CVRA: Asset Pricing via the Conditional Variational Recurrent Autoencoder in Asian Market

Abstract

We proposed a novel probabilistic dynamic factor model for investment strategies in Asian market, which improves handling of complex and noisy market conditions. Our model integrates deep learning's variational recurrent autoencoder with advanced temporal dependency modeling. A key innovation of our model is its prior-posterior learning method, which refines the model using future data to optimize posterior factors. Designed for volatile stock markets, our model effectively estimates variances from latent space distributions and predicts returns. Statistical and empirical experiments on China and Japan stock market data demonstrate the model's superior performance compared to established traditional and machine learning methods.

Keywords: Asian stock market, Dynamic factor model, Machine learning, Neural networks, Variational autoencoder, Factor investing

1. Introduction

Forecasting financial market dynamics has become essential, particularly for the challenging task of predicting stock market returns in the unique and rapidly evolving markets of Asia, specifically China and Japan. These markets differ structurally from the developed market like U.S. market, with distinct characteristics like higher retail investor participation (Jones et al., 2021; Naseem et al., 2021; Xiong et al., 2020), government influence (Tan et al., 2023; Su et al., 2023), and more dynamic market cycles (Zhang, 2018; Urasawa, 2018). Traditional asset pricing models, such as static factor models, often struggle in Asia's complex environment due to high-dimensional datasets and noisy, volatile financial data.(Callot et al., 2021; Caner and Daniele, 2022)

Unlike the U.S. market, where institutional investors dominate and market dynamics are more predictable(Chan et al., 2022; Weber, 2023), China and Japan's markets are more volatile, have lower signal-to-noise ratios, and exhibit strong temporal dependencies. These factors necessitate advanced methods for capturing complex, nonlinear relationships within high-dimensional data. Recent advancements in machine learning, especially deep learning, have shown great promise in addressing these challenges and significantly outperforming traditional models by adapting to Asia's unique market structures.(Leippold et al., 2022; Ye et al., 2024)

Traditional static factor models from Ross (1976), $r_{i,t} = \beta f_t + u_{i,t}$, rely on a linear framework. Yet, they face limitations with dynamic factors, where betas are not static and non-linearity as market conditions change rapidly. To solve the non-linearity, more deep learning skills have recently been employed. Caner and Daniele (2022) develop a deep neural network model to get a consistent estimator of the precision matrix of asset returns in large portfolios based on deep learning residuals formed from non-linear factor models, and the model works well even in low signal-to-noise environment. Gu et al. (2021) utilize a deep learning technique known as autoencoder, supplemented with additional explanatory variables. They introduce a novel asset pricing model named the conditional autoencoder (CAE), which enables f_t and $\beta\left(z_{i,t-1}\right)$ to vary in a nonlinear and entirely data-driven way based on returns or asset characteristics. The CAE model outperforms many alternative models, such as IPCA. Their latent factor models attempt to introduce dynamic betas through a high-dimensional vector of asset characteristics, but their model stills struggle to challenges like heavy-tailedness conditional heteroscedasticity, and heterogeneity.(Yang et al., 2024) Besides, their deep learning research does not account for temporal dependencies in time series data, where current values often hinge on past data.(ZHOU and Wang)

To solve those problems, especially in Asia's markets, we explore deep learning method, which offers a superior, data-driven approach to latent factor models, automating latent factor extraction to match market dynamics. Advanced models, such as Recurrent Neural Networks (RNN) and their variants (e.g., GRU, LSTM), can capture temporal dependencies that are crucial in noisy, complex markets like China and Japan. However, these models often falter in environments with high noise, low signal, and complex intersections across time series. Our proposed model, CVRA (Conditional Variational Recurrent Autoencoder), addresses these challenges by combining the temporal advantages of RNNs with the factor-extraction strengths of Variational Autoencoders (VAEs) into latent factor model. This structure enables better modeling of latent factors while maintaining robust predictive accuracy.

CVRA uses a variational recurrent autoencoder framework that treats factors as latent variables, capturing data noise effectively within CVRA's latent space. The encoder-decoder structure, trained on historical stock returns, optimally reconstructs returns and extracts factors for predictive accuracy. The model avoids future data leakage by relying on a predictor trained on past data. Additionally, CVRA offers probabilistic risk estimation, accounting for inherent market randomness, making it uniquely suited for the Asian markets' structural complexity. Novel integration of CVRA in an encoder-decoder framework specifically for Asia's noisy markets, enabling a probabilistic approach for accurate return and risk estimation. Our statistical and empirical results on real stock data from China and Japan, which demonstrate that CVRA surpasses traditional factor models and machine and deep learning approaches in predictive performance for cross-sectional returns.

In summary, our work underscores the potential of machine and deep learning for Asian markets, where structural differences necessitate advanced, nonlinear approaches that outperform traditional methods used in the more stable U.S. market.

The contributions of our paper are as follows:

- We introduce CVRA as an innovative dynamic latent factor model tailored to extract key factors from noisy market data efficiently. This model integrates Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) architectures to capture the temporal dependencies in time series data. Additionally, it incorporates a Variational Autoencoder (VAE) to address challenges in markets with low signal-to-noise ratios.

- To the best of our knowledge, we are the first to apply CVRA to blend RNN and VAE in an encoder-decoder framework specifically for Asia's noisy markets, enabling a probabilistic approach for accurate return and risk estimation.

- Our comprehensive experiments on actual stock market data reveal that CVAR outperforms traditional linear and non-linear factor models as well as various machine learning and deep learning-based models in predicting cross-sectional returns, which indicates at the broad application prospects of machine learning and deep learning methods in Asian capital markets.

2. Related Work

2.1. Dynamic Factor model

Factor models, which aim to explain the variation in stock returns, can be broadly classified into two categories: static models and dynamic models. In static factor models, the factor exposure of a stock remains time-invariant. The foundational static factor model is the Capital Asset Pricing Model, proposed by Sharpe (1964). However, traditional static factor model has been challenged by time-varying beta and noisy market data. (Jensen, 1972)

In dynamic factor models, the factor exposure of stocks varies over time. This variation is often calculated based on firm or asset characteristics such as market capitalization, book-to-market ratio, and asset liquidity. Kelly et al. (2019) introduces instrumented principal components analysis into dynamic factor models, allowing factor exposure to depend partially on observable asset characteristics with a linear relationship.

Going beyond linear relationships, Gu et al. (2021) proposes a latent dynamic factor asset pricing model, $r_t = \beta\left(z_{i,t-1}\right)' f_t + u_t$ incorporating a conditional autoencoder network. This model introduces non-linearity into return dynamics, addressing limitations found in linear factor models. The authors demonstrated that their non-linear factor model outperformed leading linear methods in terms of effectiveness.

2.2. Recurrent Neural Network (RNN)

However, challenges faced by factor models are focused on two sides: the presence of temporal dependency between stock return data and noisy market data with low signal-to-noise ratios. Traditional feedforward neural network (FFN) structure treats each input as independent and identical distribution, which neglects the intersection between time series data. Not considering temporal dependencies can lead to poor predictions and misunderstanding of the data. Therefore, there is ongoing research dedicated to addressing the issue of learning temporal dependency between time series data.

