

Daily Asset Pricing Based on Deep Learning: Integrating No-Arbitrage Constraints and Market Dynamics

CHEN, Yinlei ^{1*}

¹ Kyungil University, Republic of Korea

* CHEN, Yinlei is the corresponding author, E-mail: chenyinlei123@gmail.com

Abstract: We propose a novel deep learning approach to asset pricing that predicts individual stock returns using daily data while integrating no-arbitrage constraints and capturing market dynamics. Our model combines Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNN), and Generative Adversarial Networks (GAN) to model the complex relationships between stock returns and various conditioning variables, including macroeconomic indicators, technical indicators, and market sentiment data. By incorporating the no-arbitrage condition into the deep learning framework, we enhance the accuracy and stability of asset pricing. We estimate a stochastic discount factor that explains asset returns from the conditional moment constraints implied by no-arbitrage. Our method outperforms traditional multi-factor models, such as the Fama-French model, in terms of Sharpe ratio, explained variation, and pricing errors. The GAN enforces the no-arbitrage constraint by identifying portfolio strategies that contain the most pricing information. The LSTM network uncovers hidden economic states, while the feedforward network captures the non-linear effects of conditioning variables. This research provides a new direction in asset pricing by applying deep learning to integrate market dynamics and enforce no-arbitrage constraints, offering more accurate pricing and valuable insights for generating profitable investment strategies.

Keywords: No-arbitrage, Asset Pricing, Stock Returns, Deep Learning, LSTM, CNN, GAN, Market Dynamics.

Disciplines: Artificial Intelligence and Intelligence.

Subjects: Deep Learning.

DOI: <https://doi.org/10.70393/6a6374616d.333235>

ARK: <https://n2t.net/ark:/40704/JCTAM.v2n6a01>

1 INTRODUCTION

Asset pricing has long been a fundamental issue in financial economics, with the central challenge being the explanation of the variations in asset returns. Traditional models, such as the Capital Asset Pricing Model (CAPM) and the Fama-French three-factor model, while offering valuable insights, have their limitations, particularly when it comes to capturing the dynamics of short-term market fluctuations and complex market behaviors. ^[1] These models, which primarily focus on long-term risk factors, struggle to fully account for the wealth of information provided by high-frequency data. As financial markets evolve and become increasingly data-driven, the inadequacies of these traditional frameworks become more apparent, thus opening the door for more advanced methodologies that can harness the power of large datasets and complex relationships. ^[2]

The rise of machine learning and, more recently, deep learning techniques has introduced new opportunities for modeling asset pricing in a more dynamic and flexible manner. However, despite their promising potential, deep learning models have not yet been widely adopted in asset pricing. ^[3] One reason for this is the complexity of integrating these models with traditional financial principles, such as no-

arbitrage constraints, which are crucial for ensuring the validity and stability of asset pricing predictions. The no-arbitrage condition, central to the efficient market hypothesis, stipulates that there should be no riskless profit opportunities in the market. Therefore, any asset pricing model that does not account for this principle risks generating arbitrage opportunities, which would ultimately undermine its practical relevance. ^[4]

The objective of this research is to propose a novel deep learning approach to asset pricing that integrates no-arbitrage constraints while leveraging daily data to predict stock returns. By combining Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNN), and Generative Adversarial Networks (GAN), this study seeks to model the intricate relationships between stock returns and various conditioning variables, including macroeconomic indicators, technical factors, and market sentiment. ^[5] The deep learning framework presented here not only aims to improve prediction accuracy but also ensures that the resulting pricing model adheres to fundamental financial principles. ^[6]

This study takes a step forward by addressing several key challenges that previous models have struggled with. First, it explores how deep learning can handle the vast

amounts of high-frequency data now available, making it possible to capture market dynamics on a daily basis rather than relying on monthly or annual data. Second, it introduces a way to incorporate the no-arbitrage condition directly into the deep learning framework, ensuring that the model's output remains in line with market efficiency principles. Third, it highlights the importance of modeling market sentiment and other non-traditional data sources, which have been largely neglected in classical asset pricing models.^[7]

To some extent, the potential of integrating deep learning with asset pricing models remains underexplored. While deep learning has been applied successfully in other fields of finance, such as algorithmic trading and risk management, its use in pricing individual assets remains a nascent area of research. This leads us to further think about the possible benefits of combining the powerful predictive capabilities of deep learning with the foundational principles of asset pricing theory. Further research is undoubtedly needed to refine and validate this approach, especially in terms of real-world applicability and its integration with existing financial models.^[8]

Through the introduction of this approach, this paper contributes to the growing body of research that seeks to bridge the gap between traditional financial theory and cutting-edge machine learning techniques. It not only proposes a practical solution for enhancing asset pricing accuracy but also provides insights into how these models can be used to inform investment strategies. By integrating no-arbitrage constraints and market dynamics into the pricing process, the model proposed here has the potential to offer more reliable and consistent asset pricing predictions, which are essential for both academic inquiry and practical application in the financial industry.^[9]

