (2020) A three-frequency dynamic factor model for nowcasting Canadian provincial GDP growth. Int J Forecast 36:851–872
Chernov M, Ronald Gallant A, Ghysels E, Tauchen G (2003) Alternative models for stock price dynamics. J Econom 116:225–257 (Frontiers of financial econometrics and financial engineering)
Dayta DB, Kubo T, Ikeda K(2024) You only accept a single observation per variational iteration: a fast self-correcting algorithm for stochastic variational inference. SSRN:5062129
Department of Environment and Natural Resources. LOPEZ ORDERS CLOSURE OF 23 METALLIC MINES|DENR. https://denr.gov.ph/news-events/lopez-orders-closure-of-23-metallic-mines/
Elbannan M (2014) The capital asset pricing model: an overview of the theory. Int J Econ Financ 7:p216
Elshqirat DM (2019) An empirical examination of the arbitrage pricing theory: evidence from Jordan
Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. J Financ Econ 33:3–56
Feder-Sempach E, Szczepocki P, Bogolebska J (2024) Global uncertainty and potential shelters: gold, bitcoin, and currencies as weak and strong safe havens for main world stock markets. Financ Innov 10(1):67
Feng F et al (2019) Temporal relational ranking for stock prediction. ACM Trans Inf Syst 37:1–30
French J (2017) Macroeconomic forces and arbitrage pricing theory. J Comp Asian Dev 16:1–20
Geweke J (1977) The dynamic factor analysis of economic time series. In: Latent variables in socio-economic models
Ghorbani M, Chong EK (2020) Stock price prediction using principal components. PLoS ONE 15:e0230124
Giannone D, Reichlin L, Small DH (2005) Nowcasting GDP and inflation: the real-time informational content of macroeconomic data releases
Giglio S, Kelly B, Stroebel J (2021) Climate finance. Annu Rev Financ Econ 13:15–36
Giglio S, Kelly B, Xiu D (2022) Factor models, machine learning, and asset pricing. Annu Rev Financ Econ 14:337–368
Gil-Alana LA, Infante J, Martín-Valmayor MA (2023) Persistence and long run co-movements across stock market prices. Q Rev Econ Finance 89:347–357
Gunduz H (2021) An efficient stock market prediction model using hybrid feature reduction method based on variational autoencoders and recursive feature elimination. Financ Innov 7:28
Hayashi F, Tachi Y (2022) Nowcasting Japan's GDP. Empir Econ 64:1699–1735
Htun HH, Biehl M, Petkov N (2023) Survey of feature selection and extraction techniques for stock market prediction. Financ Innov 9:26
Htun HH, Biehl M, Petkov N (2024) Forecasting relative returns for s &p 500 stocks using machine learning. Financ Innov 10:118
Hull JC, Basu S (2016) Options, futures, and other derivatives. Pearson Education India, New Delhi
Jolliffe I (2002) Principal component analysis. Springer, New York
Jolliffe I, Cadima J (2016) Principal component analysis: a review and recent developments. Philos Trans R Soc A Math Phys Eng Sci 374:20150202
Kendall MG (1953) The analysis of economic time-series—part I: prices. J R Stat Soc Ser A (Gen) 116:11–25
Kothari S, Zimmerman JL (1995) Price and return models. J Account Econ 20:155–192

Kumari B, Patnaik S, Swarnkar T (2023) Feature selection for stock price prediction: a critical review. Int J Intell Enterp 10:48–72
Lawi A, Mesra H, Amir S (2022) Implementation of long short-term memory and gated recurrent units on grouped time-series data to predict stock prices accurately. J Big Data 9:89
Li B (2023) An explanation for the distribution characteristics of stock returns. ArXiv:2312.02472
Li YD, İşcan TB, Xu K (2010) The impact of monetary policy shocks on stock prices: evidence from Canada and the United States. J Int Money Financ 29:876–896
Li M, Zhu Y, Shen Y, Angelova M (2023) Clustering-enhanced stock price prediction using deep learning. World Wide Web 26:207–232