In recent years, more and more scholars use Recurrent Neural Network (RNN) to explore temporal dependency in market data. Fischer and Krauss (2018) explored the use of LSTM networks in predicting stock market movements and found that these models significantly outperform traditional time series models. Dey et al. (2019) integrates sentiment analysis from financial news with RNNs to predict stock price movements. Nelson et al. (2017) focuses on the application of RNNs in high-frequency trading. They found that

RNNs, particularly LSTMs, are highly effective in capturing the dynamics of the market at higher frequencies

2.3. Variational Autoencoder (VAE)

For non-sequential data, VAEs have recently been shown to be an effective modelling paradigm to recover complex multimodal distributions over the data space (Kingma, 2013). A VAE introduces a set of latent random variables z, designed to capture the variations in the observed variables x.

The joint distribution is defined as:

$$p\left(x,z\right)=p\left(x|z\right)p(z) \tag{1}$$

The prior over the latent random variables, $p(z)$, is generally chosen to be a simple Gaussian distribution and the conditional $p\left(x,z\right)$ is an arbitrary observation model whose parameters are computed by a parametric function of z. Importantly, the VAE typically parameterizes $p\left(x,z\right)$ with a highly flexible function approximator such as a neural network. While latent random variable models of the form are not uncommon, endowing the conditional $p\left(x,z\right)$ as a potentially highly non-linear mapping from z to x is a rather unique feature of the VAE. However, introducing a highly non-linear mapping from z to x results in intractable inference of the posterior $p\left(x,z\right)$. Instead, the VAE uses a variational approximation $q\left(z,x\right)$ to approximate $p\left(x,z\right)$.

The approximate posterior is a Gaussian $N\left(\mu,diag\left(\sigma^2\right)\right)$ with the mean μ and variance σ^2 are the output of a highly non-linear function of x, once again typically a neural network. The generative model $p\left(x,z\right)$ and inference model $q\left(z,x\right)$ are then trained jointly by maximizing the evidence lower bound (ELBO) with respect to their parameters, where the integral with respect to $q\left(z,x\right)$ is approximated stochastically. The gradient of this estimate can have a low variance estimate by reparameterizing z.

$$E_{q(z|x)}\left(logp\left(x|z\right)\right)=E_{p(\varepsilon)}\left(logp\left(x|z=\mu+\sigma\varepsilon\right)\right) \tag{2}$$

Figure 1 shows the structure of VAE.

2.4. Variational Recurrent Autoencoder

Though VAE has already shown its advantage in efficient learning of complex distribution, it is not inherently designed to handle sequential data, such as time series, music, or text. Chung et al. (2015) and Fabius and

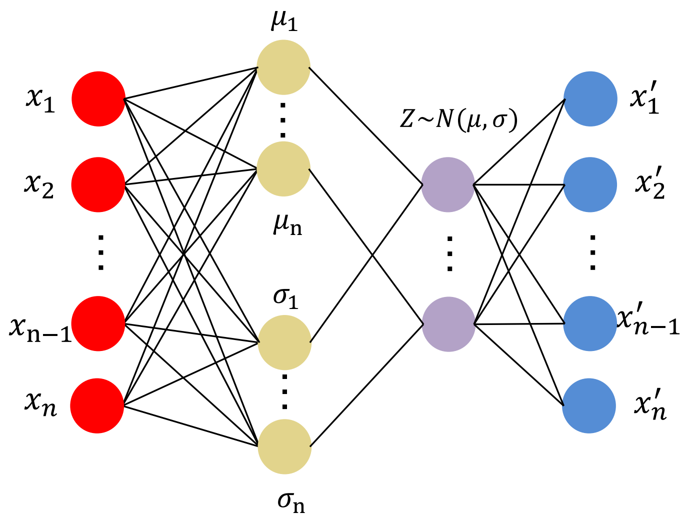

Figure 1: Variational Autoencoder Architecture

Van Amersfoort (2014) propose that a new recurrent structure which contains a VAE at each time step and VAE is conditioned at the state variable h_{t-1} of an RNN. The prior on the latent random variable is no longer a standard Gaussian distribution:

$$Z \sim N\left(\mu_{0,t},\ diag\left(\sigma_{0,t}^{2}\right)\right) \tag{3}$$

where $[\mu_{0,t},\ \sigma_{0,t}] =\ \varphi_{\tau}^{prior}\left(h_{t-1}\right)$ Moreover, the generating distribution will not only be conditioned on z but also on h_{t-1}.

$$x_t|z_t \sim N\left(\mu_{x,t},\ diag\left(\sigma_{x,t}^{2}\right)\right. \tag{4}$$

where $[\mu_{x,t},\ \sigma_{x,t}] =\ \varphi_{\tau}^{dec}\left(\varphi_{\tau}^{z}(z_t)\,,\ h_{t-1}\right)$. Then, RNN are used to update hidden states at the last step.

$$h_t =\ f_{\theta}\left(\varphi_{\tau}^{x}\left(x_t\right),\ \varphi_{\tau}^{z}\left(z_t\right), h_{t-1}\right) \tag{5}$$

VRAE has shown its success in many industries. In recent years, VRAE's scope has expanded to include anomaly detection, with notable advancements in identifying irregular events in video data (Yan et al., 2020), in monitoring anomalies in solar energy systems (Ferreira et al., 2018), and in estimating stock volatility (Luo et al., 2018).

3. Methodology

According to Ross (1976),

$$r_t = \ \beta_t{}' f_t + u_t \tag{6}$$

where where $t = 1, \ldots, T$, f_t is risk premia, β_t is factor exposures, and u_t is the idiosyncratic error. We use CVRA, which contains VAE in each timestep. The formulation of our task is to learn a dynamic factor model with parameter θ to predict future cross-sectional returns from historical data.

Our network model is composed of two modules:beta network and factor network.

In the first module about factor network, we combine recurrent structure with variational autoencoder, assume that the vector of cross-sectional returns r_t is driven by a vector of latent variables z_t and hidden state variables h_{t-1}, then apply the regularized conditional variational recurrent autoencoder (CVRA) to learn the latent factor f_t from the approximated posterior distribution of z_t. Specifically, r_t as the input has the following marginal likelihood:

$$p_\theta\left(r_t\right) = \int p_\theta\left(r_t|z_t\right) p_\theta\left(z_t\right) dz_t \tag{7}$$

where $p_\theta\left(z_t\right)$ is the prior probability. In vanilla VAE, $p_\theta\left(z_t\right)$ is assumed as a standard Gaussian distribution. Then we can get posterior probability of z_t:

$$p_\theta\left(z_t|r_t\right) = \ \frac{p_\theta\left(r_t|z_t\right) p_\theta\left(z_t\right)}{p_\theta\left(r_t\right)} \tag{8}$$

where $p_\theta\left(r_t|z_t\right)$ is the likelihood function and θ is a vector of unknown parameters. Up till now, the method to compute $p_\theta\left(z_t|r_t\right)$ is intractable, especially when $p_\theta\left(r_t|z_t\right)$ has a complex structure. To solve the problem, we use a new function $q_\varnothing\left(z_t|r_{\ t}\right)$ with unknown parameters ϕ to approximate the true posterior $p_\theta\left(z_t|r_t\right)$. During this process, given r_t and an initial hidden state variable h_0, it produces a distribution over the possible values of z_t from which r_t could have been reconstructed. Thus, we can construct:

$$q_\varnothing\left(z_t|r_t\right) = N\left(\mu_{zt},\ diag\left(\sigma_{zt}^2\right)\right) \tag{9}$$

where $\left[\mu_{zt},\ \sigma_{zt}\right] = \ \varphi_\tau^{encoder}\left(\varphi_\tau^r(r_t)\ ,\ h_{t-1}\right)$, μ_{zt}, σ_{zt} are outputs of neural network with returns r_t and state variables h_{t-1}. $\varphi_\tau^{encoder}$ can be any highly flexible function such as neural networks and φ_{τ}^{r} means function can extract features from r_t.