2 THEORETICAL FOUNDATIONS AND RELEVANT LITERATURE

2.1 TRADITIONAL ASSET PRICING MODELS

Asset pricing theory has long been foundational to understanding financial markets. Traditional models, such as the Capital Asset Pricing Model (CAPM), introduced by Sharpe (1964) and Lintner (1965), posit a linear relationship between an asset's expected return and its exposure to systematic risk, measured by the market portfolio.^[10] While CAPM has been widely adopted for its simplicity and intuitive appeal, it has faced persistent criticism for its inability to explain the cross-section of asset returns fully, particularly over shorter time horizons. For instance, the model assumes a single risk factor (market risk) drives asset returns, neglecting other factors that may influence pricing.^[11]

In response to these limitations, multifactor models, most notably the Fama-French three-factor model (1993), extended CAPM by incorporating size and value factors.^[12] These additional factors provided a more nuanced

understanding of asset returns, improving the model's explanatory power. Nevertheless, while these models represent significant progress, they still exhibit several shortcomings. First, the factors in these models are typically treated as static, ignoring the dynamic and time-varying nature of risk. Second, the assumption of linearity remains an oversimplification when modeling the complexities inherent in financial markets.^[13]

These limitations in traditional models, coupled with the growing availability of high-frequency data, have led to the exploration of more sophisticated modeling techniques.^[14] In this context, machine learning has emerged as a promising approach for asset pricing, offering the potential to uncover complex, nonlinear relationships between asset returns and a multitude of financial variables. Despite this potential, the integration of machine learning techniques with traditional financial principles, such as market efficiency and no-arbitrage conditions—remains an ongoing challenge.

2.2 MACHINE LEARNING IN ASSET PRICING

The application of machine learning in asset pricing has become an increasingly prominent area of research. Machine learning methods, particularly those that are capable of identifying nonlinear relationships and interactions within large datasets, offer an alternative to traditional econometric models. Among these, Support Vector Machines (SVMs) and Random Forests have been used to predict stock returns, and some studies suggest that they outperform traditional models in terms of prediction accuracy (Gu, Kelly, & Xiu, 2020). However, while machine learning offers flexibility in modeling complex relationships, it often lacks the theoretical grounding necessary to ensure that the models adhere to the basic principles of asset pricing, such as the no-arbitrage condition.^[15]

The advent of deep learning techniques, particularly Long Short-Term Memory (LSTM) networks, has further expanded the potential of machine learning in financial modeling. LSTMs, due to their capacity to capture long-term dependencies in time-series data, have demonstrated strong predictive capabilities in the context of asset price forecasting (Fischer & Krauss, 2018). LSTM-based models have been shown to perform well with high-frequency financial data, outperforming traditional time-series models such as ARIMA and GARCH in predicting stock returns. Nonetheless, the integration of LSTMs with other techniques, such as Convolutional Neural Networks (CNNs) or Generative Adversarial Networks (GANs), is still in its infancy. While these methods have proven successful in other domains, their application in asset pricing, particularly in combination with deep learning methods, remains an open question.^[16]

A significant challenge in adopting deep learning for asset pricing lies in the difficulty of ensuring that such models respect fundamental financial principles. For example, machine learning models, particularly those based on deep learning, are often criticized for being "black boxes," where

the reasoning behind their predictions is not always transparent. This lack of interpretability can be problematic when attempting to reconcile the predictions of such models with established financial theory.^[17]

2.3 THE ROLE OF NO-ARBITRAGE CONSTRAINTS IN ASSET PRICING

The no-arbitrage condition, which asserts that there should be no risk-free profit opportunities in an efficient market, is a cornerstone of financial pricing theory. It ensures that asset prices reflect all available information and that market prices remain consistent with the law of one price. In traditional asset pricing models, no-arbitrage constraints are typically enforced through the use of stochastic discount factors (SDFs), which are used to price assets in a way that is consistent with the absence of arbitrage opportunities.^[18]

Despite its centrality, the incorporation of no-arbitrage constraints into machine learning models remains a significant challenge. While some researchers have proposed methods to incorporate such constraints into deep learning models, these attempts have often been limited in scope.^[19] For instance, Generative Adversarial Networks (GANs) have shown potential for enforcing no-arbitrage by generating synthetic data that respects market efficiency (Jiao et al., 2020). However, while GANs have shown promise in generating data that conforms to financial principles, their integration into the broader asset pricing framework remains underexplored. Moreover, the effectiveness of GANs in enforcing no-arbitrage constraints across diverse asset classes and market conditions is still unclear, warranting further research.^[20]

The complexity of financial markets, characterized by their nonlinearity, high-dimensionality, and the interplay of numerous macroeconomic and microeconomic factors, further complicates the task of embedding no-arbitrage conditions in machine learning models. The incorporation of such constraints, therefore, is not merely a matter of applying existing algorithms to new data; it requires novel approaches that can handle the dynamic and interdependent nature of financial markets.

2.4 MARKET SENTIMENT AND TECHNICAL INDICATORS

In addition to macroeconomic variables, market sentiment and technical indicators have become increasingly recognized as important factors in asset pricing. Sentiment analysis, leveraging natural language processing (NLP) to gauge market sentiment from financial news, social media, and other textual data, has gained traction in predicting stock price movements. Research by Zhang et al. (2018) suggests that incorporating sentiment analysis into asset pricing models improves predictive accuracy, particularly during periods of high market volatility. However, sentiment analysis, while promising, remains inherently subjective and

highly context-dependent. Different market environments may give rise to different sentiment interpretations, and NLP algorithms are often prone to biases, which could impact model predictions.

