공학박사학위논문

Graph-based Empirical Asset Pricing : Impact of Network Connectedness

그래프 기반 실증적 자산 가격 결정 : 네트워크 구조의 영향을 중심으로

2022 년 8 월

서울대학교 대학원

산업공학과

손 범 호

Graph-based Empirical Asset Pricing : Impact of Network Connectedness

그래프 기반 실증적 자산 가격 결정 : 네트워크 구조의 영향을 중심으로

지도교수 이 재 욱

이 논문을 공학박사 학위논문으로 제출함

2022 년 7 월

서울대학교 대학원

산업공학과

손 범 호

손범호의 공학박사 학위논문을 인준함

2022 년 8 월

위 원 장	이 덕 주	(인)
부위원장	이 재 욱	(인)
위 원	장 우 진	(인)
위 원	김 남 형	(인)
위 원	장 희 수	(인)

Abstract

Graph-based Empirical Asset Pricing : Impact of Network Connectedness

Bumho Son
Department of Industrial Engineering
The Graduate School
Seoul National University

Financial assets are always exposed to risks. It is important to evaluate the risk properly and figure out how much each asset is compensated for its risk. Asset pricing model explains the behavior of financial asset return by evaluating the risk and risk exposure of asset return. We focused on factor model structure among asset pricing models, which explains excess return through factor and beta coefficients. While conventional factor models estimate factor or beta through various macroeconomic variables or firm-specific variables, there exist fewer studies considering the connectedness between assets. Since financial assets have connected dynamics, asset returns should be priced simultaneously considering the graph structure of assets.

In this dissertation, we proposed the AI-based empirical asset pricing model to reflect the connected structure between assets in the factor model. We first proposed the graph neural network-based multi-factor asset pricing model. As important as the structure of the model in constructing an asset pricing model that reflects the structure of the connection between assets is, how to define the connectivity. Graph

neural network requires a well-defined graph structure. We defined the connectedness between assets as the binary converted Pearson correlation coefficients of asset returns by the cutoff value. The proposed model consists of a beta estimation part and a factor estimation part, where each part is estimated with firm characteristics and excess returns, respectively. The empirical analysis of U.S equities reveals that the proposed model has more explanatory power and prediction ability than benchmark models. In addition, the most efficient stochastic discount factor can be estimated from the estimated factors.

While return is the main object of asset pricing, volatility is also important property for explaining the behavior of financial assets. Volatility can be the factor in explaining return since many studies point out that return and volatility are correlated. As with the asset pricing model, considering the connected structure between assets in volatility prediction can be of great help in explaining the dynamics of assets. In the volatility analysis, what affects between volatility is called spillover. In this aspect, we proposed the volatility prediction model that can directly reflect this spillover effect. We estimated the graph structure between asset volatility using the volatility spillover index and utilized the spatial-temporal graph neural network structure for model construction. From the empirical analysis of global market indices, we confirm that the proposed model shows the best performance in short- and mid-term volatility forecasting.

To include volatility in the asset pricing discussion, it is necessary to focus on how volatility is defined in the asset pricing model. In the asset pricing model, volatility can be interpreted as the variance of the residual of the model. However, asset pricing models with time-series estimation mostly have time-unvarying volatility

constraints. We constructed an asset pricing model with time-varying volatility by estimating variability using the prediction model and reflecting it in the training loss of the asset pricing model. We identify that the proposed model can improve the statistical performance during the low volatility period through an empirical study of U.S equities.

Currently, there are clearly structurally connected assets in the cryptocurrency market, which has grown to a scale that cannot be ignored. All of the same blockchain-based tokens are issued and traded on that blockchain, so they have strong structural connectivity. We tried to identify that an observable factor for explaining excess return exists in such connected tokens as an application of previous studies. We limited the analysis target to Ethereum-based tokens and showed that the Ethereum gas price became a factor for the macroeconomic factor model after the application of EIP-1559. Furthermore, we applied the volatility spillover index-based volatility prediction model using gas return and showed that gas return can increase the prediction performance of certain tokens' volatility.

Keywords: Asset pricing, Volatility prediction, Graph neural network

Student Number: 2017-27701

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation of the Dissertation

Financial markets around the world have experienced several financial crises in the past century. There was great repression of the 1930s in the distance, and great repression of 2007-2010 from subprime mortgages can be found nearby. In particular, the Great Recession began with the bankruptcy of leading financial companies such as Bear Steans, Lehman, and AIG, causing a worldwide stock price plunge, and it took more than three years to recover. The fact that the stock market is exposed to risks easily like this suggests that we should be able to properly evaluate and manage the risk of stocks. Institutions and individuals must be able to correctly determine what the return on individual stocks is determined by and how the factors that determine the return relate to each other to form a risk-management portfolio.

In this aspect, explaining the behavior of financial assets is one of the main goals of financial asset management. Especially, the approaches that try to explain the behavior of return of financial assets is called asset pricing problem. This makes it possible to evaluate how much an asset is exposed to risk and how valuable the exposure is. In modern portfolio theory Markowitz (1952), Markowitz argued that firm risk is divided into systemic risk and idiosyncratic risk. The former is the risk

dependent on the market to which the firm belongs, and the latter is the risk that the firm has uniquely regardless of the market. Systematic risk can be removed by constructing a portfolio to gain a diversification effect. However, the idiosyncratic risk still remains after constructing the portfolio.

Various studies of asset pricing have been done to explain the idiosyncratic risk. The beginning of the asset pricing study can be seen as the capital asset pricing model (CAPM) proposed by Sharpe (1964) and Lintner (1965). CAPM claimed that the pricing of firm return is determined by beta coefficient and market excess return, where beta coefficient denotes the idiosyncratic risk. The beta coefficient is also known as risk exposure. This implies that the excess return of a firm is given as compensation for the risk of the firm.

Although CAPM proposed a method of structurally interpreting an excess return of a firm, studies to establish an empirically more accurate model continued. Among various literature, we concentrate on factor models. The factor model clarifies the factor that explains excess return. In the case of CAPM, it can be also seen as the factor model with one factor, the market excess return. Research trends on factor models can be classified into three categories: macroeconomic factor model, fundamental factor model, and latent factor model. The macroeconomic factor model uses macroeconomic variables as factors. CAPM is also part of the macroeconomic factor model because it uses a macroeconomic feature, market excess return. Chen et al. (1986) proposed a macroeconomic multi-factor model. They showed that unexpected changes in macroeconomic variables such as industrial production and inflation have an exogenous influence on asset return.

The most important factor in the construction of the macroeconomic factor

model is which macro variable to choose. It is important to select macroeconomic variables containing the unique characteristics of a market as factors for a macroeconomic factor model targeting assets in a certain market. For instance, oil prices play an important factor in the asset pricing of stocks in the oil refining industry sector, but oil prices may not have a significant impact on stocks in the banking sector that does not deal with consumer goods. Therefore, when evaluating an asset, it is possible to identify which connected network the asset belongs to, such as a sector in the case of stock, and to construct a macroeconomic factor model by selecting the macroeconomic variable corresponding to that network.

In recent years, the size of the digital asset market, including cryptocurrency, has grown so fast that the evaluation of digital assets has also emerged as one of the important research topics. Asset pricing models that use factors such as mining cost and reward have been developed to reflect the unique characteristics of cryptocurrency based on blockchain. However, these studies are attempts to judge cryptocurrency as a unified market without classifying it. In fact, multiple cryptocurrencies can be issued on a single chain, so it is reasonable to assume that each chain to which it belongs has network connectedness. Therefore, assets belonging to a single layer one blockchain, such as Ethereum-based tokens, can have a good explanatory power to construct a macroeconomic factor model using the unique properties of the network to which they belong as macroeconomic variables.

Unlike the macroeconomic factor model, which considers all firms to be affected by the same variable, the fundamental factor model is a model in which each firm considers its unique characteristic as a factor. Fama and French (1992) identified that the market excess return, the excess return of firms with smaller market capi-

talization versus large firms (Small minus big; SMB), and the spread between value stocks and growth stocks (High minus low; HML) are the three factors that explain the excess return of individual firms. While macroeconomic factor models use linear regression to estimate risk exposure to factors, Fama and French (1992) uses a two-step procedure consisting of cross-sectional analysis and time-series regression. First, they sort firms based on firm characteristics and construct long-short portfolio which takes a long position on the top quantile and takes a short position on the low quantile. The return of the long-short portfolio becomes factor. Second, the time-series regression is done based on the constructed factor and estimates risk exposure. Starting with the Fama-French three-factor model, there are parts where the model does not fully explain access return, and many studies have been proposed to present additional factors to explain this. Carhart (1997) proposed the four-factor model, which shows that the momentum factor explains the excess return that is not fully explained by the Fama-French three-factor model. Furthermore, Fama and French (2015) suggested additionally using robust minus weak (RMW) and conservative minus aggressive (CMA) factors to the Fama-French three-factor model and thus proposed the five-factor model.

The macroeconomic factor model and fundamental factor model both constitute factors with variables that are observable while the latent factor model aims to build an unobserved latent factor from the observed data. Unlike fundamental factor models, which usually have no more than six factors, latent factor models have the advantage of being able to utilize a larger number of observable variables to make them a smaller number of latent factors. In other words, for a latent factor model, projecting high-dimensional observable variables into a low-dimensional latent fac-

tor while maintaining as many variations as possible in the covariance matrix of the dataset is a problem to be solved. Connor and Korajczyk (1986) and Connor and Korajczyk (1988) developed asymptotic principal component analysis (APCA). APCA estimates n latent factors as the first n eigenvectors of the covariance matrix of asset return.