Therefore, our encoder is as flows:

$$h_0 = 0,\ r_0 =\ r_t \tag{10}$$

$$h_t = tanh\left(w_{encoder}^{T} h_{t-1} + w_{in}^{T} r_t + b_{encoder}\right) \tag{11}$$

$$\mu_{zt} = w_{\mu}^{T} h_t + b_{\mu} \tag{12}$$

$$log\left(\sigma_z^2\right) = w_{\sigma}^{T} h_t + b_{\sigma} \tag{13}$$

$$Z \sim N\left(\mu_{zt},\ diag\left(\sigma_{zt}^2\right)\right) \tag{14}$$

Equation (10) initializes the network with the vector of individual asset returns r_t and the state variables h_t. Equation (11) shows how RNN updates its hidden state with recurrence equation. Equation (12) and (13) calculate the mean and the standard deviation of posterior factors from the mapping layer,respectively. Equation (14) shows the distribution of z_t.

Before we come to the decoder process, we move our eyes to the reparameterization trick. If we sample directly samples z_t from $q_{\varnothing}\left(z_t|r_t\right)$, the sampling behavior is undifferentiable. Thus, we use the reparameterization trick proposed by Chung et al. (2015) to solve it. However, our model samples multiple z_t at different timesteps, and for $q_{\varnothing}\left(z_t|r_t\right)$, the mean and covariance do not directly depend only on z_{t-1}, but also h_t. After using the reparameterization trick,z_t can be sampled as follows:

$$z_t\ =\ \mu_z\left(h_t, r_t\right) +\ diag\left(\sigma\left(h_t, r_t\right)_z^2\right)^{\frac{1}{2}} \varepsilon_t \tag{15}$$

where $\varepsilon_t \sim N(0,\ I)$, we can sample z_t by just sampling ε_t, where ε_t is viewed as the stochastic input of the model with a standard multivariate normal distribution depending not on any unknown parameters. Then, we pass z_T from the last timestep to decoder. For decoder process, the model needs simultaneously specify a probabilistic decoder $p_{\theta}\left(r_t|z_t\right)$, which produces a distribution over the possible values of corresponding to a given code z_t.

$$p_{\theta}\left(r_t|z_t\right) =\ N\left(\mu_{rt},\ diag\left(\sigma_{rt}^2\right)\right) \tag{16}$$

where $[\mu_{rt},\ \sigma_{rt}] =\ \varphi_{\tau}^{decoder}\left(\varphi_{\tau}^{z}(z_t)\ ,\ h_{t-1}\right)$, $\mu_{rt},\ \sigma_{rt}$ are outputs of neural network with returns r_t and state variables h_{t-1}. $\varphi_{\tau}^{decoder}$ can be any highly flexible function such as neural networks and φ_{τ}^{r} means function can extract features from z_t. Thus, we construct the decoder part:

$$z_0 = \ z_T \tag{17}$$

$$h_0 = tanh\left(w_z^T z_T \ + b_z\right) \tag{18}$$

$$h_t = tanh\ \left(w_{decoder}^T h_{t-1} + w_r^T r_t + b_{decoder}\right) \tag{19}$$

$$\hat{f}_t = sigmod\left(w_{out}^T h_t + b_{out}\right) \tag{20}$$

Equation(17) and (18) initializes the network with state variable h_t and latent variable z_t. Then similar to the encoding process, RNN updates its hidden state with recurrence (19). Equation (20) shows how we obtain the output layer f_t.

The second module of our model is to estimate β_t from firm characters x_t with a multi-head structure. Compared with traditional Deep neural network, multi-head structure has advantage in extracting higher-order features from the input and identifying some protential effect like the quantile effect. (Yang et al., 2024). The Beta network are shown as follow:

$$s_{i,t-1} = \widetilde{g}\left(W_{z1} c_{i,t-1} + b_{z1}\right) \tag{21}$$

$$\beta_{i,t-1} = ReLU(W_{beta} s_{i,t-1} + b_{z2}) \tag{22}$$

The equation (21) shows that $s_{i,t-1}$ is the hidden output to transform firm characteristics $c_{i,t-1}$. The equation (22) shows that the factor betas $\beta_{i,t-1}$ emerges from the terminal output layer $s_{i,t-1}$ on firm characteristics c_t (in orange). Figure2 shows our model structure. The factor network (right panel) describes how the latent factors f_t are obtained from an vector of asset returns r_t (in lower red) via CVRA. The upper red nodes in output layer are computed by multiplying each row from the beta network with the vector of latent factors (in blue) from the factor network.

Besides, following Gu et al. (2021), we replace the individual returns with a P-dimensional vector of dynamically re-weighted, characteristic-based portfolios, $x_t = \left(Z'_{t-1} Z_{t-1}\right)^{-1} Z'_{t-1} r_t$, to reduce the model's complexity. This method offers three benefits: a significant reduction in the sequence model's parameters due to smaller dimensions, an effective bypass of return panel imbalances from missing stocks, and enhanced latent factor determination, leveraging the proven importance of characteristic-managed portfolios in refining conditional linear factor models.

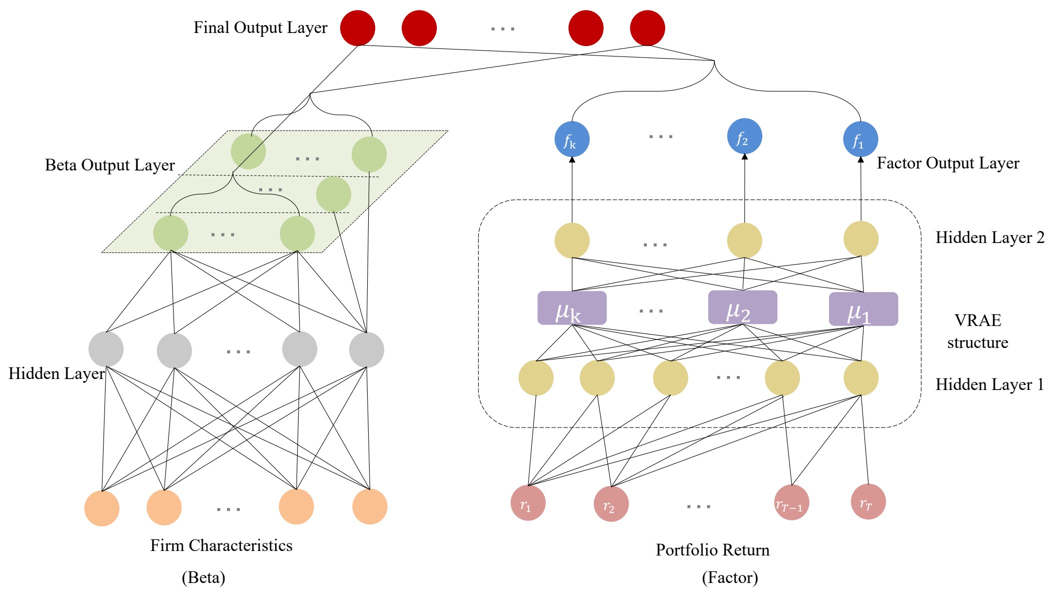

Figure 2: Our Model Architecture

4. Estimation

Our estimation consists of two parts. The first part is to train an optimal factor model, which is to reduce the reconstruction error, and the second part is to enforce the prior factors to approximate to the posterior factors with Kullback–Leibler (KL) divergence. Thus, the objective function of our model is:

$$\begin{aligned} L(\theta, \phi; r_t) = & \frac{1}{L}\sum_{l=1}^{L} \|r_t - \hat{r_t}(\theta)\|_2^2 \\ & + \lambda KL\left(q_{\varnothing}\left(z_t|r_t\right) \parallel p_\theta\left(z_t|r_t\right)\right) \end{aligned} \tag{23}$$