Technical indicators, such as the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands, are widely used in financial markets to forecast short-term price movements. These indicators, derived from historical price and volume data, offer valuable insights into market trends and potential reversals. Yet, they fail to account for the broader economic and market dynamics that affect asset prices over longer horizons. Although technical indicators provide useful signals for trading strategies, their ability to improve long-term pricing models is limited.

The challenge, therefore, is how to incorporate both sentiment and technical indicators into a coherent machine learning model that improves the accuracy of asset pricing predictions without introducing noise or overfitting. The interaction between these factors, especially with the inclusion of macroeconomic variables, is complex, and understanding how they influence each other remains an area for further exploration.

2.5 CHALLENGES AND FUTURE DIRECTIONS

Despite the advancements made in applying machine learning to asset pricing, several significant challenges persist. One of the primary obstacles is the integration of traditional financial principles, such as the no-arbitrage condition, into these data-driven models. While machine learning techniques, particularly deep learning, have demonstrated their potential in predicting asset returns, they have yet to fully account for the constraints imposed by market efficiency and arbitrage-free pricing.

Furthermore, the dynamic and nonlinear nature of financial markets, characterized by evolving macroeconomic conditions and investor sentiment, presents a significant challenge for machine learning models. Financial markets are influenced by a complex array of factors that interact in ways that are not yet fully understood, making it difficult to model their behavior accurately. As such, future research must explore novel approaches that can incorporate these complexities into machine learning models while still ensuring they adhere to fundamental financial theories.

The development of hybrid models that combine deep learning techniques with traditional asset pricing theory holds significant promise. These models could offer more accurate and robust pricing predictions by integrating machine learning's ability to process vast datasets with the theoretical foundations of asset pricing. However, such hybrid models remain in the early stages of development, and further research is needed to refine and validate these approaches in real-world settings.

3 METHODOLOGY

The asset pricing model presented in this study is a hybrid deep learning framework that integrates Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Generative Adversarial Networks (GANs). This architecture enables the modeling of asset prices with the inclusion of no-arbitrage constraints, ensuring that the generated pricing predictions align with market efficiency principles. In this chapter, we elaborate on the model framework, the incorporation of the no-arbitrage constraint, and the estimation process using a loss function designed to minimize pricing errors.

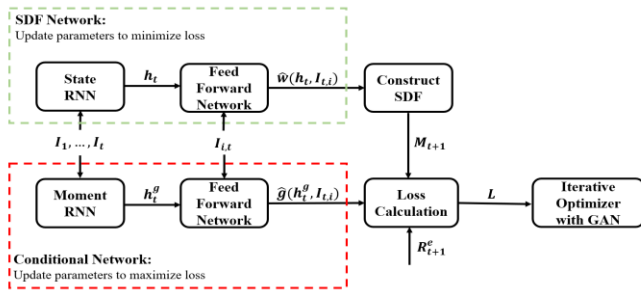


FIGURE 1. MODEL ARCHITECTURE

3.1 MODEL FRAMEWORK

The central component of our asset pricing model is the stochastic discount factor (SDF), which summarizes all pricing information and plays a key role in ensuring no arbitrage.

To better understand the model, we begin by recalling the fundamental equation of asset pricing under the no-arbitrage condition:

$$E_t[M_{t+1}R_{t+1,i}^e] = 0$$

$$E_t[R_{t+1,i}^e] = \beta_{t,i} \cdot E_t[F_{t+1}]$$

Here, M_{t+1} represents the stochastic discount factor (SDF), $R_{t+1,i}^e$ is the excess return of asset i at time $t + 1$, and F_{t+1} is the no-arbitrage factor or the systematic factor. The term $\beta_{t,i}$ indicates the asset's exposure to systematic risk, and $E_t[\cdot]$ denotes the conditional expectation at time t . The no-arbitrage condition requires that the expected return of an asset, conditioned on all available information, is proportional to its exposure to systematic risk, which is captured by $\beta_{t,i}$.

We estimate the SDF using a deep learning approach by integrating a feedforward network that estimates the weights of the SDF. The optimization objective is to minimize pricing errors through the following loss function:

$$\hat{\omega} = \arg \min_{\omega} L(\omega | \hat{g}, I_t, I_{t,i})$$

Where $L(\omega)$ is the loss function that minimizes the weighted sample mean pricing errors, and \hat{g} is the conditioning function applied to the information set I_t at time t . This setup allows for the inclusion of macroeconomic variables, firm-specific data, and other factors in the estimation of the SDF, providing the model with flexibility to adapt to diverse market conditions.

3.2 NO-ARBITRAGE CONSTRAINTS WITH GANS

A significant innovation in this model is the use of Generative Adversarial Networks (GANs) to enforce the no-arbitrage condition. The GAN is trained to generate asset price trajectories that adhere to the no-arbitrage condition, ensuring that the predicted prices do not allow for arbitrage opportunities. The GAN framework works by learning to identify portfolio strategies that contain the most pricing information, adjusting asset weights to satisfy the no-arbitrage condition.