In recent years, there have been many related studies of latent factor model research that certain firm characteristics have the explanatory power to explain risk exposure, which can explain access return. Kelly et al. (2019) showed empirical evidence that asset characteristics (also called anomalies) can estimate the time-varying risk exposure. The asset pricing model, instrumented PCA (IPCA), proposed by Kelly et al. (2019) assumes that risk exposure beta can be estimated as the linear function of firm characteristics. However, many theoretical studies (e.g. Campbell and Cochrane (1999), Bansal and Yaron (2004)) claim that risk exposure can have a nonlinear complex structure of firm characteristics. To overcome the linear structure limitation, Gu et al. (2020a) introduced autoencoder formation that allows nonlinearity of risk exposure by using a nonlinear activation function in neural network structure. The authors of Gu et al. (2020a) generalized the PCA method often used for asset pricing by deep learning-based autoencoder because PCA is theoretically same with one layer autoencoder.

Even though Gu et al. (2020a) adopted a neural network for the asset pricing model, there exist only a few numbers of asset pricing studies based on deep learning. However, considering the characteristics of the latent factor model, the latent factor model has a curse of dimensionality problem because it deals with high-dimension features, and the neural network is an effective way to solve this problem (Bengio

et al. (2006), Poggio et al. (2017)), so further research needs to be done. As far as we know, Chen et al. (2019), Feng et al. (2020), and Gu et al. (2020a) are the only attempts to use neural networks as the estimation function for the asset pricing model. All of the above studies have a problem that the relationship between assets cannot be considered in the model since the propagation rule of basic neural network structure only supports column-wise or row-wise calculation. Since Ozsoylev and Walden (2011) and Herskovic (2018) have shown empirically that excess return of firms exchange relationships with each other, identifying the connection relationship between assets can help improve the performance of the asset pricing model.

Deep learning technology has developed various models that fit the characteristics of each domain, such as convolution neural network (CNN) in the image recognition field and recurrent neural network (RNN) in the NLP field. To handle data that each component has a relational structure, Kipf and Welling (2016) proposed the graph convolution network (GCN). GCN takes the adjacency matrix of graph-structured data as input and produces linked output considering the multi-step linkage between data components. While Kipf and Welling (2016) used citation data with a clear graph structure because the linkage between papers can be defined as quotation status, GCN can be applied to data with no pre-defined graphical structure. Cai et al. (2019) and Doosti et al. (2020) estimated the adjacency matrix of objects from the data and utilized it as the input of GCN. Therefore, the asset return data also does not have a clearly pre-defined relation, but GCN can be applied if the adjacency matrix is estimated in an appropriate way.

While asset pricing studies mostly concentrate on asset return itself, volatility is also the feature to be considered. Many researches as Jin (2017); Berument and

Doğan (2011) and Li (2011) showed that there exists a negative relationship between return and volatility. Therefore, volatility can be a useful factor that explains excess return. In this perspective, Herskovic et al. (2016) pointed out that shocks to the common idiosyncratic volatility high affect asset return. However, the limitation is that it uses historical volatility to construct factors. Considering the time-varying property of volatility, using predicted future volatility to explain future return can have more explanatory power. (Engle III and Ng (1991)).

Volatility prediction is being actively conducted in the field of time-series analysis. One of the biggest features of volatility research is that it should start with how to define volatility. Volatility is considered as the latent feature of asset return. The generalized autoregressive conditional heteroskedasticity (GARCH) model, proposed by Engle (1982), defined volatility as a hidden process embedded in the residual term of the autoregressive (AR) model for asset return. GARCH-based models as exponential GARCH and integrated GARCH predict daily volatility effectively (Nelson (1991); Engle and Bollerslev (1986). Numerous estimation methods for latent volatility have been proposed alongside GARCH volatility, as well as stochastic volatility (SV) and exponentially weighted moving average (EWMA) (see Taylor (2008) and Morgan (1996)).

However, Bollerslev (1987); Malmsten et al. (2010) and Carnero et al. (2004) pointed out that latent factors cannot capture decreasing autocorrelations in the squared returns, which is one of the important dynamics of return. Misunderstanding return dynamics can lead to the inadequate prediction of volatility. In this sense, Andersen and Bollerslev (1998) introduced the concept of observable realized volatility. While GARCH volatility is estimated by the AR model, the realized volatility is

directly calculated by the squared sum of returns. Theoretical results from Barndorff-Nielsen and Shephard (2002); Andersen et al. (2003) and Meddahi (2002) claim that observable volatility allows better prediction performance of volatility than latent volatility in high-frequency return data. To predict observable realized volatility (RV), Corsi (2009) proposed the heterogeneous autoregressive model for realized volatility (HAR-RV). HAR-RV has a formulation of AR model with daily lagged RV, weekly averaged lagged RV, and monthly averaged lagged RV as independent variables. Empirical prediction results of HAR-RV type models proposed by Chen and Ghysels (2011) and Patton and Sheppard (2015) showed good performance on global stock markets.

In volatility prediction, it is important to consider the connection relationship between assets as in asset pricing. The connection relationship between volatility is expressed as spillover, which indicates the effect of the impact on the volatility of other markets when an impact is applied to each market. Karolyi (2001) and Diebold and Yilmaz (2009) reported that there both exists a volatility spillover effect in individual stocks and global stock markets. Liang et al. (2020) and Wilms et al. (2021) showed that the volatility spillover effect also affects volatility prediction. Attempts to utilize the volatility spillover effect for realized volatility prediction mainly attempted to extend the HAR-RV model to a multivariate structure, as shown in Bubák et al. (2011) and Degiannakis et al. (2018). The limitation of these models is that they do not use well pre-defined volatility spillover. Although the multivariate model reflects the spillover effect in that it also uses the volatility of another asset when predicting the volatility of one asset, it does not reflect its exact degree. Diebold and Yilmaz (2009) and Diebold and Yilmaz (2012) measured the volatil-

ity spillover based on the error decomposition of the vector autoregressive model and defined the volatility spillover index. Therefore, employing well-defined volatility spillover index for volatility prediction have the potential to improve prediction performance.

The mixture of asset pricing and volatility prediction begins with considering the role of volatility in the asset pricing model. While traditional asset pricing models such as CAPM and Fama-French factor models assume constant volatility of model since they use linear regression for model parameter estimation, empirical evidence shows that volatility varies over time (see Justiniano and Primiceri (2008); Lee and Ohk (1992); Lewis (2021)). Kim and Kim (2016) proposed CAPM with a time-varying volatility framework to overcome this limitation, but it models volatility as latent volatility. Utilizing time-varying realized volatility for asset pricing model can guarantee better volatility prediction accuracy and thus more accurate asset pricing on return.

In this dissertation, we first focus on developing a deep learning-based asset pricing model that captures connectedness between assets. This study explores how the relationship between assets can be estimated and how the relationship can be reflected in the model structure. Next, we developed a volatility prediction method to explore return dynamics by considering the volatility spillover effect. Based on the previously proposed models, we proposed a deep learning-based asset pricing model with time-varying volatility prediction. Finally, we identified the macroeconomics factor for connected assets in the cryptocurrency market as an application.

1.2 Aims of the Dissertation

This thesis aims to develop an AI-based asset pricing model reflecting the network connectedness of assets. We first suggest a graph convolutional network-based asset pricing model that reflects the graphical relationship between assets. After developing the graph-based multi-factor asset pricing model, this thesis focuses on realized volatility prediction that contains volatility spillover effect. Then, we utilized the proposed methods to develop an asset pricing model with network connectedness and time-varying realized volatility. Finally, we showed that the connected assets in the cryptocurrency market have a common macroeconomic factor of gas fee as an application in the cryptocurrency market. The detailed summaries of this thesis are presented as follows:

Graph-based multi-factor asset pricing model (Chapter 2) In this chapter, we propose the graph-based multi-factor asset pricing model to make asset pricing model reflect the connectedness between assets. We estimate the network structure of assets by the Pearson correlation coefficients of asset returns and cutoff value for binary classification. Estimated adjacency matrix of assets is used as the input of proposed graph convolutional network-based asset pricing model. Subsequently, we propose the forward stagewise adaptive factor constructing algorithm for sequential factor modelling. We performed experiments on individual U.S. equities. The results demonstrate that our proposed method outperforms benchmark models in terms of explanatory power, prediction power, and Sharpe ratio of factor portfolio.

Volatility prediction with volatility spillover index (Chapter 3) Considering the volatility spillover effect in multivariate volatility forecasting is widely known to improve the prediction performance as it reflects the linkages between asset volatility. In this chapter, we propose a method that uses the volatility spillover index to construct a graph between global market indices and apply the graph directly through the spatial-temporal graph neural network model. An empirical analysis is conducted on eight representative global market indices. From the out-of-sample results, we found the following features. First, the proposed spatial-temporal GNN spillover model outperforms the benchmark models in short- and mid-term forecasting. Second, the forecasting accuracy highly depends on the inclusion of the market index with a high volatility spillover effect. Including SPX, which contains the highest net spillover index, effectively helps to forecast the volatility of other markets. Third, setting the mid-term KPPS step for constructing the graph performs the best for mid- and long-term forecasting tasks because the volatility spillover effect persists up to the mid-term.