Where θ and ϕ include all unknown weights and bias parameters. L is the number of Monte Carlo method ,and λ is the weight of KL divergence loss. According to Yang et al. (2020), we can replace KL divergence loss with ELBO. Formally, given input r_t, the ELBO of our model is defined as:

$$\begin{aligned} ELBO = & E_{z\sim q_{\varnothing}(z_t|r_t)}\left(log p_\theta\left(r_{i,t}|z_T\right)\right) \\ - & \frac{1}{T}\sum_{t=1}^{T} KL\left(q_{\varnothing}\left(z_t|r_t\right) \parallel p_\theta\left(z_t\right)\right) \end{aligned} \tag{24}$$

The model passes a single z_t at the final timestep of encoder to the decoder. Moreover, compared with vanilla function to calculate ELBO, there is a crucial difference that while existing models only impose KL regularization on the last timestep, we impose timestep KL regularization and average the KL loss over all timesteps, which allows more robust model learning and can effectively mitigate posterior collapse. Through the above objective function, we can use Adam algorithm to obtain unbiased θ and ϕ.

5. Empirical Result

5.1. Data

We study the Asia–Pacific equity markets, focusing on China and Japan, where structural frictions, heterogeneous information environments, and varying degrees of market liberalization provide a natural testing ground for dynamic factor models. Our dataset covers the period from January 2000 to December 2024. Monthly stock returns for all listed firms in the Shanghai Stock Exchange (SSE), Shenzhen Stock Exchange (SZSE), and Tokyo Stock Exchange (TSE) are obtained from the Wind Financial Terminal and Bloomberg. These exchanges collectively span firms across diverse ownership structures, liquidity regimes, and industry compositions, allowing us to evaluate model performance under heterogeneous market conditions.

We construct a panel of 46 stock-level characteristics widely used in the empirical asset pricing literature, drawn from Gu et al. (2020, 2021); Freyberger et al. (2020); Leippold et al. (2022). The characteristics include measures of liquidity, volatility, profitability, investment, growth, and market-based signals. Following Gu et al. (2021), we align each month-t characteristic with the return realized in month $t+1$ to avoid forward-looking bias and to ensure a clean predictive structure. To address the heavy-tailed and skewed nature of characteristics documented in Gu et al. (2021), we apply the standard rank-normalization transformation:

$$c_{ij,t} = \frac{2}{N_t + 1} \operatorname{rank}(c^*_{ij,t}) - 1, \tag{25}$$

where $c^*_{ij,t}$ is the raw characteristic of firm i, characteristic j at time t, and N_t is the number of firms available at month t. This transformation maps each characteristic into the interval $(-1, 1)$ and mitigates the impact of outliers, improving numerical stability during training.

Rolling-Window Estimation and Real-Time Forecasting. Rather than using a static train–validation–test split, we adopt a rolling-window estimation design consistent with empirical asset pricing practice and real-time forecasting constraints. At each month τ, the model is trained using an expanding window of all historical data from 2000 up to month τ:

$$\mathcal{D}_\tau^{\text{train}} = \{(r_{i,t+1}, c_{i,t}) : t = 2000{:}\tau\}.$$

Model parameters are updated recursively as new data arrive, and a one-step-ahead forecast is generated for month $\tau + 1$:

$$\widehat{r}_{i,\tau+1} = \mathbb{E}(r_{i,\tau+1} \mid c_{i,\tau}, \mathcal{D}_\tau^{\text{train}}).$$

To prevent overfitting and maintain comparability across models, hyperparameters are tuned using a 10-year sliding validation block immediately preceding the forecasting origin. This procedure mimics the forecasting environment faced by practitioners and ensures that no future information is used at any point in the training pipeline.

Finally, to maintain comparability across model classes, we set the number of latent factors to $K = 5$ unless otherwise noted. This specification is consistent with the range commonly estimated for Asia–Pacific markets and aligns with prior empirical findings that 4–6 latent factors capture most priced variation in these markets.

Our empirical analysis focuses on two of the largest and most structurally distinctive equity markets in Asia—China and Japan—where retail participation, heterogeneous trading motives, and intermittent policy interventions generate a return environment markedly different from that of developed Western markets. To ensure broad market coverage, we compile monthly returns for all A-share firms listed on the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE), as well as for firms listed on the Tokyo Stock Exchange (TSE). Chinese return and accounting data are obtained from the Wind Financial Terminal, the principal commercial database for mainland China, while Japanese returns come from the Bloomberg Terminal. The sample spans January 2000 to December 2024, a period that encompasses episodes of financial liberalization, regulatory tightening, and pronounced speculative cycles in both markets, making it well suited for evaluating a model that emphasizes nonlinear dynamics and time-varying uncertainty.

Alongside returns, we construct a panel of 46 firm-level and macroeconomic characteristics drawn from Gu et al. (2021, 2020); Freyberger et al.

(2020); Leippold et al. (2022), covering well-established predictors such as size, value, profitability, investment, liquidity, and past return trends. These characteristics are matched to returns using a one-month lag to avoid forward-looking bias, following the convention in Gu et al. (2021). The resulting database includes firms of varying listing ages, leading to an unbalanced return panel—an empirically relevant setting in which sequential latent-variable models such as CVRA can leverage their ability to operate on characteristic-managed portfolios rather than on raw returns.

Because many asset characteristics exhibit strong skewness, heavy tails, and cross-sectional outliers—features that are especially pronounced in the Chinese A-share market—each raw characteristic $c^*_{ij,t}$ is transformed using the rank-normalization procedure of Gu et al. (2021):

$$c_{ij,t} = \frac{2}{N+1}\,\mathrm{rank}(c^*_{ij,t}) - 1, \tag{26}$$

where $\mathrm{rank}(c^*_{ij,t})$ denotes the cross-sectional rank of the characteristic across all N firms at time t. The transformation maps all characteristics to the interval $(-1, 1)$ and mitigates the undue influence of extreme observations, thereby stabilizing the learning of characteristic-based exposures in the beta network.

For each firm i and month t, we denote the return by $r_{i,t}$ and the vector of normalized characteristics by $c_{i,t} = (c_{i1,t}, \ldots, c_{i,46,t})'$. To evaluate out-of-sample predictive performance, the full sample is partitioned into a 12-year training set, a 4-year validation set, and a 4-year test set. Unless otherwise noted, the number of latent factors is set to five ($K = 5$), consistent with the benchmark configurations in Gu et al. (2021) and with the empirical evidence suggesting that a small number of factors can explain a substantial fraction of variation in Asian equity returns.