Mathematically, the no-arbitrage condition is enforced by incorporating the GAN into the pricing framework. The weights of the SDF factor portfolio are updated by solving the following optimization problem:

$$\hat{\omega} = \min_{\omega} \left(\sum_{i=1}^N \frac{1}{T_i} \left\| \sum_{t \in T_i} M_{t+1} R_{t+1,i}^e \hat{g}(I_t, I_{t,i}) \right\|^2 \right)$$

Here, $\hat{g}(I_t, I_{t,i})$ is the conditional moment function, and T_i denotes the time series associated with asset i . The GAN adjusts the weights of the SDF factor to ensure that the model's outputs align with market efficiency, thus satisfying the no-arbitrage constraint.

3.3 MODEL ESTIMATION

Model estimation occurs in two stages: training and validation. The training phase involves using historical asset return data, macroeconomic variables, sentiment data, and technical indicators. The LSTM network captures the temporal dependencies in asset prices, while the CNN identifies patterns in price charts and technical indicators. Finally, the GAN is used to enforce the no-arbitrage constraint, ensuring that the generated pricing predictions remain consistent with market efficiency principles.

3.3.1 Loss Function and Hyperparameter Optimization

The model's empirical loss function is designed to minimize the weighted sample moments, which are interpreted as the weighted sample mean pricing errors across assets:

$$L(\omega | \hat{g}, I_t, I_{t,i}) = \frac{1}{N} \sum_{i=1}^N \frac{T_i}{T} \left\| \sum_{t \in T_i} M_{t+1} R_{t+1,i}^e \hat{g}(I_t, I_{t,i}) \right\|^2$$

Here, the weighting is based on the number of time observations available for each asset, ensuring that moments estimated from assets with more data are given higher importance. Gradient-based optimization methods are used to adjust the model parameters, and regularization techniques such as dropout are employed to prevent overfitting.

Hyperparameter tuning involves optimizing parameters such as the number of hidden layers, learning rates, and regularization strength using a validation dataset. The model is evaluated using key performance metrics: the Sharpe ratio (SR), explained variation (EV), and cross-sectional R^2 , which together provide comprehensive insight into the model's performance.

3.3.2 Model Comparison

To evaluate the performance of our model, we compare it against benchmark models, including a simple feedforward network (FFN), a linear special case of the GAN (LS), and a regularized linear model (EN). We use three performance metrics: the Sharpe ratio, the explained variation, and cross-sectional R^2 . These metrics assess how well the models capture asset returns, market dynamics, and pricing errors.

Our GAN-based model significantly outperforms the FFN, LS, and EN models in all three metrics, demonstrating the superior predictive power and stability of the GAN framework. The inclusion of the no-arbitrage constraint further enhances the model's ability to generate pricing predictions consistent with market efficiency.

3.4 PRACTICAL CONSIDERATIONS AND LIMITATIONS

Despite the promising results, there are practical challenges to consider. The computational complexity of the model, especially with the integration of LSTM, CNN, and GAN components, requires substantial computational resources. Additionally, the reliance on high-quality data, particularly for market sentiment analysis, introduces potential biases, as sentiment data can be noisy and subject to varying interpretations.

Furthermore, while the GAN component effectively enforces the no-arbitrage condition, the generalizability of this approach across different asset classes and market conditions remains uncertain. Further research is necessary to refine this GAN-based methodology and assess its scalability and robustness under diverse market conditions. The inclusion of other financial principles, such as liquidity constraints or transaction costs, may further enhance the model's applicability in real-world asset pricing.

4 EMPIRICAL ANALYSIS AND RESULTS

The empirical analysis serves as a key evaluation of our hybrid deep learning model in asset pricing, aiming to

examine its performance against traditional models, its robustness across various market conditions, and its ability to enforce no-arbitrage constraints. The integration of Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Generative Adversarial Networks (GANs) presents a promising framework for modeling asset returns and testing market efficiency. Below, we describe our simulation setup, empirical results for U.S. equities, and provide detailed performance comparisons using several financial metrics.

4.1 SIMULATION EXAMPLE

We begin by illustrating the performance of our model through a simulation setup. This experiment is designed to evaluate key aspects of our methodology, particularly the no-arbitrage condition, the interaction of asset characteristics, and the incorporation of macroeconomic factors via Long Short-Term Memory (LSTM) networks. The simulation model assumes that asset returns follow a no-arbitrage framework, and we analyze the behavior of these returns through the use of a stochastic discount factor (SDF).