Graph-based multi-factor model with time-varying volatility (Chapter 4) Allowing factor model to have time-varying volatility can improve the conformity to the real data of the model. In this chapter, we propose the graph-based multi-factor asset pricing model with the relaxation of the fixed volatility constraint. Realized volatility estimation is used for time-varying volatility prediction and worked as a regularization term of the training loss. The empirical analysis result on U.S individual stocks shows that the proposed model has a large increase in explanation and prediction power during the low volatility period. Furthermore, the proposed model

estimates factors that can construct the most efficient stochastic discount factor among benchmark models, which can be confirmed as the Sharpe ratio of tangency factor portfolios.

Macroeconomic factor model and spillover-based volatility prediction for ERC-20 tokens (Chapter 5) In this chapter, we extend the literature on identifying internal factors in cryptocurrencies by demonstrating that Ethereum-based token prices have a relationship with Ethereum gas prices. We applied the volatility prediction model with a volatility spillover index based on the identified relationship between tokens' volatility and gas return. Based on the relationship analysis, we constructed the macroeconomic two-factor model of market return and Ethereum gas return. An empirical analysis was performed using daily data from Ethereum, five ERC-20 tokens, and seven ERC-721 tokens. The results are shown in two periods divided by EIP-1559. The empirical results highlight the following features: First, the gas returns and the Ethereum returns and volatility are strongly correlated; Second, Ethereum and ERC-20 tokens' returns Granger-causes gas returns in the pre-EIP-1559 period, whereas gas returns have a causal effect on ERC-20 tokens' returns in the post-EIP-1559 period; Third, Ethereum volatility is the only asset volatility with predictive power for gas returns over the entire period; Fourth, the ERC-721 tokens did not show any constant pattern in their influence on the gas returns; and finally, the constructed two-factor model reveals that the EIP-1559 has made Ethereum gas price as the consistent factor for connected assets on Ethereum blockchain.

1.3 Organization of the Dissertation

The remainder of this dissertation is organized as follows. In chapter 2, we propose a graph-based multi-factor asset pricing model that reflects the graph structure of asset return. In chapter 3, we propose a volatility prediction model using a spatial-temporal graph neural network and volatility spillover index. In chapter 4, we propose a multi-factor model with time-varying volatility. Chapter 5 investigates the existence of observable factors for connected tokens in the cryptocurrency market as an application of research from previous chapters. Chapter 6 concludes the dissertation along with the contribution and future plan of the research.

Chapter 2

Graph-based multi-factor asset pricing model

2.1 Chapter Overview

"What determines the excess return of assets?" is a long-standing problem. Asset pricing models try to explain why different assets have different expected returns. Fama and French (2015) claim that firm characteristics such as market excess return, size and B/E ratio can explain expected asset returns proposing a three-factor model. Sanusi and Ahmad (2016) provide empirical evidence of a multi-factor model for stocks in specific sectors and research on the magnitude of each factor in the model was also conducted by Bank and Insam (2019). In addition to the study of observable factor models, study of latent factor models has also been conducted (Chamberlain and Rothschild (1982) and Connor and Korajczyk (1988)). Kelly et al. (2019) and Gu et al. (2020a) showed that firm characteristics affect risk exposure rather than risk factors.

Finding out the exact function to estimate risk exposure from firm characteristics is a difficult issue to solve using conventional asset pricing models, because observable firm characteristics have very high dimensions. Deep neural networks are known to perform well on similar tasks that deal with high-dimensional data (Goodfellow et al., 2016). Therefore, studies such as Gu et al. (2020a) and Chen et al. (2019) have tried

to adopt a deep learning approach to estimate excess returns from high-dimension asset characteristics.

However, deep-learning-based studies have thus far had two problems. First, the connectedness between firms is not reflected in the previously proposed models. Ozsoylev and Walden (2011) and Herskovic (2018) showed that the excess return of an individual firm is affected by those of other firms connected to it, and some asset pricing factors can be determined by the network structure of assets. However, recent deep-learning-based approaches, like those in Gu et al. (2020a) and Feng et al. (2020), have used an architecture in which different firms do not affect each other during layer-wise propagation. Second, as the number of factors increases, the influence of the added factors cannot be accurately measured, because the $(K+1)$-factor model does not inherit the factors of the K-factor model. Because existing deep-learning-based models start learning from a new initialization point each time, they cannot remember the factors from the previous model. Research on observable factor models has been conducted by adding a new factor to the existing model and checking whether it has explanatory power. For example, Carhart (1997) added the momentum factor to Fama and French (1992)'s three-factor model and determined whether the added factor explained the part of the data that the existing factors could not explain, as regards expected returns. In research on the statistical latent factor model, PCA-based models such as Lettau and Pelger (2020b) and Kelly et al. (2019) have performed empirical analysis by adding factors starting with a one-factor model to find the sufficient number of factors. This makes it difficult to compare an observable factor model or PCA-based latent factor models with a deep-learning-based latent factor model according to the number of factors on the same line.

In this chapter, we estimated the multi-factor asset pricing model of individual US equities using firm-specific information and deep learning architecture. Our contribution is two-fold. First, to reflect the connectedness between asset returns for risk exposure estimation, we used a graph convolutional network (GCN), which takes into account the graph structure between firms, rather than a simple neural network. We also presented a way to construct a graph of firms that fits our goal, asset pricing, using the correlation of returns. Second, we proposed a forward stagewise modeling architecture that sequentially adds latent factors to inherit factors from the latent factor model of the previous step, such as an observable factor model or PCA-based latent factor model. Moreover, we can accurately evaluate the value of a newly added factor as the increase in the model's explanatory power and prediction performance while increasing the number of factors because the $(K+1)$-factor model and the K-factor model share the same K factors.

Empirically, we conducted an analysis of 119 individual U.S. equities. In the out-of-sample analysis, our proposed graph factor model shows 29% explanatory power and 5.7% prediction R^2, and thus outperforms every other benchmark model. In terms of economic implications, we showed that our proposed model achieves the highest Sharpe ratio of the tangency factor portfolio.

This chapter is organized as follows: Section 2.2 describes background of our proposed model. In Section 2.3, we represent our proposed model. Section 2.4 presents the data for the empirical analysis, benchmark models, and the empirical results. Finally, Section 2.5 explains the corresponding discussion.

2.2 Preliminaries

2.2.1 Graph Neural Network

Lots of variations of artificial neural network, such as convolutional neural network and recurrent neural network, have been proposed for different purposes. However, most of those approaches are not specialized to deal with graph data, where nodes and their edges contain features and certain relationship information between nodes.

Proposed by Scarselli et al. (2008), the graph neural network (GNN) applies neural networks in graph data that is cyclic, directed, or undirected, such as social links and citation networks. Its framework attaches a state at each node based on its features, neighborhood information, and relationships with its neighbors. Node states x_n are calculated through function f_w with neighbor states and labels of nodes, edges, and neighbors, each denoted as $x_{ne[n]}$, l_n, $l_{co[n]}$, $l_{ne[n]}$.

$$x_n = f_w(l_n, l_{co[n]}, x_{ne[n]}, l_{ne[n]}) \tag{2.1}$$

Node representations are iteratively propagated by contraction maps until they reach a stable fixed point. Khamsi and Kirk (2011) supported the existence and uniqueness of convergence using Banach's fixed-point theorem.

Node states are used to compute their outputs, and the cost function can be set by the difference between their outputs and real outputs.

$$e_w = \sum_{i=1}^{p} \sum_{j=1}^{q} (t_{i,j} - \phi_w(G_i, n_{i,j}) \tag{2.2}$$

The gradient-descent algorithm is then applied to update the weights until the state

converges.

2.2.2 Graph Convolutional Network

GNN was later developed to apply the concepts of convolution in graph data by stacking multiple graph convolutional layers to extract node representations, which is known as GCN. GCN is based on the first order approximation of spectral graph convolution. Spectral convolutions on graph is defined as follows:

$$g_\theta \star x = U g_\theta U^T x \tag{2.3}$$

where g_θ denotes filter, U denotes eigenvector matrix of the normalized graph Laplacian, and x is the signal. Since g_θ can be approximated by a Chebyshev polynomials, Equation 2.3 can be written in K-th order Chebyshev polynomials formulation:

$$g_\theta \star x \approx \sum_{k=0}^{K} \theta'_k T_k(\tilde{L}) x \tag{2.4}$$

where T_k is k-th order Chebyshev polynomial, $\tilde{\Lambda} = \frac{2}{\lambda_{max}}\Lambda - I_N$, Λ is a diagonal matrix with eigenvalues of L, and λ_{max} is the largest eigenvalue of L. By assuming $\lambda_{max} \approx 2$ and $k = 1$, Equation 2.4 becomes as follows:

$$\begin{aligned} g_\theta \star x &\approx \theta'_0 x + \theta'_1 (L - I_N) x \\ &= \theta'_0 x - \theta'_1 D^{-\frac{1}{2}} A D^{-\frac{1}{2}} x \\ &= \theta (I_N + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}) x \end{aligned} \tag{2.5}$$

where $\theta = \theta'_0 - \theta'_1$.