We apply our model to analyze the Asia equity market, especially China and Japan market. We collect daily and monthly individual A-share stock returns from the Wind Database, the largest financial data provider in China. We also collect individual stock returns from Bloomberg Terminal. The returns dataset is from January 2000 to December 2019 for firms listed in the three major exchanges in China and Japan: Shanghai Stock Exchange (SSE), Shenzhen Stock Exchange (SZSE), and Tokyo Stock Exchange (TSE). In addition, we collect 46 different firm and macroeconomic characteristics for each stock. All listed characteristics are chosen from Gu et al. (2021, 2020); Freyberger et al. (2020); Leippold et al. (2022). Following Gu et al. (2021), we match stock characteristics with corresponding returns to eliminate forward-looking bias in our dataset. Finally, we split data into 12-year training set, 4-year validation set and 4-year test set. We also take 5 as the number of factor to construct our model.

For each stock i at month t, $r_{i,t}$ represents the return, and $c^*_{ij,t}$, $j = 1, \ldots, 46$, denotes 46 distinct characteristics. To counteract the skewness and leptokurtosis in asset characteristics observed by Gu et al. (2021), and mitigate outlier effects, we employ their recommended rank normalization method:

$$c_{ij,t} = \frac{2}{N+1} rank(c^*_{ij,t}) - 1 \tag{27}$$

where $rank\left(c^*_{ij,t}\right)$ is the rank of $c^*_{ij,t}$ within all characteristics $c^*_{ij,t}$. After rank-normalization, $c_{ij,t}$ values fall within $(-1,\ 1)$. We analyze using the vector $c_{i,t}$, composed of normalized asset characteristics, where each j-th entry is $c_{ij,t}$.

5.2. Model Comparison

In our empirical analysis, we compare other mainstream machine learning models: NN, CA, FactorVAE, ALSTM, and our model, CVRA. The NN model, applied in Gu et al. (2020); Leippold et al. (2022) is the neural network model. According to **?**, we take their best performace model NN4 as one of benchmark model here. The CA model, introduced by Gu et al. (2021), uses the Autoencoder and neural network for latent factors and nonlinear factor loading, respectively. We use their CA_2 model, which has two hidden layers in their beta network. The CA_2 model has the best overall performance in both statistical and economic estimation in Gu et al. (2021). FactorVAE is a probabilistic dynamic factor model based on variational Autoencoder from Duan et al. (2022). The model use vanilla VAE but doesn't consider time series's temporal dependency. ALSTM is a the time series model with LSTM with an attention layer for stock returns prediction. (Qin et al., 2017). CVRA is our model which uses LSTM structure and Variational recurrent autoencoder to construct beta network and factor network, respectively.

5.3. Statistical Performance

We follow Kelly et al. (2019) and Gu et al. (2021) to use total and predictive R^2 to evaluate the model's capacity to explain stock risk and predict future returns, with total R^2 assessing contemporaneous factor identification and predictive R^2 gauging accuracy in risk compensation variation, respectively.

$$R^2_{total} = 1 - \frac{\sum \left(r_{i,t} - \widehat{\beta'}_{i,t-1}\widehat{f_t}\right)^2}{\sum r^2_{i,t}} \tag{28}$$

and

$$R^2_{pred} = 1 - \frac{\sum \left(r_{i,t} - \widehat{\beta'}_{i,t-1}\widehat{\mu_{t-1}}\right)^2}{\sum r^2_{i,t}} \tag{29}$$

where μ_{t-1} is the prevailing sample average of f_t up to month $t-1$.

Table 1 reports the out-of-sample total R^2's for managed portfolios x_t in China A-Share market. We noticed that Among the models, CVRA achieves the highest R^2 across most configurations, with a peak value of 86.26% at $K = 5$, outperforming the other methods in explaining the portfolio returns. This is closely followed by CA and FactorVAE, which also exhibit strong performance, with R^2 values of 85.61% and 85.37% at $K = 5$ and $K = 6$, respectively. NN and ALSTM show slightly lower R^2 values, with their performance plateauing at $K = 6$. Besides, As K increases from 1 to 6, the total R^2 improves consistently across all models, indicating that incorporating additional factors enhances the explanatory power of the models.

Table 1: Total R^2 in China Market

Model	Test assets	*Number of Factors(K)*					
		1	2	3	4	5	6
NN	x_t	72.10	75.12	78.46	80.43	82.73	85.14
CA	x_t	73.32	76.41	80.51	82.56	85.61	85.29
FactorVAE	x_t	72.89	75.98	80.23	81.87	85.32	85.37
ALSTM	x_t	73.35	75.34	79.72	81.26	84.73	84.26
CVRA	x_t	74.12	76.22	81.05	83.43	86.26	86.01

Table 2 reports the out-of-sample total R^2's for managed portfolios x_t in Japan market. CVRA again achieves the highest R^2, with a peak value of 88.63% at $K = 5$, demonstrating superior predictive capabilities. FactorVAE follows closely, with an R^2 of 88.03% at $K = 6$, slightly outperforming CA at the same factor level. NN and ALSTM display lower R^2 values relative to CVRA but still show a steady improvement as K increases. Compared to the China market, models in Japan consistently exhibit higher R^2 values

Table 2: Total R^2 in Japan Market

Model	Test assets	*Number of Factors(K)* 1	2	3	4	5	6
NN	x_t	74.41	77.31	80.53	82.63	85.93	87.64
CA	x_t	76.05	79.82	82.79	85.65	88.14	87.91
FactorVAE	x_t	76.12	80.35	82.98	85.73	87.92	88.03
ALSTM	x_t	75.84	78.96	81.62	84.53	86.31	86.54
CVRA	x_t	78.12	81.22	83.05	86.85	88.63	88.36

across all K, suggesting that the Japan market might have less noise or more structured factor relationships that these models can exploit effectively.

Table 3 reports the out-of-sample predictive R^2's for managed portfolios x_t. Predictive R^2 measures the ability of a model to forecast portfolio returns accurately, providing insight into the models' out-of-sample performance.

As the number of factors increases from $K = 1$ to $K = 6$, all models exhibit a general improvement in predictive R^2, though the gains diminish at higher values of K. CVRA consistently achieves the highest predictive R^2 values across all factor configurations, with a peak of 2.52% at $K = 6$. This indicates its superior ability to model the complex dynamics of portfolio returns. FactorVAE also performs well, with a maximum predictive R^2 of 2.30% at $K = 6$, followed closely by CA at 2.24%. ALSTM and NN trail behind, suggesting that these models might be less effective in capturing the intricate relationships between factors and portfolio returns.

Notably, CVRA demonstrates a significant jump in performance at $K = 6.$, where its predictive R^2 rises to 2.49%, outperforming other models by a considerable margin. This suggests that CVRA benefits more from an optimal number of factors, leveraging its ability to learn latent factor representations effectively. In contrast, the performance of other models, such as CA and FactorVAE, plateaus or improves only marginally at higher values of K, indicating their relative limitations in handling additional complexity.