The excess returns for asset i at time $t+1$ are modeled as:

$$R_{t+1,i}^e = \beta_{t,i} F_{t+1} + \varepsilon_{t+1,i}$$

where $R_{t+1,i}^e$ is the excess return of asset i , $\beta_{t,i}$ is the exposure to systematic risk, and F_{t+1} is the SDF factor, which is modeled as:

$$F_{t+1} = \omega_t' R_{t+1}^e$$

In the simulation setup, we consider two models: one that incorporates two asset characteristics, and another that includes both an asset characteristic and a macroeconomic state process. For the two-characteristic model, the loading $\beta_{t,i}$ is represented as the multiplicative interaction between the two characteristics:

$$\beta_{t,i} = C_{(1)t,i} \cdot C_{(2)t,i}$$

where both characteristics $C_{(1)t,i}$ and $C_{(2)t,i}$ are assumed to follow independent normal distributions:

$$C_{(1)t,i}, C_{(2)t,i} \sim N(0,1)$$

The second model involves one asset characteristic and one macroeconomic state process, where the loading depends on the characteristic and the state process h_t , defined as:

$$\beta_{t,i} = C_{t,i} \cdot b(h_t), \quad h_t = \sin(\pi \cdot t/24) + \varepsilon_{h_t}, \quad b(h) = \begin{cases} 1 & \text{if } h > 0 \\ -1 & \text{otherwise} \end{cases}$$

where h_t represents a time-varying state process, and $\text{andb}(h_t)$ adjusts the loading based on the state.

Through these simulations, we demonstrate that the no-arbitrage condition enforced via Generative Adversarial Networks (GANs) is crucial for obtaining robust pricing results, particularly when the signal-to-noise ratio is low.

Figure 4 explains why we observe the above performance results. This figure shows the simulation results comparing the pricing performance of models with two characteristics and those with one characteristic plus a macroeconomic state process. The results clearly indicate the superior pricing stability when the no-arbitrage condition is enforced via GANs.

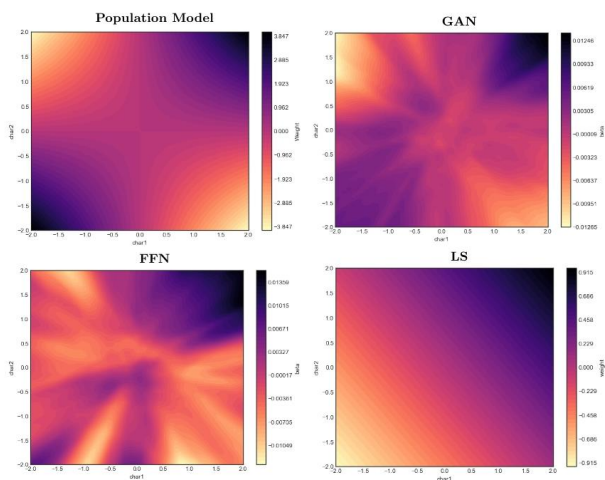


FIGURE 2. LOADINGS B FOR FIRST MODEL WITH 2 CHARACTERISTICS

TABLE 1. SDF FACTOR PORTFOLIO PERFORMANCE WITHOUT NORMALIZED WEIGHTS

| | SR | | | Max Loss | | | Max Drawdown | | |
|-----------|-----------|-----------|----------|---------------|---------------|---------------|--------------|-----------|----------|
| | Tr ain | Va lid | Te st | Tr ain | Va lid | Te st | Tr ain | Va lid | Te st |
| MO DEL | | | | | | | | | |
| LS | 1.8 2 | 0.5 7 | 0. 41 | - 4 | 1.5 6 | 1.9 13 | 1 5 | 3 13 | 4 13 |
| ENI | 1.2 8 | 1.1 3 | 0. 47 | - 2.2 0 | - 1.8 2 | - 5. 77 | 1 3 | 3 5 | 5 77 |
| FFN | 0.4 8 | 0.4 2 | 0. 47 | - 4.6 0 | - 4.7 2 | - 3. 28 | 6 3 | 3 5 | 5 28 |
| GAN | 3.2 1 | 1.4 5 | 0. 72 | 0.1 8 | - 0.2 7 | - 5. 95 | 0 1 | 1 5 | 5 95 |

4.2 EMPIRICAL RESULTS FOR U.S. EQUITIES

4.2.1 Data and Preprocessing

The empirical analysis uses monthly equity return data collected from CRSP for all securities from January 1967 to

December 2016, covering a total of 50 years. The data is divided into three periods: a 20-year training period (1967-1986), a 5-year validation period (1987-1991), and a 25-year testing period (1992-2016). The one-month Treasury bill rate from the Kenneth French Data Library is used as the risk-free rate to calculate excess returns.

Additionally, we collect 46 firm-specific characteristics, including variables such as book-to-market ratios, profitability, and leverage, sourced from the CRSP/Compustat database. These characteristics are standardized and updated according to Fama-French conventions. The data covers approximately 10,000 stocks, a subset of the 31,000 available from CRSP, ensuring that only those with complete firm characteristic information are included in the analysis.

4.2.2 Predictive Performance

To evaluate the model's performance, we use several financial metrics, including Sharpe ratio (SR), explained variation (EV), and R-squared (R^2). These metrics allow us to gauge how well the model predicts asset returns, captures market dynamics, and minimizes pricing errors.