Then, the layer-wise propagation rule of GCN for graph $\mathcal{G}$ with corresponding adjacency matrix A is derived from Equation 2.5 as follows:

$$\boldsymbol{H}^{(\ell+1)} = \boldsymbol{\phi}\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}\boldsymbol{H}^{(\ell)}\boldsymbol{W}^{(\ell)}\right) \tag{2.6}$$

Here, $\boldsymbol{H}^{(\ell)}$ and $\boldsymbol{W}^{(\ell)}$ are the matrix of hidden state and trainable weight in the ℓ^{th} layer, respectively. $\tilde{A} = A + I$ is the self-connection-added adjacency matrix, and $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$ is the normalization matrix of $\tilde{A}$. ϕ is an activation function, such as sigmoid, ReLU, or tanh.

2.3 Methodology

2.3.1 Multi-factor asset pricing model

It is well-known (Back (2010)) that under the no arbitrage assumption, a stochastic discount factor (SDF) satisfying the unconditional asset pricing equation exists:

$$\mathbb{E}_t[m_{t+1}R_{i,t+1}] = 0 \Leftrightarrow \mathbb{E}_t[R_{i,t+1}] = \frac{\text{cov}_t(m_{t+1}, R_{i,t+1})}{\text{var}_t(m_{t+1})}\left(-\frac{\text{var}_t(m_{t+1})}{\mathbb{E}_t[m_{t+1}]}\right), \tag{2.7}$$

where $R_{i,t+1}$ denotes the excess return of asset i at time $t+1$, orthogonal to the SDF, m_{t+1}. $\beta_{i,t} = \frac{\text{cov}_t(m_{t+1},R_{i,t+1})}{\text{var}_t(m_{t+1})}$ implies risk exposure, and $\lambda = -\frac{\text{var}_t(m_{t+1})}{\mathbb{E}_t[m_{t+1}]}$ is the price of risk.

Given random variable F, the orthogonal projection of $R_{i,t+1}$ on the span of F and constant is:

$$\mathbb{E}_t[R_{i,t+1}] + \beta_{i,t}^T(F_{t+1} - \mathbb{E}_t[F_{t+1}]) \tag{2.8}$$

where $\beta_{i,t} = \sum_{F_{t+1}}^{-1} Cov(F_t, R_{i,t+1})$. The existence of factor model with factor F is

equivalent as the existence of λ_{t+1} such that,

$$\mathbb{E}_t[R_{i,t+1}] = \lambda_{t+1}^T \beta_{i,t}. \tag{2.9}$$

We can find that the formulation of Equation 2.7 and Equation 2.9 derive the same result. If there is a factor model with factor F, the affine function of the factor F becomes an SDF. It leads to the following multi-factor model:

$$R_{i,t+1} = \beta_{i,t}^T F_{t+1} + \epsilon_{i,t+1}, \tag{2.10}$$

where $\beta_{i,t} \in \Re^{K\times 1}$, $F_{t+1} \in \Re^{K\times 1}$, $\epsilon_{i,t+1} \in \Re$ and $\mathbb{E}_t[\epsilon_{i,t+1}] = \mathbb{E}_t[\epsilon_{i,t+1}F_{t+1}] = 0$.

Recent literature, like Gu et al. (2020a); Feng et al. (2020) and Kelly et al. (2019), has tried to estimate β or F using firm characteristics Z. In this chapter, we follow this trend and use Z to estimate the risk exposure β.

$$R_{i,t+1} = \beta(Z_{i,t})' F_{t+1} \tag{2.11}$$

In this form, β becomes the embedding function of firm characteristics $Z_{i,t} \in \Re^{P\times 1}$.

We expand Eq. (2.11) to make function β conditioned by the characteristics of firms related to firm i.

$$R_{i,t+1} = \beta(Z_{j_1^i,t}, \ldots, Z_{j_n^i,t})' F_{t+1}, \tag{2.12}$$

where firms $j_1^i, \ldots, j_n^i$ are the firms that have a relationship with firm i. If we consider the relation of firms as graph $\mathcal{G}$ with adjacency matrix A, where nodes are firms and edges mean the connected nodes have a relationship, Eq. (2.12) can be expressed as

follows:

$$R_{t+1} = \boldsymbol{\beta}(\boldsymbol{Z}_t, A)F_{t+1}, \tag{2.13}$$

where $R_{t+1} = (R_{1,t+1}, ..., R_{N,t+1})^T \in \Re^{N\times 1}$, $\boldsymbol{\beta}(Z_t, A) \in \Re^{N\times K}$ whose i-th row vector is $\beta(Z_{j_1^i,t}, \ldots, Z_{j_n^i,t})'$.

2.3.2 Proposed method

The purpose of this chapter is to estimate the time-varying risk exposure from firm characteristics and connectivity between firms. To develop a deep-learning-based model for Eq. (2.12), we constructed a graph factor model (GF) consisting of two parts: a risk exposure model and a factor model similar to Gu et al. (2020a).

Figure 2.1 shows the architecture of our proposed model. We considered the GCN to estimate the risk exposure function β and set factors as a linear combination of excess returns using a single-layer neural network. The mathematical representation of the layer-wise risk exposure model is as follows:

$$\begin{aligned}
&\boldsymbol{H}^{(0)} = \boldsymbol{Z}_{t-1} \\
&\boldsymbol{H}^{(\ell)} = \boldsymbol{\phi}(\hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}\boldsymbol{H}^{(\ell-1)}\boldsymbol{W}^{(\ell-1)}), \quad l = 1, ..., L_\beta \\
&\boldsymbol{\beta}_{t-1} = \boldsymbol{H}^{(L_\beta)}
\end{aligned} \tag{2.14}$$

Notations here are the same as those in chapter 2.2.2: $\boldsymbol{\phi}$ can be any activation function. In this study, we used ReLU activation to afford non-linearity in the model. When $L_\beta = 2$, the model can be interpreted as a network that utilizes characteristics of other firms that are neighbors of firm i or are connected to a neighbor of firm i to calculate the β_i of firm i.

In a traditional graph convolutional network, A denotes the adjacency matrix of the graph. However, we do not know the exact connectivity between firms because there is no agreement to come up with a clear definition of connectivity. To address this issue, we estimated the adjacency matrix A using the correlation matrix of R.

$$A_{ij} = \begin{cases} 1, & \text{if } rank(|\rho_{ij}|) \geq c \\ 0, & \text{otherwise} \end{cases}, \tag{2.15}$$

where ρ_{ij} is the correlation of $R_{i,(0,...,t)}$ and $R_{j,(0,...,t)}$. $rk(x)$ denotes the percentile rank of x, and c is the cutoff value.

We used a simple neural network structure for the factor model to estimate factors as a linear combination of excess returns. The recursive formula is as follows:

$$\begin{aligned} &S^{(0)} = R_t \\ &S^{(\ell)} = b^{(\ell-1)} + \tilde{\boldsymbol{W}}^{(\ell-1)} S^{(\ell-1)}, \quad l = 1, ..., L_f, \\ &f_t = S^{(L_f)}, \end{aligned} \tag{2.16}$$

where $S^{(\ell)} \in \Re^{N\times 1}$, $\tilde{\boldsymbol{W}}^{(\ell)} \in \Re^{N\times K}$, and $b^{(\ell)} \in \Re^{K\times 1}$ are the hidden state vector, trainable weight matrix, and bias vector in the ℓ^{th} layer, respectively.

The output of Eq. (2.14) is the factor exposure matrix $\boldsymbol{\beta}_{t-1} \in \Re^{N\times K}$ and the output of Equation 2.16 is factor vector $f_t \in \Re^{K\times 1}$. Finally, we dot product $\boldsymbol{\beta}_{t-1}$ and f_t as in Eq. (2.11) to produce a prediction of asset return.

$$R_t = (\boldsymbol{H}^{(L_\beta)}) S^{(L_f)} = \boldsymbol{\beta}_{t-1} f_t \tag{2.17}$$

2.3.3 Forward stagewise additive factor modeling

We proposed a forward stagewise additive factor modeling architecture that maintains K factors from the K-factor model while constructing the $(K+1)$-factor model. Because our graph factor model is constructed of neural networks, we set the training initialization point of the $(K+1)$-factor model as the K-factor model that has undergone training. Then, we made corresponding gradients of K factors that were to be newly trained as not trainable, which makes the K factors stay the same when training the $(K+1)$-factor model. The entire training process can be seen in Algorithm 1.

Algorithm 1: Forward stagewise additive factor modeling

set K_{max};

while $K <= K_{max}$ **do**

  if $K = 1$ **then**

    train 1-factor model $R = \beta^{(1)} f^{(1)}$;

  else

    load $(K-1)$-factor model's factor $f^{(K-1)}$;

    initialize $f^{(K)}$'s 1 $(K-1)$ part with $f^{(K-1)}$;

    set $f^{(K)}$'s 1 $(K-1)$ part as requires_grad = False;

    train K-factor model $R = \beta^{(K)} f^{(K)}$;

  end

end

2.4 Empirical Studies

2.4.1 Data

We analyzed the monthly return data of firms from March 1957 to December 2016. Asset returns are collected from CRSP. The object of data collection is firms in NYSE, AMEX, and NASDAQ, which existed throughout the 60-year window, totaling 119 firms. Even though previous studies, Gu et al. (2020a); Kozak et al. (2017), and Gu et al. (2020b) used all available firms over the past 60 years, we have to select certain firms because our proposed model needs a fixed number of firms during the training, validation, and test periods. We used the three-month Treasury bill rate as the risk-free rate, which is also collected from CRSP. The average price of target assets are shown in Figure 2.2.