Table 4 reports the out-of-sample predictive R^2's for managed portfolios x_t in Japan market. As K increases from 1 to 6, all models demonstrate improved predictive R^2, though the rate of improvement diminishes at higher K. CVRA still outperforms all other models, achieving the highest R^2 of 2.62% at $K = 4$, with performance remaining strong at $K = 5$ and $K = 6$ (2.54% and 2.58%, respectively). FactorVAE and CA also perform well, with

Table 3: Predictive R^2 in China Market

Model	Test assets	*Number of Factors(K)*					
		1	2	3	4	5	6
NN	x_t	1.71	1.92	2.07	2.04	2.14	2.13
CA	x_t	1.85	2.01	2.17	2.26	2.21	2.24
FactorVAE	x_t	1.87	2.06	2.23	2.32	2.26	2.30
ALSTM	x_t	1.77	1.96	2.11	2.24	2.17	2.18
CVRA	x_t	1.91	2.05	2.24	2.49	2.36	2.52

CA attaining a peak R^2 of 2.53% at $K = 6$, while FactorVAE reaches 2.51% at $K = 4$. NN and ALSTM trail slightly, highlighting their relative inefficiency in capturing the complex dynamics of factor-based portfolio returns compared to CVRA.

When compared to the China market (Table 3), the Japan market exhibits slightly higher predictive R^2 values across most models, indicating that models achieve better forecasting performance in Japan. This difference suggests that the Japan market might be more structured, with factor-based relationships that are easier for these models to exploit. CVRA's predictive advantage is consistent across both markets, but its performance improvement at higher K values is more pronounced in Japan, suggesting better alignment between its latent variable framework and the characteristics of the Japan market.

Table 4: Predictive R^2 in Japan Market

Model	Test assets	*Number of Factors(K)*					
		1	2	3	4	5	6
NN	x_t	1.94	2.12	2.23	2.24	2.35	2.31
CA	x_t	2.06	2.23	2.37	2.45	2.42	2.53
FactorVAE	x_t	2.03	2.26	2.43	2.51	2.41	2.49
ALSTM	x_t	1.98	2.16	2.33	2.41	2.36	2.38
CVRA	x_t	2.11	2.28	2.47	2.62	2.54	2.58

The results from Tables 1 to 4 provide a comprehensive comparison of model performance across the China and Japan markets for out-of-sample total and predictive 3. Key findings highlight consistent improvements in R^2 as the number of factors increases, reflecting the value of incorporating more factors to explain and predict portfolio returns.

In the total R^2 analysis, the CVRA model outperforms other methods in both markets, achieving peak values of 86.26% in China and 88.63% in Japan. Models such as FactorVAE and CA also perform well but trail CVRA, while NN and ALSTM show slightly lower performance. Notably, the Japan market exhibits higher total R^2 values across models, suggesting a more structured factor relationship. For predictive R^2, CVRA again demonstrates superior forecasting ability, achieving the highest values in both markets, with 2.52% in China and 2.62% in Japan. Japan's higher predictive R^2 values across models suggest that factor-based predictions may be more effective in this market.

Overall, CVRA consistently delivers the best performance, leveraging its latent variable framework effectively. The higher R^2 values in Japan suggest market-specific differences, emphasizing the importance of tailoring models to regional characteristics.

5.4. Economic Performance

We evaluate five models economically by examining Sharpe ratios of portfolios with and without transaction costs, informed by their predictions on the conditional mean and variance of returns in the testing sample. We sort stocks into deciles, forming two portfolio types: a long-short portfolio from buying top 10% (decile 10) and selling bottom 10% (decile 10), and a long-only portfolio comprising only the top 10% (decile 10), using equal weighting.

Table 5 presents the equal weighting annualized Sharpe ratios for both long-short and long-only portfolios across five models at various K values, excluding transaction costs. For long-short portfolios, a clear trend emerges: as the number of factors increases, the Sharpe ratios improve significantly. CVRA achieves the highest Sharpe ratios across all K, peaking at 2.85 for $K = 6$, indicating superior risk-adjusted performance. In contrast, NN and ALSTM exhibit slower growth, plateauing at Sharpe ratios of 2.36 and 2.47, respectively.

For long-only portfolios, the results are more nuanced. Although CVRA continues to lead, with the highest Sharpe ratio of 2.04 at $K = 5$, its margin over competitors narrows compared to long-short portfolios. CA and FactorVAE closely follow, with strong performances at higher K, achieving Sharpe ratios of 1.96 and 1.94, respectively. NN and ALSTM deliver lower returns, showing limited improvement at higher K.

Table 5: China Market:Out-of-sample Sharpe ratios of long-short and long-only portfolios (without transaction costs)

	Number of Factors(K)											
Model	1	2	3	4	5	6	1	2	3	4	5	6
	Long-short portfolios						Long-only portfolios					
NN	1.68	1.76	1.95	2.10	2.37	2.36	0.92	1.24	1.60	1.75	1.83	1.77
CA	1.72	1.94	2.17	2.44	2.73	2.66	1.07	1.32	1.67	1.84	1.96	1.95
FactorVAE	1.70	1.86	2.07	2.35	2.67	2.71	1.02	1.33	1.64	1.72	1.94	1.90
ALSTM	1.65	1.77	1.98	2.25	2.45	2.47	0.88	1.26	1.55	1.62	1.85	1.83
CVRA	1.77	2.01	2.29	2.64	2.79	2.85	1.22	1.56	1.73	1.81	2.04	2.01

Besides, we incorporate 30 basis points transaction costs to mitigate slippage in small-cap stocks, as Engle et al. (2006), Table 6 shows the adjusted annualized Sharpe ratios for both portfolio types with equal weights, factoring in these costs. Compared to Table 5, which excluded transaction costs, the Sharpe ratios here are notably lower, reflecting the impact of trading costs on portfolio performance.

For long-short portfolios, CVRA remains the best-performing model, achieving the highest adjusted Sharpe ratio of 1.44 at $K = 5$, though slightly reduced compared to its unadjusted value of 2.85. The reduction is consistent across all models, but CVRA retains its leading edge. FactorVAE and CA follow, with peak adjusted Sharpe ratios of 1.23 and 1.33 at $K = 5$ and $K = 6$, respectively. In long-only portfolios, CVRA also dominates, peaking at 1.32 for $K = 5$. However, the impact of transaction costs appears less pronounced in long-only strategies, as these portfolios typically incur fewer trades. Other models, such as CA and FactorVAE, perform reasonably well but remain behind CVRA.

Table 7 and Table 8 show the annualized Sharpe ratios for both portfolio types without and with transcation costs in Japan market, respectively. Corroborating the Sharpe Ratios from Table 5 and Table 6, our model still dominates other machine learning models.