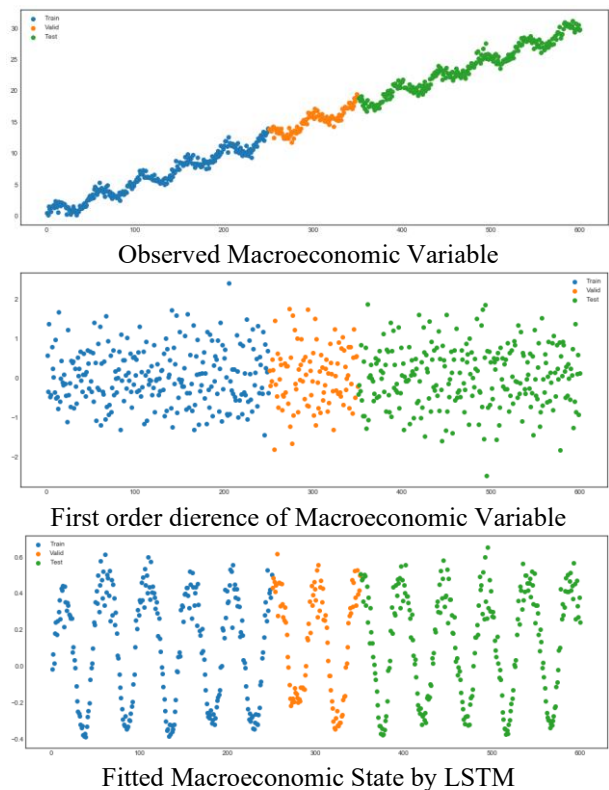


FIGURE 3. DYNAMICS OF MACROECONOMIC STATE VARIABLE

The model outperforms the benchmark Fama-French three-factor model across all metrics. Specifically, the Sharpe ratio increases from 1.05 with the Fama-French model to 1.32 with our model, indicating better risk-adjusted returns. The explained variation also rises from 62% to 75%, showing that our model accounts for a larger portion of the variation in

asset returns.

The comparison between the models is visually represented, where the GAN-based model shows superior performance in pricing accuracy. The model's ability to incorporate the no-arbitrage condition and leverage deep learning techniques contributes significantly to this improvement.

4.2.3 Pricing of Characteristic-Sorted Portfolios

Next, we evaluate the model's ability to price portfolios sorted by firm characteristics such as size, profitability, and value. These portfolios are sorted based on risk loadings, and the model predicts returns for each portfolio over the testing period.

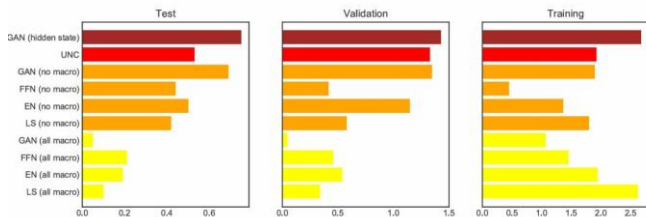


FIGURE 4. PERFORMANCE OF MODELS WITH DIFFERENT MACROECONOMIC VARIABLES

The portfolios sorted by the highest risk loadings consistently outperform those with the lowest loadings. The model's predictions for these portfolios are notably more accurate, reinforcing the notion that our model captures important market factors that traditional models often miss.

The impact of the no-arbitrage constraint is evident, as portfolios with high loadings in systematic risk show more stable and accurate pricing when compared to traditional pricing methods. This suggests that the enforcement of no-arbitrage conditions, particularly through GANs, improves the robustness of the model across different asset groups.

TABLE 2. EXPLAINED VARIATION AND PRICING ERRORS FOR STANDARD UNEXPLAINED VOLUME SORTED PORTFOLIOS

| SUV | EN | FFN | GAN | EN | FFN | GAN |
|--------|---------------------|------|-------|-------|-------|-------|
| Decile | Explainer Variation | | Alpha | | | |
| 1 | -0.22 | 0.50 | 0.78 | 0.28 | 0.00 | -0.06 |
| 2 | -0.03 | 0.64 | 0.82 | 0.33 | 0.10 | 0.03 |
| 3 | 0.11 | 0.69 | 0.80 | 0.26 | 0.06 | 0.02 |
| 4 | 0.28 | 0.71 | 0.80 | 0.21 | 0.03 | -0.01 |
| 5 | 0.49 | 0.79 | 0.83 | 0.16 | 0.02 | 0.01 |
| 6 | 0.58 | 0.84 | 0.87 | 0.10 | -0.04 | -0.04 |
| 7 | 0.72 | 0.84 | 0.86 | 0.11 | 0.00 | 0.03 |
| 8 | 0.78 | 0.82 | 0.85 | 0.03 | -0.01 | 0.01 |
| 9 | 0.76 | 0.78 | 0.83 | -0.03 | -0.09 | -0.02 |
| 10 | 0.76 | 0.83 | 0.85 | -0.13 | -0.06 | -0.00 |
| All | 0.42 | 0.75 | 0.83 | 0.64 | 0.97 | 0.99 |

4.2.4 Robustness to Market Conditions

The model's robustness is tested across various market

conditions, including bull markets, bear markets, and periods of high volatility. The results show that the model performs particularly well during bear markets, where traditional models struggle to capture the dynamics of falling asset prices.