For firm characteristics, we used the dataset of Gu et al. (2020b). It includes 94 characteristics; the full list and details of characteristics are presented in Tables 2.1, 2.2 and 2.3. We rank-normalized each firm characteristics from (-1, 1) in each timestep.

We set data from 1957 to 1974 as the training set, the 1975 to 1986 data as the validation set, and the 1987-2016 data as the test set. The validation set is used to tune the hyperparameter of the model without observing any data in the test set.

2.4.2 Benchmark models

We chose the conditional autoencoder model of Gu et al. (2020a) for the benchmark of the latent factor model, and chose the Fama-French factor model for the benchmark of the observable factor model. The conditional autoencoder model has the same factor network structure as our proposed model, and uses a multi-layer per-

Table 2.1: List and reference of firm characteristics

Variable name	Firm characteristic	Reference
absacc	Absolute accuruals	Bandyopadhyay et al. (2010)
acc	Working capital accruals	Sloan (1996)
aeavol	Abnormal earning announcement volume	Lerman et al. (2007)
age	Number of years since first Compustat coverage	Jiang et al. (2005)
agr	Asset growth	Cooper et al. (2008)
baspread	Bid-ask spread	Amihud and Mendelson (1989)
beta	Beta	Fama and MacBeth (1973)
betasq	Beta squared	Fama and MacBeth (1973)
bm	Book-to-market	Rosenberg et al. (1985)
bm_ia	Industry-adjusted book-to-market	Asness et al. (2000)
cash	Cash holdings	Palazzo (2012)
cashdebt	Cash flow to debt	Ou and Penman (1989)
cashpr	Cash productivity	Chandrashekar et al. (2009)
cfp	Cash flow to price ratio	Desai et al. (2004)
cfp_ia	Industry-adjusted cash flow to price ratio	Asness et al. (2000)
chatoia	Industry-adjusted change in asset turnover	Soliman (2008)
chcsho	Change in shares outstanding	Pontiff and Woodgate (2008)
chempia	Industry-adjusted change in employees	Asness et al. (2000)
chinv	Change in inventory	Thomas and Zhang (2002)
chmom	Change in 6 month momentum	Gettleman and Marks (2006)
chpmia	Industry-adjusted change in profit margin	Soliman (2008)
chtx	Change in tax expense	Thomas and Zhang (2011)
cinvest	Corporate investment	Titman et al. (2004)
convind	Convertible debt indicator	Valta (2016)
currat	Current ratio	Ou and Penman (1989)
depr	Depreciation	Holthausen and Larcker (1992)
divi	Dividend initiation	Michaely et al. (1995)
divo	Dividend omission	Michaely et al. (1995)
dolvol	Dollar trading volume	Chordia et al. (2001)
dy	Dividend to price	Litzenberger and Ramaswamy (1982)
ear	Earnings announcement return	Brandt et al. (2008)

Table 2.2: List and reference of firm characteristics (Continued)

Variable name	Firm characteristic	Reference
egr	Growth in commom shareholder equity	Richardson et al. (2005)
ep	Earnings to price	Basu (1977)
gma	Gross profitability	Novy-Marx (2013)
grCAPX	Growth in capital expenditures	Anderson and Garcia-Feijoo (2006)
grltnoa	Growth in long term net operating assets	Fairfield et al. (2003)
herf	Industry sales concentration	Hou and Robinson (2006)
hire	Employee growth rate	Belo et al. (2014)
idiovol	Idiosyncratic return volatiltiy	Ali et al. (2003)
ill	Illiquidity	Amihud (2002)
indmom	Industry momentum	Moskowitz and Grinblatt (1999)
invest	Capital expenditures and inventory	Chen and Zhang (2010)
lev	Leverage	Bhandari (1988)
lgr	Growth in long-term debt	Richardson et al. (2005)
maxret	Maximum daily return	Bali et al. (2011)
mom12m	12 month momentum	Jegadeesh (1990)
mom1m	1 month momentum	Jegadeesh and Titman (1993)
mom36m	36 month momentum	Jegadeesh and Titman (1993)
mom6m	6 month momentum	Jegadeesh and Titman (1993)
ms	Financial statement score	Mohanram (2005)
mvel1	Size	Banz (1981)
mve_ia	Industry adjusted size	Asness et al. (2000)
nincr	Number of earnings increases	Barth et al. (1999)
operprof	Operating profitability	Fama and French (2015)
orgcap	Organizational capital	Eisfeldt and Papanikolaou (2013)
pchcapx_ia	Indstury adjusted percentage change in capital expenditures	Abarbanell and Bushee (1998)
pchcurrat	Percentage change in current ratio	Ou and Penman (1989)
pchdepr	Percentage change in depreciation	Holthausen and Larcker (1992)
pchgm_pchsale	Percentage change in gross margin - percentage change in sales	Abarbanell and Bushee (1998)
pchquick	Percentage change in quick ratio	Ou and Penman (1989)
pchsale_pchinvt	Percentage change in sales - percentage change in inventory	Abarbanell and Bushee (1998)
pchsale_pchrect	Percentage change in sales - percentage change in A/R	Abarbanell and Bushee (1998)

Table 2.3: List and reference of firm characteristics (Continued)

Variable name	Firm characteristic	Reference
pchsale_pchxsga	Percentage change in sales - percentage change in SG&A	Abarbanell and Bushee (1998)
pchsaleinv	Percentage change sales-to-inventory	Ou and Penman (1989)
pctacc	Percent accruals	Hafzalla et al. (2011)
pricedelay	Price delay	Hou and Moskowitz (2005)
ps	Financial statements score	Piotroski (2000)
quick	Quick ratio	Ou and Penman (1989)
rd	R&D increase	Eberhart et al. (2004)
rd_mve	R&D to market capitalization	Guo et al. (2006)
rd_sale	R&D to sales	Guo et al. (2006)
realestate	Real estate holdings	Tuzel (2010)
retvol	Return volatility	Ang et al. (2006)
roaq	Return on assets	Balakrishnan et al. (2010)
roavol	Earnings volatility	Francis et al. (2004)
roeq	Return on equity	Hou et al. (2015)
roic	Return on invested capital	Brown and Rowe (2007)
rsup	Revenue surprise	Kama (2009)
salecash	Sales to cash	Ou and Penman (1989)
saleinv	Sales to inventory	Ou and Penman (1989)
salerec	Sales to receivables	Ou and Penman (1989)
secured	Secured debt	Valta (2016)
securedind	Secured debt indicator	Valta (2016)
sgr	Sales growth	Lakonishok et al. (1994)
sin	Sin stocks	Hong and Kacperczyk (2009)
sp	Sales to price	Barbee Jr et al. (1996)
std_dolvol	Volatility of liquidity (in dollar trading volume)	Chordia et al. (2001)
std_turn	Volatility of liquidity (in share turnover)	Chordia et al. (2001)
stdacc	Accrual volatility	Bandyopadhyay et al. (2010)
stdcf	Cash flow volatility	Huang (2009)
tang	Debt capacity / firm tangibility	Almeida and Campello (2007)
tb	Tax income to book income	Lev and Nissim (2004)
turn	Share turnover	Datar et al. (1998)
zerotrade	Zero trading days	Liu (2006)

ceptron for a factor exposure network, while our model uses a graph convolutional network. For a fair comparison, we used the best-performing hyperparameters noted in Gu et al. (2020a).

The Fama-French factor model has varying factors from one to six; the six factors in turn are excess market returns, SMB, HML, CMA, RMW, and UMD. The Fama-French K-factor model we used for comparison in this study is composed of K factors described above. The factor returns are collected from Ken French's website.

We compared our graph factor model while changing the parameter c, which determines the sparsity of the asset graph's adjacency matrix as in Equation 2.15.

2.4.3 Empirical results

We performed the model performance evaluation through out-of-sample data to ascertain the true explanatory power and predictive power of our proposed model. Revisiting Euler equation 2.11, the asset pricing model is the explanation model for individual firms. We used R^2_{total} to evaluate statistical explanatory power.

$$R^2_{total} = 1 - \frac{\sum_{(i,t)\in OOS}(R_{i,t} - \hat{\beta}_{i,t-1}\hat{f}_t)^2}{\sum_{(i,t)\in OOS} R^2_{i,t}} \tag{2.18}$$

As deep-learning-based models are known to be good at prediction tasks, we evaluated the predicted performance by replacing $\hat{f}_t$ as $\hat{\beta}_{i,t-1} = \frac{1}{t-1}\sum_s \hat{\beta}_{i,s}$, which is the time series mean of $\hat{f}_{1:t-1}$.

$$R^2_{pred} = 1 - \frac{\sum_{(i,t)\in OOS}(R_{i,t} - \hat{\beta}_{i,t-1}\bar{f}_{1:t-1})^2}{\sum_{(i,t)\in OOS} R^2_{i,t}} \tag{2.19}$$

To achieve robust result and show confidence interval, we used bootstrap method

with 200 repeats. We computed the p-value of performance difference between graph factor model with cutoff value 0.9 and conditional autoencoder model by Welch's t-test using the empirical distribution from bootstrapping.