5.5. Characteristics importance

Table 9 shows which part of overall 10 most influential characteristics are chosen by the individual model in China market. We notice that liquidity-related features such as std_dolvol, std_turn, and baspread, and price trend indicators such as mom1m, mom6m, and mom12m are universally chosen by

Table 6: China Market: Out-of-sample Sharpe ratios of long-short and long-only portfolios (with 30 bps transaction costs)

	Number of Factors(K)											
Model	1	2	3	4	5	6	1	2	3	4	5	6
	Long-short portfolios						Long-only portfolios					
NN	0.63	0.68	0.79	0.96	1.04	1.07	0.53	0.76	0.89	0.97	1.04	1.12
CA	0.74	0.98	1.05	1.23	1.33	1.33	0.65	0.83	0.96	1.15	1.26	1.25
FactorVAE	0.81	1.00	1.09	1.17	1.23	1.20	0.62	0.74	0.87	0.96	1.06	1.04
ALSTM	0.77	0.94	1.01	1.05	1.13	1.13	0.56	0.69	0.88	0.92	1.07	1.05
CVRA	0.87	1.16	1.23	1.36	1.44	1.40	0.72	0.92	1.13	1.24	1.32	1.28

Table 7: Japan Market: Out-of-sample Sharpe ratios of long-short and long-only portfolios (without transaction costs)

	Number of Factors(K)											
Model	1	2	3	4	5	6	1	2	3	4	5	6
	Long-short portfolios						Long-only portfolios					
NN	1.88	1.95	2.14	2.30	2.57	2.52	1.02	1.15	1.71	1.93	1.93	1.85
CA	1.92	2.04	2.25	2.58	2.86	2.83	1.14	1.36	1.73	1.94	2.05	2.02
FactorVAE	1.81	1.95	2.12	2.37	2.70	2.69	1.12	1.43	1.77	1.88	2.04	1.99
ALSTM	1.73	1.87	2.08	2.37	2.56	2.53	0.96	1.36	1.59	1.71	1.89	1.93
CVRA	1.95	2.08	2.39	2.74	2.89	2.88	1.16	1.46	1.80	1.94	2.14	2.11

Table 8: Japan Market: Out-of-sample Sharpe ratios of long-short and long-only portfolios (with 30 bps transaction costs)

	Number of Factors(K)											
Model	1	2	3	4	5	6	1	2	3	4	5	6
	Long-short portfolios						Long-only portfolios					
NN	0.72	0.80	1.00	1.11	1.20	1.20	0.60	0.77	0.93	1.07	1.16	1.15
CA	0.81	1.02	1.13	1.27	1.38	1.36	0.72	0.86	1.01	1.16	1.29	1.27
FactorVAE	0.86	1.06	1.13	1.21	1.27	1.25	0.68	0.80	0.92	1.00	1.12	1.09
ALSTM	0.81	0.99	1.07	1.11	1.16	1.15	0.60	0.73	0.92	0.95	1.08	1.07
CVRA	0.93	1.18	1.26	1.38	1.47	1.44	0.76	0.97	1.16	1.27	1.36	1.33

most models, reflecting their critical role in explaining market dynamics.

In contrast, certain characteristics exhibit variability in selection. For instance, idiovol is identified by CA and CVRA, while zerotrade is selected by CA, FactorVAE, and CVRA, indicating divergent interpretations of idiosyncratic risk. Similarly, rd_sale—a growth indicator—is excluded by NN but recognized by the remaining models. In China market, liquidity and price trends dominate the influential characteristics.

Table 9: Top 10 most influential characteristics chosen by individual models in China

Characteristic	**Category**	**NN**	**CA**	**FactorVAE**	**ALSTM**	**CVRA**
std_dolvol	Liquidity	✓	✓	✓	✓	✓
mom1m	Price trend	✓	✓	✓	✓	✓
baspread	Liquidity	✓	✓	✓	✓	✓
std_turn	Liquidity	✓	✓	✓	✓	✓
mom6m	Price trend	✓	✓	✓	✓	✓
beta	Risk measure	✓	✓	✓	✓	✓
mom12m	Price trend	✓	✓	✓	✓	✓
idiovol	Risk measure		✓			✓
zerotrade	Risk measure		✓	✓		✓
rd_sale	Growth		✓	✓	✓	✓

The results from the Japan market (Table 10) highlight liquidity (std_turn, std_dolvol, baspread) and risk measures (beta, zerotrade) as dominant characteristics, with universal or near-universal selection across models. However, differences emerge compared to China. For example, mom1m and mom6m (price trend indicators) are less universally chosen in Japan, with mom1m omitted by NN. Additionally, rd_sale is selected less consistently, reflecting a reduced emphasis on growth compared to China.Besides, a unique feature in Japan is the inclusion of mvel1 and atr, highlighting a preference for alternative liquidity metrics not prominent in China. These differences may reflect market-specific dynamics and differences between China and Japan market.

6. Conclusion

In this paper, we show how to learn an effective dynamic factor model with Variational autoencoder and recurrent structure for predicting cross-sectional stock returns. Specifically, in view of the noisy market environment, we propose a dynamic factor model based on conditional variational recurrent

Table 10: Top 10 most influential characteristics chosen by individual models in Japan

Characteristic	**Category**	**NN**	**CA**	**FactorVAE**	**ALSTM**	**CVRA**
std_turn	Liquidity	✓	✓	✓	✓	✓
beta	Risk measure	✓	✓	✓	✓	✓
std_dolvol	Liquidity	✓	✓	✓	✓	✓
atr	Liquidity	✓	✓	✓		✓
zerotrade	Risk measure	✓	✓	✓		✓
rd_sale	Growth		✓		✓	✓
mvel1	Liquidity	✓		✓	✓	✓
mom1m	Price trend		✓	✓	✓	✓
baspread	Liquidity	✓	✓	✓	✓	✓
mom6m	Price trend		✓	✓	✓	✓

autoencoder (CVRA). By treating factors as the latent random variables in CVRA, the proposed model with inherent randomness can model the noisy data and estimate stock risk. We extract useful factor exposures from a lot of firm characteristics and obtain factor loadings relying on the estimation of the conditional distribution of the returns by CVRA. Our statistical performance with higher total and predictive R^2 and financial performance with higher out-of-sample Sharpe Ratio demonstrate the effectiveness of our model. Overall, our model serves as a new tool for asset pricing, and its superior capacity allows us to get informative features from rich datasets.

Appendix A. Table of Firm Characteristics

Table A.11: Details on stock characteristics

No.	Acronym	Stock Characteristics	Frequency
1	absac	Absolute accruals	Semi-annual
2	acc	Working capital accruals	Semi-annual
3	agr	Asset growth	Annual
4	atr	Abnormal Turnover Ratio	Monthly
5	beta	Beta	Monthly
6	bm	Book-to-market	Quarterly
7	baspread	Bid-ask spread	Monthly
8	cash	Cash holdings	Quarterly
9	cashdebt	Cash flow to debt	Quarterly
10	cfp	Cash flow to price ratio	Quarterly
11	chato	Change in asset turnover	Quarterly
12	chcsho	Change in shares outstanding	Monthly
13	chinv	Change in inventory	Quarterly
14	chpm	Change in profit margin	Quarterly
15	dy	Dividend to price	Annual
16	egr	Growth in common shareholder equity	Quarterly
17	er-trend	Trend factor	Monthly
18	gma	Gross profitability	Quarterly
19	idivool	Idiosyncratic return volatility	Monthly
20	ill	Illiquidity	Monthly
21	invest	Capital expenditures and inventory	Annual
22	lev	Leverage	Quarterly
23	lgr	Growth in long-term debt	Quarterly
24	mom1m	1-month momentum	Monthly
25	mom6m	6-month momentum	Monthly
26	mom12m	12-month momentum	Monthly
27	mvell	Size	Monthly
28	nincr	Number of earnings increases	Quarterly
29	pchcurrat	% Change in current ratio	Quarterly
30	pchgm_pchsale	% Change in gross margin - % change in sales	Quarterly

Table A.12: Details on stock characteristics (continued)