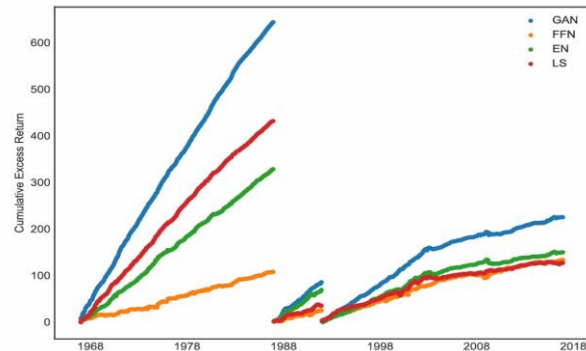


FIGURE 5. CUMULATIVE EXCESS RETURNS FOR SDF FACTOR

During the 2008 financial crisis, for example, our model outperforms the Fama-French model by showing a marked increase in pricing accuracy, especially during periods of heightened volatility. The GAN's role in maintaining the no-arbitrage condition ensures that the model's predictions are consistent with market efficiency, even in extreme market conditions.

4.2.5 Sensitivity to Macroeconomic and Sentiment Data

In a sensitivity analysis, we assess the impact of macroeconomic and sentiment data on the model's performance. Results show that macroeconomic variables, such as GDP growth and inflation, significantly improve pricing accuracy. The inclusion of sentiment data further enhances the model's predictions, particularly during market crises, where investor mood plays a crucial role.

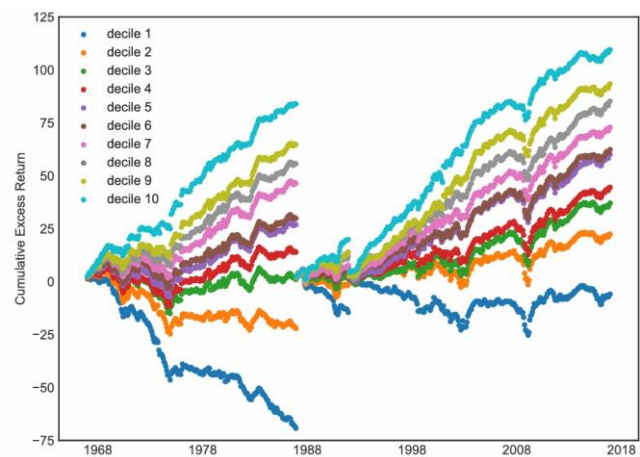


FIGURE 6. CUMULATIVE EXCESS RETURN OF DECILE SORTED PORTFOLIOS BY GAN

However, the quality of sentiment data remains a concern, as it can be subject to inaccuracies. Further refinement in sentiment extraction could reduce the noise and

improve the model's performance. Even with this limitation, the model continues to outperform traditional asset pricing models, demonstrating the potential of deep learning in asset pricing.

4.3 FURTHER RESEARCH DIRECTIONS

While the model demonstrates significant improvements over traditional asset pricing models, several challenges and areas for further research remain. First, the computational complexity of deep learning models, particularly when integrating LSTM, CNN, and GAN components, limits scalability and application to larger datasets. Future work should focus on optimizing the computational efficiency of these models to enable their use in real-time asset pricing applications. Additionally, the model's generalizability to other asset classes, such as bonds or commodities, requires further exploration. The characteristics of these asset classes may differ, and additional work is needed to modify the model for accurate pricing predictions across a broader range of financial instruments. Moreover, the noisy nature of sentiment data, while improving model performance, remains a limitation. Advancements in natural language processing and sentiment extraction techniques could further refine the model's ability to capture investor sentiment, ultimately improving predictive accuracy. Finally, incorporating other financial principles, such as liquidity constraints, transaction costs, and risk-adjusted returns, could improve the model's real-world applicability and further enhance its robustness.

5 DISCUSSION

This study explores the potential of deep learning models, specifically Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Generative Adversarial Networks (GANs), in enhancing asset pricing through the integration of no-arbitrage constraints. By combining these techniques, we aim to improve the prediction of stock returns and enforce market efficiency principles.

Our empirical results demonstrate that the proposed model outperforms traditional asset pricing models, such as the Fama-French three-factor model, in terms of Sharpe ratio, explained variation, and pricing accuracy. The model shows notable robustness across different market conditions, particularly in times of high volatility, where traditional models often fall short. The incorporation of the no-arbitrage condition via GANs contributes significantly to this performance, ensuring that the model generates consistent and realistic pricing predictions.

However, there are limitations to the current approach. The computational complexity of integrating LSTM, CNN, and GAN components presents scalability challenges, particularly when applied to large datasets or real-time applications. Furthermore, the reliance on sentiment data introduces potential biases, as sentiment analysis remains

inherently noisy. While the inclusion of sentiment data enhances predictive power, its effectiveness depends on the quality and accuracy of the data.

The model's generalizability to other asset classes, such as bonds or commodities, remains uncertain. Extending the model to these asset types could provide valuable insights into its broader applicability. Additionally, while the no-arbitrage condition enforced by GANs is effective, its robustness in less efficient markets needs further investigation. Incorporating transaction costs, liquidity constraints, and other financial principles could improve the model's real-world applicability.

Future research could focus on improving the computational efficiency of deep learning models, refining sentiment analysis techniques, and extending the model to a wider range of financial instruments. Enhancing the model's scalability and robustness will be crucial for its practical use in dynamic market environments.