The results are shown in Table 2.4. In terms of the out-of-sample total R^2 and the prediction R^2, respectively, the graph factor model with a cutoff value of 0.9 and six factors showed the best performance, at 29% and 5.7%. The Fama-French factor model is known to explain characteristic-sorted portfolios, but shows little explanatory power for individual assets. The R^2_{total} of the conditional autoencoder model with four factors is 28.05%, which outperforms the graph factor model with cutoff values of 0.1 and 0.5. At the same time, p-value 0.75 shows that conditional autoencoder model with four factors has similar performance with graph factor model with cutoff value 0.9. However, in terms of R^2_{pred}, the conditional autoencoder model shows worse prediction performance than the graph factor model for all factor counts. Overall, the graph factor model with a cutoff value of 0.9 outperformed all other models. Table 2.5 shows additional metrics for evaluating goodness of fit and Table 2.6 shows significance of average coefficients of each model with six factors.

The cutoff value c of the graph factor model determines the sparsity of the firm graph's adjacency matrix. When $c = 0.9$, only 10% of firms are connected. We can see the visualization of the adjacency matrix with various cutoff values in Figure 2.3. Recalling that our proposed model estimates firm i's risk exposure by firm characteristics connected across up to two edges from firm i, $c = 0.1$, and $c = 0.5$, all firms' risk exposures are affected by nearly all other firms. The fact that the graph factor model with $c = 0.9$ dominates those with $c = 0.1$ and $c = 0.5$ in terms of R^2 from Figure 2.3 shows that the graph factor model requires a sparse, clear

Table 2.4: Comparison of Out-of-sample R^2_{total} and R^2_{pred}

	# of Factors	1	2	3	4	5	6
Out-of-sample total R-square (%)	FF	0.25	0.33	0.41	0.41	0.62	0.63
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
	CA	23.21	27.76	27.5	**28.05**	25.84	26.32
		(0.12)	(0.35)	(0.2)	(0.24)	(0.27)	(0.83)
	GF 0.1	21.36	22.48	22.9	24.28	24.72	25.42
		(0.12)	(0.08)	(0.31)	(0.19)	(0.19)	(0.17)
	GF 0.5	23.93	25.46	26.69	27.78	27.46	27.5
		(0.09)	(0.21)	(0.13)	(0.14)	(0.25)	(0.18)
	GF 0.9	**24.94**	**27.9**	**29**	28	**28.52**	**28.93**
		(0.1)	(0.15)	(0.19)	(0.11)	(0.31)	(0.11)
	t-statistic	20.4	0.76	8.02	-0.32	15.94	6.05
	p-value	0.00*	0.45	0.00*	0.75	0.00*	0.00*
Out-of-sample prediction R-square (%)	FF	<0	<0	<0	<0	<0	<0
	CA	1.13	2.72	2.98	3.28	2.63	2.56
		(0.13)	(0.31)	(0.14)	(0.2)	(0.21)	(0.44)
	GF 0.1	2.45	3.58	4.02	4.16	4.37	4.83
		(0.1)	(0.08)	(0.16)	(0.13)	(0.12)	(0.13)
	GF 0.5	3.14	4.23	4.6	4.54	4.74	5.33
		(0.11)	(0.13)	(0.08)	(0.09)	(0.16)	(0.13)
	GF 0.9	**3.79**	**4.98**	**5.21**	**5.58**	**5.64**	**5.7**
		(0.09)	(0.12)	(0.14)	(0.16)	(0.1)	(0.15)
	t-statistic	32.38	13.75	20.52	18.7	23.96	13.33
	p-value	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*

Notes. The upper part of table represents the empirical results of out-of-sample total R^2 and lower part represents the out-of-sample prediction R^2. 95% confidence intervals derived by bootstrapping are in parenthesis. In the second column, FF, CA, and GF each denotes Fama-French model, conditional autoencoder, and graph factor model. GF 0.1 means graph factor model with cutoff value 0.1. The best performing models along the fixed number of factors are shown in bold. For each metric, t-statistic and p-value are derived from comparison between CA model and GF 0.9 model. We marked * to p-value<0.05

adjacency matrix to be trained as intended.

Moreover, we can see that the R^2_{pred} of the graph factor model subsequently increases as the number of factors increases, while that of the conditional autoencoder model decreases when the number of factors increases from four to five. This is due to the proposed forward stagewise additive modeling scheme. Maintaining the previous stage model's factors produces more consistent predictive power in statistical

Table 2.5: Comparison of MAE, RMSE, RAE, and RSE

	# of Factors	1	2	3	4	5	6
Out-of-sample MAE	FF	0.074	0.074	0.074	0.074	0.074	0.074
	CA	0.056	**0.053**	0.054	**0.053**	0.055	0.054
	GF 0.1	0.057	0.057	0.057	0.056	0.055	0.054
	GF 0.5	0.055	0.054	0.054	**0.053**	0.053	0.053
	GF 0.9	**0.054**	**0.053**	**0.053**	**0.053**	**0.052**	**0.052**
Out-of-sample RMSE	FF	0.116	0.116	0.116	0.116	0.116	0.116
	CA	0.1	**0.096**	0.097	**0.096**	0.098	0.097
	GF 0.1	0.101	0.1	0.1	0.099	0.098	0.098
	GF 0.5	0.099	0.098	0.097	**0.096**	0.097	0.097
	GF 0.9	**0.098**	**0.096**	**0.096**	**0.096**	**0.096**	**0.096**
Out-of-sample RAE	FF	1.061	1.06	1.057	1.058	1.057	1.055
	CA	0.811	0.766	0.772	0.769	0.794	0.774
	GF 0.1	0.825	0.815	0.814	0.8	0.791	0.781
	GF 0.5	0.798	0.779	0.773	**0.759**	0.764	0.759
	GF 0.9	**0.787**	**0.76**	**0.758**	0.763	**0.753**	**0.751**
Out-of-sample RSE	FF	1.04	1.038	1.036	1.036	1.036	1.034
	CA	0.876	0.85	0.851	**0.848**	0.861	0.858
	GF 0.1	0.887	0.88	0.878	0.87	0.868	0.864
	GF 0.5	0.872	0.863	0.856	0.85	0.852	0.851
	GF 0.9	**0.866**	**0.849**	**0.843**	0.849	**0.845**	**0.843**

Notes. This table shows four metrics for evaluating goodness of fit. The first of the four sections in the table represents the out-of-sample Mean Absolute Error (MAE) $= \frac{1}{N}\sum|R_{i,t} - \hat{\beta}_{i,t-1}\hat{f}_t|$. The second, third, and fourth section each shows out-of-sample Root Mean Squared Error (RMSE) $= \sqrt{\frac{1}{N}\sum(R_{i,t} - \hat{\beta}_{i,t-1}\hat{f}_t)^2}$, Relative Absolute Error (RAE) $= \frac{\sum|R_{i,t} - \hat{\beta}_{i,t-1}\hat{f}_t|}{\sum|R_{i,t}|}$, and Relative Standard Error (RSE) $= \frac{\sum(R_{i,t} - \hat{\beta}_{i,t-1}\hat{f}_t)^2}{\sum R_{i,t}^2}$. For each metric, the best performing models along the fixed number of factors are shown in bold.

terms.

R^2 evaluation measures statistical evidence on asset pricing models. We can evaluate the economic meaning of the factor models through the Sharpe ratio test of the factor portfolio. According to Hansen and Jagannathan (1991), SDF is the linear span of excess returns. Since we estimated factors as a linear combination of excess returns in the graph factor model, SDF becomes the linear span of factors

Table 2.6: Significance of average coefficient of each factor

	t-statistics of average coefficient of					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
FF						
Mean	0.037	-0.073	-0.058	-0.108	0.374	0.016
Std error	0.014	0.019	0.031	0.043	0.044	0.015
t-statistic	2.696*	-3.89*	-1.866	-2.534*	8.55*	1.052
CA						
Mean	-0.004	0.008	0.006	0.023	0.052	0.1
Std error	0.001	0.001	0.005	0.01	0.017	0.009
t-statistic	-28.2*	116.6*	14.096*	24.71*	33.381*	122.244*
GF 0.1						
Mean	0.631	0.322	0.345	-1.991	-0.087	-0.349
Std error	0.069	0.025	0.096	0.149	0.355	0.038
t-statistic	9.122*	12.693*	3.593*	-13.337*	-0.243	-9.211*
GF 0.5						
Mean	-0.382	-0.075	0.057	-0.109	0.031	-0.007
Std error	0.012	0.017	0.009	0.015	0.004	0.006
t-statistic	-33.059*	-4.295*	6.507*	-7.189*	7.164*	-1.201
GF 0.9						
Mean	-0.118	-0.09	0.091	0.322	0.153	0.173
Std error	0.009	0.006	0.034	0.026	0.015	0.031
t-statistic	-13.705*	15.817*	2.632*	12.427*	10.163*	5.596*

Notes. This table shows mean, standard error and t-statistic of coefficients of FF, CA, and GF model with six factors. The mean coefficient is the cross-sectional average of $\bar{\beta}_i^{(k)}$, which is the time-series average of each firm's model coefficient $\beta_{i,t}^{(k)}$. In the same way, the standard error is calculated from $[\bar{\beta}_1^{(k)}, ..., \bar{\beta}_N^{(k)}]$ and the t-statistic is the mean coefficient divided by its standard error. t-statistics with its p-value smaller than 0.05 are marked as *

according to Kozak et al. (2017). Therefore, we construct the factor portfolio as the mean-variance efficient portfolio, considering each factor as asset returns, and analyzed the annualized Sharpe ratio. Table 2.7 shows the results. The Fama-French factors produce a maximum of 0.82, with six factors. The models with time-varying betas (CA and GF) showed much higher Sharpe ratios than the Fama-French model.