31	pchsale_pchinv	% Change in sales - % change in inventory	Quarterly
32	pchsaleinv	% Change sales-to-inventory	Quarterly
33	pricedelay	Price delay	Monthly
34	quick	Quick ratio	Quarterly
35	rd	R&D increase	Quarterly
36	rd_sale	R&D to sales	Quarterly
37	volatility	Return volatility	Monthly
38	roaq	Return on assets	Quarterly
39	roeq	Return on equity	Quarterly
40	salecash	Sales to cash	Quarterly
41	saleinv	Sales to inventory	Quarterly
42	sgr	Sales growth	Quarterly
43	sp	Sales to price	Quarterly
44	std_dolvol	Volatility of liquidity (yuan trading volume)	Monthly
45	std_turn	Volatility of liquidity (share turnover)	Monthly
46	zerotrade	Zero trading days	Monthly

Appendix B. Robustness check

In this section, we check whether our best model CVRA has a robust performance with respect to the selection of assets in the training and testing samples. To be specific, we split all stocks into two groups with odd or even permnos from China market, respectively. Accordingly, we divide the training and test sample into two subsamples. Then, we retrain our model with $K = 5$ on each sub-training sample and assess the out-of-sample performance on each sub-testing sample. Table B.13 reports the out-of-sample results of total and predictive R^2's and long-short and long-only Sharpe ratios without transaction costs under four different scenarios. It can be seen that the performance of CVRA model is robust with respect to the sub-training and sub-testing samples, even when the sub-training and sub-testing samples are not the same.

Table B.13: Statistical and Economic Performance for Robustness Check

	Statistical performance			
	Total R^2		Predictive R^2	
	Odd	Even	Odd	Even
Odd	15.26	16.30	1.12	1.14
Even	16.11	15.74	1.26	1.24
	Economic performance			
	Sharpe ratio for Long-short portfolios		Sharpe ratio for Long-only portfolios	
	Odd	Even	Odd	Even
Odd	2.54	2.62	2.01	1.97
Even	2.71	2.66	2.05	2.02

References

Callot, L., Caner, M., Önder, A.Ö., Ulaşan, E., 2021. A nodewise regression approach to estimating large portfolios. Journal of Business & Economic Statistics 39, 520–531.

Caner, M., Daniele, M., 2022. Deep learning based residuals in non-linear factor models: Precision matrix estimation of returns with low signal-to-noise ratio. arXiv preprint arXiv:2209.04512 .

Chan, K., Cheng, S., Hameed, A., 2022. Investor heterogeneity and liquidity. Journal of Financial and Quantitative Analysis 57, 2798–2833.

Chung, J., Kastner, K., Dinh, L., Goel, K., Courville, A.C., Bengio, Y., 2015. A recurrent latent variable model for sequential data. Advances in neural information processing systems 28.

Dey, P., Nahar, N., Hossain, B., 2019. News impact on stock trend. International Journal of Education and Management Engineering 9, 40–49.

Duan, Y., Wang, L., Zhang, Q., Li, J., 2022. Factorvae: A probabilistic dynamic factor model based on variational autoencoder for predicting cross-sectional stock returns, in: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 4468–4476.

Engle, R.F., Ferstenberg, R., Russell, J.R., 2006. Measuring and modeling execution cost and risk .

Fabius, O., Van Amersfoort, J.R., 2014. Variational recurrent auto-encoders. arXiv preprint arXiv:1412.6581 .

Ferreira, A., Kunh, S.S., Fagnani, K.C., De Souza, T.A., Tonezer, C., Dos Santos, G.R., Coimbra-Araújo, C.H., 2018. Economic overview of the use and production of photovoltaic solar energy in brazil. Renewable and Sustainable Energy Reviews 81, 181–191.

Fischer, T., Krauss, C., 2018. Deep learning with long short-term memory networks for financial market predictions. European journal of operational research 270, 654–669.

Freyberger, J., Neuhierl, A., Weber, M., 2020. Dissecting characteristics nonparametrically. The Review of Financial Studies 33, 2326–2377.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. The Review of Financial Studies 33, 2223–2273.

Gu, S., Kelly, B., Xiu, D., 2021. Autoencoder asset pricing models. Journal of Econometrics 222, 429–450.

Jensen, M.C., 1972. Capital markets: Theory and evidence. The Bell Journal of Economics and Management Science , 357–398.

Jones, C.M., Shi, D., Zhang, X., Zhang, X., 2021. Understanding retail investors: evidence from china. Available at SSRN 3628809.

Kelly, B.T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics 134, 501–524.

Kingma, D.P., 2013. Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 .

Leippold, M., Wang, Q., Zhou, W., 2022. Machine learning in the chinese stock market. Journal of Financial Economics 145, 64–82.

Luo, R., Zhang, W., Xu, X., Wang, J., 2018. A neural stochastic volatility model, in: proceedings of the AAAI conference on artificial intelligence.

Naseem, S., Mohsin, M., Hui, W., Liyan, G., Penglai, K., 2021. The investor psychology and stock market behavior during the initial era of covid-19: a study of china, japan, and the united states. Frontiers in Psychology 12, 626934.

Nelson, D.M., Pereira, A.C., De Oliveira, R.A., 2017. Stock market's price movement prediction with lstm neural networks, in: 2017 International joint conference on neural networks (IJCNN), Ieee. pp. 1419–1426.

Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., Cottrell, G., 2017. A dual-stage attention-based recurrent neural network for time series prediction. arXiv preprint arXiv:1704.02971 .

Ross, S.A., 1976. The arbitrage theory of capital asset pricing. Journal of Economic Theory 13, 341–360.

Sharpe, W.F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. The journal of finance 19, 425–442.

Su, Y., Li, J., Yang, B., An, Y., 2023. The impacts of policy uncertainty on asset prices: Evidence from china's market. Asia-Pacific Financial Markets , 1–47.

Tan, L., Zhang, X., Zhang, X., 2023. Retail and institutional investor trading behaviors: Evidence from china. Annual Review of Financial Economics 16.

Urasawa, S., 2018. Structural change and business cycles in japan: Revisiting the stylized facts. Journal of Business Cycle Research 14, 243–281.

Weber, R., 2023. Institutional investors, households, and the time-variation in expected stock returns. Journal of Financial and Quantitative Analysis 58, 352–391.

Xiong, X., Wang, C., Shen, D., 2020. Market participation willingness and investor's herding behavior: evidence from an emerging market. Asia-Pacific Financial Markets 27, 439–452.

Yan, M., Meng, J., Zhou, C., Tu, Z., Tan, Y.P., Yuan, J., 2020. Detecting spatiotemporal irregularities in videos via a 3d convolutional autoencoder. Journal of Visual Communication and Image Representation 67, 102747.

Yang, X., Liu, W., Zhou, D., Bian, J., Liu, T.Y., 2020. Qlib: An ai-oriented quantitative investment platform. arXiv preprint arXiv:2009.11189 .

Yang, X., Zhu, Z., Li, D., Zhu, K., 2024. Asset pricing via the conditional quantile variational autoencoder. Journal of Business & Economic Statistics 42, 681–694.

Ye, J., Goswami, B., Gu, J., Uddin, A., Wang, G., 2024. From factor models to deep learning: Machine learning in reshaping empirical asset pricing. arXiv preprint arXiv:2403.06779 .

Zhang, Y., 2018. China, japan and the us stock markets and the global financial crisis. Asia-Pacific Financial Markets 25, 23–45.

ZHOU, X., Wang, Y., . Learning from memory: Asset pricing via recurrent neural network and attention mechanism. Available at SSRN 4999011 .