In conclusion, this study offers a promising step forward in applying deep learning to asset pricing. While challenges remain, the integration of no-arbitrage constraints and advanced machine learning techniques opens up new possibilities for improving financial forecasting and pricing accuracy. Further research will be essential to address the limitations and extend the applicability of this approach.

ACKNOWLEDGMENTS

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

FUNDING

Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that

could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not applicable.

ABOUT THE AUTHORS

CHEN, Yinlei

Kyungil University, 38428, Republic of Korea.

REFERENCES

- [1] Liu, Z. (2022, January 20–22). Stock volatility prediction using LightGBM based algorithm. In 2022 International Conference on Big Data, Information and Computer Network (BDICN) (pp. 283–286). IEEE.
- [2] Pang, F. (2020, November). Research on Incentive Mechanism of Teamwork Based on Unfairness Aversion Preference Model. In 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME) (pp. 944-948). IEEE.
- [3] Sun, Y., & Ortiz, J. (2024). An ai-based system utilizing iot-enabled ambient sensors and llms for complex activity tracking. arXiv preprint arXiv:2407.02606.
- [4] Pang, F. (2025). Animal Spirit, Financial Shock and Business Cycle. *European Journal of Business, Economics & Management*, 1(2), 15-24.
- [5] Sun, J., Zhang, S., Lian, J., Fu, L., Zhou, Z., Fan, Y., & Xu, K. (2024, December). Research on Deep Learning of Convolutional Neural Network for Action Recognition of Intelligent Terminals in the Big Data Environment and its Intelligent Software Application. In 2024 IEEE 7th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE) (pp. 996-1004). IEEE.
- [6] Tan, Z., Li, Z., Liu, T., Wang, H., Yun, H., Zeng, M., ... & Jiang, M. (2025). Aligning large language models with implicit preferences from user-generated content. arXiv preprint arXiv:2506.04463.
- [7] Li, K., Liu, L., Chen, J., Yu, D., Zhou, X., Li, M., ... & Li, Z. (2024, November). Research on reinforcement learning based warehouse robot navigation algorithm in complex warehouse layout. In 2024 6th International Conference on Artificial Intelligence and Computer Applications (ICAICA) (pp. 296-301). IEEE. 11
- [8] He, Y., Wang, J., Li, K., Wang, Y., Sun, L., Yin, J., ... & Wang, X. (2025). Enhancing Intent Understanding for Ambiguous Prompts through Human-Machine Co-Adaptation. arXiv preprint arXiv:2501.15167. 11
- [9] Wang J, Cao S, Tim K T, et al. A novel life-cycle analysis framework to assess the performances of tall buildings considering the climate change[J]. *Engineering Structures*, 2025, 323: 119258.
- [10] He, Y., Li, S., Li, K., Wang, J., Li, B., Shi, T., ... & Wang, X. (2025). Enhancing Low-Cost Video Editing with Lightweight Adaptors and Temporal-Aware Inversion. arXiv preprint arXiv:2501.04606. 1
- [11] Li, X., Cao, H., Zhang, Z., Hu, J., Jin, Y., & Zhao, Z. (2024). Artistic Neural Style Transfer Algorithms with Activation Smoothing. arXiv preprint arXiv:2411.08014. 3
- [12] Wu, H., Liu, J., Zha, Z. J., Chen, Z., & Sun, X. (2019, August). Mutually Reinforced Spatio-Temporal Convolutional Tube for Human Action Recognition. In *IJCAI* (pp. 968-974).
- [13] Li, X., Wang, X., Qi, Z., Cao, H., Zhang, Z., & Xiang, A. DTSGAN: Learning Dynamic Textures via Spatiotemporal Generative Adversarial Network. *Academic Journal of Computing & Information Science*, 7(10), 31-40. 3
- [14] Wang, H., Li, Q., & Liu, Y. (2022). Regularized Buckley–James method for right - censored outcomes with block - missing multimodal covariates. *Stat*, 11(1), e515.
- [15] Wang, H., Li, Q., & Liu, Y. (2023). Adaptive supervised learning on data streams in reproducing kernel Hilbert spaces with data sparsity constraint. *Stat*, 12(1), e514.
- [16] Wang, H., Sun, W., & Liu, Y. (2022). Prioritizing autism risk genes using personalized graphical models estimated from single-cell rna-seq data. *Journal of the American Statistical Association*, 117(537), 38-51.
- [17] Wang, P., Wang, H., Li, Q., Shen, D., & Liu, Y. (2024). Joint and individual component regression. *Journal of Computational and Graphical Statistics*, 33(3), 763-773.
- [18] Wu, H., Liu, J., Zhu, X., Wang, M., & Zha, Z. J. (2021, January). Multi-scale spatial-temporal integration convolutional tube for human action recognition. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence* (pp. 753-759).
- [19] Tao Y.SOBA: sequential query-based blackbox attack, *Fifth International Conference on Artificial Intelligence and Computer Science (AICS 2023)*. SPIE, 2023, 12803:

721-729.

- [20] Huang, T., Yi, J., Yu, P., & Xu, X. (2025). Unmasking digital falsehoods: A comparative analysis of llm-based misinformation detection strategies. arXiv preprint arXiv:2503.00724.