The best performing model is the graph factor model with factors more than four. Compared to the graph factor model, with a cutoff value of 0.9, the conditional autoencoder model produced a 1.74 Sharpe ratio. Through this analysis, we are able to confirm that the graph factor model produces factors that are closer to SDF than benchmark models.

Table 2.7: Comparison of tangency portfolio sharpe ratio

# of Factors	**1**	**2**	**3**	**4**	**5**	**6**
FF	0.51	0.41	0.53	0.71	0.71	0.82
CA	**0.52**	1.2	0.95	1.6	1.77	1.74
GF 0.1	0.47	1.62	1.84	1.86	1.86	1.89
GF 0.5	**0.52**	1.78	1.78	1.79	1.79	1.98
GF 0.9	**0.52**	**2.04**	**2.04**	**2.05**	**2.05**	**2.05**

Notes. Results in table denotes the annual Sharpe ratio of tangency portfolio of factors. The best performing models along the fixed number of factors are shown in bold.

2.5 Chapter Summary

In this chapter, we have proposed a new asset pricing model that estimates risk exposure by individual firm characteristics considering the connectivity between firms. We estimated the graph structure of assets using the correlation of asset returns and applied a graph convolutional network architecture using the estimated relationship between assets. To clearly evaluate the statistical and economic value of a new factor, we proposed a turn-based method of making a model while accumulating factors. Our empirical analysis is performed on monthly data for individual US equities. The results show that our graph factor model achieves 29% of the total R-squared and 5.7% of prediction R-squared, and thus that it dominates the Fama-French factor model and the conditional autoencoder model. The graph factor model has a large

cutoff value of 0.9, which means estimating the graph structure of assets as a sparse matrix outperformed others. This shows that considering connectivity between assets helps to estimate the risk exposure and helps explain excess returns. In terms of the economic meaning of the model, the graph factor model shows a 2.05 Sharpe ratio of the tangency factor portfolio, while the conditional autoencoder model shows a 1.74. Through these results, we can confirm that our graph factor model can generate stochastic discount factor by spanning model factors.

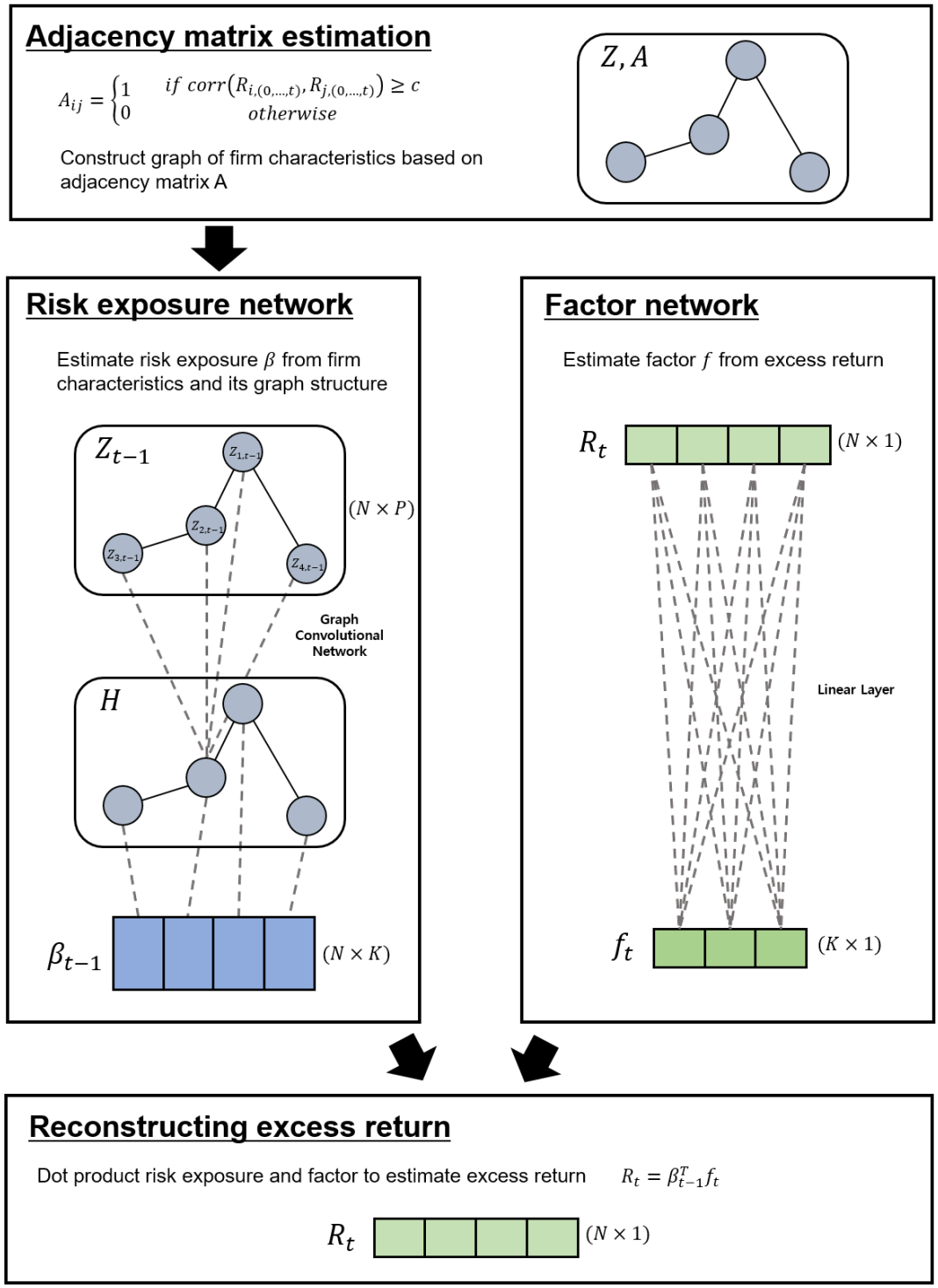

Figure 2.1: Overall architecture of graph based multi-factor model

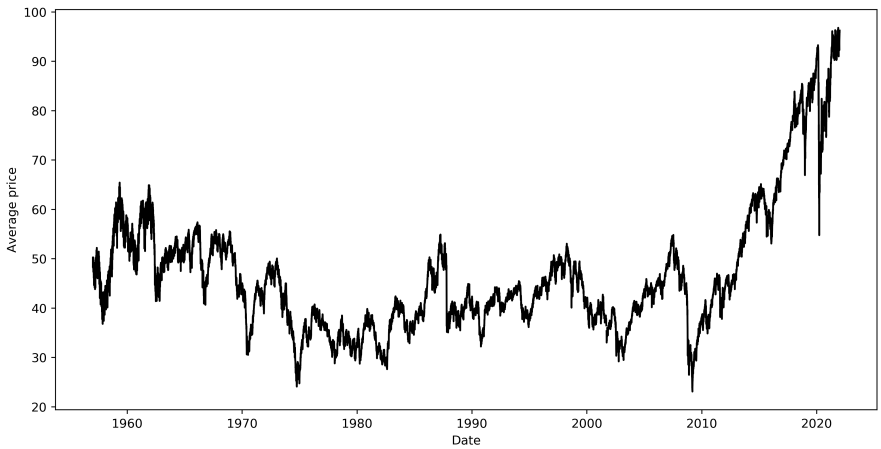

Figure 2.2: The average price of target assets during whole period

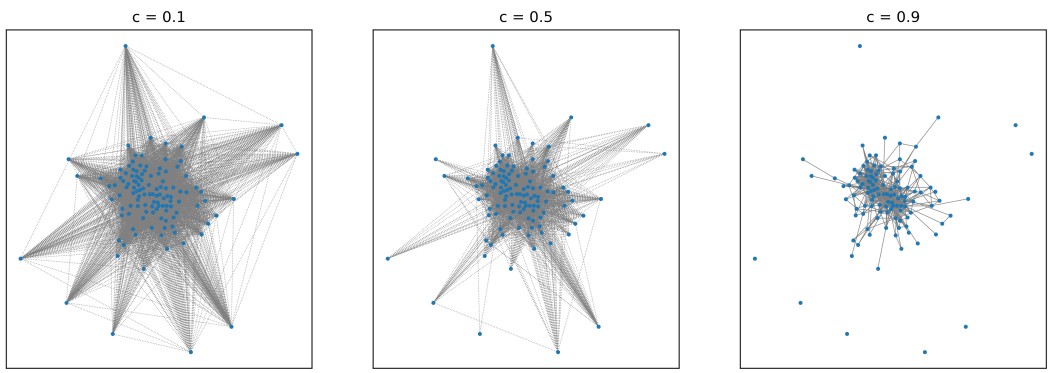

Figure 2.3: The illustration of adjacency matrix of firm characteristics as cutoff value differs

Chapter 3

Volatility prediction with volatility spillover index

3.1 Chapter Overview

Stock market volatility highly affects the decision-making of market participants. Individual investors and hedge funds particularly care about managing their portfolio weight based on the volatility of assets to manage the risk level of the portfolio (Engle (1993)). Volatility is also important for pricing options and derivatives, where stochastic volatility is used as the main component of pricing models (Ball and Roma (1994); Bouchaud and Potters (2003)).

Future volatility is more important than historical volatility. For example, the latter is used to find the maximum Sharpe ratio portfolio and efficient frontier for the portfolio optimization problem, but it cannot guarantee its efficiency in the future because it is not a sufficient estimate for future volatility (Choudhury et al. (2014)). Therefore, accurate volatility forecasting is one of the most important problems in managing risk.