Lim BG, Ong HJ, Tan RR, Ikeda K (2023) Dynamic principal component analysis for the construction of high-frequency economic indicators. In Proceedings of the 4th International Conference on Advances in Computational Science and Engineering. Springer Nature Singapore, Singapore, pp. 645–663
Lintner J (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Rev Econ Stat 47:13–37
Long W, Lu Z, Cui L (2019) Deep learning-based feature engineering for stock price movement prediction. Knowl-Based Syst 164:163–173
Luciani M, Pundit M, Ramayandi A, Veronese G (2018) Nowcasting Indonesia. Empir Econ 55:597–619
Marathe RR, Ryan SM (2005) On the validity of the geometric Brownian motion assumption. Eng Econ 50:159–192
Molero-González L, Trinidad-Segovia J, Sánchez-Granero M, García-Medina A (2023) Market beta is not dead: an approach from random matrix theory. Financ Res Lett 55:103816
Muhammad T et al (2023) Transformer-based deep learning model for stock price prediction: a case study on Bangladesh stock market. Int J Comput Intell Appl 0:2350013
Neszveda G (2025) Aspiration level, probability of success, and stock returns: an empirical test. Financ Innov 11(1):91
Officer RR (1972) The distribution of stock returns. J Am Stat Assoc 67:807–812
Onatski A (2009) Testing hypotheses about the number of factors in large factor models. Econometrica 77:1447–1479
Osborne MFM (1959) Brownian motion in the stock market. Oper Res 7:145–173
Pearson K (1901) On lines and planes of closest fit to system of points in space. Lond Edinb Dubl Philos Mag J Sci 2:559–572
Peiró A (1994) The distribution of stock returns: international evidence. Appl Financ Econ 4:431–439
Perold AF (2004) The capital asset pricing model. J Econ Perspect 18:3–24
Phelan MJ (1997) Probability and statistics applied to the practice of financial risk management: the case of JP Morgan's riskmetrics^{TM}. J Financ Serv Res 12:175–200
Reinganum MR (1981) The arbitrage pricing theory: some empirical results. J Financ 36:313–321
Roll R, Ross SA (1980) An empirical investigation of the arbitrage pricing theory. J Financ 35:1073–1103
Rossi M (2016) The capital asset pricing model: a critical literature review. Glob Bus Econ Rev 18:604
Shahvaroughi Farahani M, Razavi Hajiagha SH (2021) Forecasting stock price using integrated artificial neural network and metaheuristic algorithms compared to time series models. Soft Comput 25(13):8483–8513
Sharpe WF (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. J Financ 19:425–442
Shen J, Shafiq MO (2020) Short-term stock market price trend prediction using a comprehensive deep learning system. J Big Data 7:66
Stock JH, Watson MW (1999) Forecasting inflation. J Monet Econ 44:293–335
Stock JH, Watson MW (2002) Forecasting using principal components from a large number of predictors. J Am Stat Assoc 97:1167–1179
Stock J, Watson M (2016) Chapter 8—dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In: Taylor JB, Uhlig H (eds) Volume 2 of handbook of macroeconomics. Elsevier, pp 415–525
Wang C (2024) Stock return prediction with multiple measures using neural network models. Financ Innov 10:72
Xi J, Wen C, Tang Y, Zhao F (2024) A factor score clustering approach to analyze the biopharmaceutical sector in the Chinese market during Covid-19. Financ Innov 10:135
Yu Z et al (2023) Engineering a data processing pipeline for an ultra-lightweight lensless fluorescence imaging device with neuronal-cluster resolution. Artif Life Robot 28:483–495
Zhong X, Enke D (2019) Predicting the daily return direction of the stock market using hybrid machine learning algorithms. Financ Innov 5:24

Publisher's Note