The autoregressive conditional heteroskedasticity (ARCH) class of models was first widely used for forecasting volatility. Bollerslev (1986) introduced the generalized autoregressive conditional heteroskedasticity (GARCH) model, and Nelson (1991) proposed the exponential GARCH (EGARCH) model to solve some draw-

backs of GARCH as a negative correlation between current returns and future volatility. Engle and Bollerslev (1986) followed this research and introduced integrated GARCH (IGARCH). While GARCH class models show good performance for daily volatility, as shown in Harrison and Moore (2012), they do not perform well for high-frequency intraday volatility data. Andersen and Bollerslev (1998) introduced realized volatility (RV), which is constructed as the summation of squared intraday returns. RV is a highly efficient estimator of latent return volatility. For empirical forecasting of RV, Corsi (2009) proposed a heterogeneous autoregressive type of model (HAR-RV). Despite its simple structure, the HAR-RV model showed good empirical performance. Numerous studies have proposed extensions of the HAR-RV model to improve forecasting accuracy (e.g. Wang et al. (2017), Andersen et al. (2007), Chen and Ghysels (2011), Patton and Sheppard (2015)). In particular, our research focuses on Liang et al. (2020) and Wilms et al. (2021), who highlighted that volatility spillover between assets affects the prediction of realized volatility. Engle III et al. (1988); Baillie and Bollerslev (1991); Forbes and Rigobon (2002) and Bhar and Hamori (2003) suggested that when certain assets or markets experience shocks, the effect spreads to other assets or markets, which is known as the volatility spillover effect. Therefore, including the volatility spillover information in the forecasting model would increase the prediction performance.

BEKK-GARCH (Engle and Kroner (1995)) and HVS-GARCH (Wang et al. (2018)) represent GARCH-type models that successfully capture volatility spillover effects and generate accurate forecasts. They divided volatility from the return process into two components and constructed one component as a spillover component. However, they cannot be used for forecasting observed high-frequency real-

ized volatility because a GARCH-based structure is designed for unobserved latent volatility. To address this problem, various studies have proposed a multivariate HAR-based model that reflects spillover between different assets. Bubák et al. (2011) generalized multivariate HAR by allowing the error term to follow a multivariate GARCH process and analyzed volatility transmission between Central European currencies and the EUR/USD foreign exchange. While Bubák et al. (2011) concentrated on changing the model structure, Liang et al. (2020) and Wilms et al. (2021) utilized different predictive information in addition to the realized variance. Following Degiannakis et al. (2018), Liang et al. (2020) demonstrated that the information flow of implied volatility (IV) offers better performance in forecasting underlying volatility than RV. Garvey and Gallagher (2012) empirically shows that implied volatility is an effective metric of volatility across medium-term horizons. Wilms et al. (2021) extended the multivariate HAR model to allow jump, continuous, and IV information as components of volatility. These recent studies only capture the volatility spillover effect through a multivariate model structure.

In this chapter, we focus on implying the spillover effect directly in the forecasting model. Diebold and Yilmaz (2012) suggested a measure of the directional spillover index using a decomposition of error variance from a vector autoregressive regression (VAR) framework. As they successfully characterized the spillover effect across different market domains to the net pairwise spillover index, we designed a volatility forecasting model that involved volatility and the net pairwise spillover index as well. When we want to directly reflect the net pairwise spillover index in the model, HAR-based models can only use it as an additional variable in the regression equation, as in the HAR-RV-KS model of Liang et al. (2020). This causes the model

to consider only the spillover effect of each asset adjacent to the asset, and not the spillover effect that leads to multiple stages. As an extension of this discussion, we adopted a deep-learning-based model graph convolution network (GCN) to apply the spillover effect as a graph. The GCN structure can reflect the spillover effect of each market linked in multiple steps because these can be represented as multiple link connections between nodes in a graph.

Along with the success of deep-learning models in computer vision tasks, long short-term memory recurrent neural networks (LSTM-RNN) have gained popularity for time-series forecasting (Gers et al. (2000)) because of their ability to capture both long- and short-term memory in time-series data. Liu (2019) and Hu et al. (2020) suggested that LSTM-RNN can be used for volatility prediction of stocks and copper. On the other hand, deep-learning has also been developed to learn graphical relationships for relational data. Kipf and Welling (2016) proposed a GCN for a semi-supervised classification task on graph-structured data. Furthermore, Seo et al. (2018) and Li et al. (2017) designed a spatial-temporal GNN, the model combining RNN and GCN, to deal with multivariate time-series data where agents have physical or geological relationships. Chen et al. (2021a) and Chen et al. (2018) demonstrated that the graph convolutional approach can successfully capture the relationship between corporations, thus leading to better stock price or trend prediction. However, the major problem of applying the GCN approach to financial time-series data is constructing a graph. Unlike citation data or traffic data, financial volatility does not originally contain a graphical relationship. Therefore, defining and constructing graph relationships between markets has become the most important task. As a solution, we propose using the volatility spillover index as the graph structure of

global market indices.

In this chapter, we design a spatial-temporal GNN-based model for forecasting the volatility of global market indices. We propose a method for constructing graph edges of market indices with a net pairwise spillover index Diebold and Yilmaz (2012) and compare it to the method using Pearson correlation coefficients, which is a common approach. We conducted an empirical analysis of eight global market indices and present out-of-sample results. Out-of-sample analysis indicates that our proposed spatial-temporal GNN model using the spillover index outperforms other benchmark models in short- and mid-term forecasting. We also highlight that the results of our proposed model are highly affected by the inclusion of markets with large spillover effects on others.

The remainder of this chapter is organized as follows. In Section 3.2, we explain benchmark time-series models, deep-learning models, and volatility spillover measurements. Section 3.3 presents the data for the empirical analysis, the proposed prediction methodology, and the empirical results. Finally, concluding remarks are presented in Section 3.4.

3.2 Preliminaries

3.2.1 Realized Volatility

Andersen and Bollerslev (1998) introduced realized variance to approximate the integrated variance of intraday high-frequency return data . It is defined as the sum of intraday squared returns with equal time intervals. The mathematical formula for

the realized variance at day t is as follows:

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2 \tag{3.1}$$

where $\frac{1}{M}$ is the time interval and $r_{t,i}$ represents the return between $(i-1)M$ and iM. As discussed in Andersen et al. (2007), the standard deviation form of RV is closer to normal distribution. Therefore, we used $(RV_{i,t})^{1/2}$ as the realized volatility for the remainder of the paper.

3.2.2 Volatility Spillover Measurements

With the growing need to measure the financial market volatility spillovers, Diebold and Yilmaz (2009) introduced a general framework for measuring linkages between asset return volatilities, also known as volatility spillovers. Their framework is based on forecast error variance decomposition from VAR models, and it is capable of capturing spillover trends and cycles across individual assets, portfolios, markets, and so on. However, this framework has some limitations. First, the measuring of the volatility spillover is dependent on the order of the variables. Second, it only measures the *total* spillovers; therefore, one cannot examine the *directional* spillovers of a particular market. As a result, Diebold and Yilmaz (2012) extended the previous framework to measure the *directional* spillovers in a generalized VAR framework, which solves the variable ordering dependency problem (the DY framework hereafter).

A covariance stationary N-variable VAR(p) model can be defined as $x_t = \sum_{i=1}^{P} \Phi_i x_{t-i} + \epsilon_t$, where $\epsilon \sim (0, \Sigma)$ is an independently and identically distributed disturbance. The moving average of such a VAR(p) model can be expressed as $x_t = \sum_{i=0}^{\infty} A_i \epsilon_{t-i}$

where A_i represents the coefficient matrices that satisfy $A_i = \Phi_1 A_{i-1} + \Phi_2 A_{i-2} + \cdots + \Phi_p A_{i-p}$, with A_0 being an identity matrix and $A_i = 0$ for $i < 0$. The DY framework exploits the generalized VAR model by Koop et al. (1996) and Pesaran and Shin (1998) (KPPS) to solve the variable ordering dependency problem of the previous framework, as the KPPS produces variance decompositions that are invariant to the variable ordering. The KPPS H-step ahead forecast error variance decompositions $\theta_{ij}^g(H)$ are given as:

$$\theta_{ij}^g(H) = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma A_h' e_i)} \tag{3.2}$$

where σ_{jj} is the standard deviation of the error term for the j-th equation, e_i is the selection vector with one in the i-th element and zero otherwise, A_h is the coefficient matrix of the VAR model, and Σ is the variance matrix of the error ϵ. However, the KPPS approach does not orthogonalize the shocks of each variable, so the row sum of the elements of the variance decomposition may not be equal to one. Therefore, normalization by the row sum of the variance decomposition matrix is required:

$$\tilde{\theta}_{ij}^g(H) = \frac{\theta_{ij}^g(H)}{\sum_{j=1}^{N} \theta_{ij}^g(H)} \tag{3.3}$$

Then, $\sum_{j=1}^{N} \tilde{\theta}_{ij}(H) = 1$ and $\sum_{i,j=1}^{N} \tilde{\theta}_{ij}^g(H) = N$.

Using $\tilde{\theta}_{ij}^g(H)$, the *directional* volatility spillover by market i from all other mar-

kets j can be measured as:

$$S_{i\cdot}^{g}(H) = \frac{\sum_{\substack{j=1 \\ j\neq i}}^{N} \tilde{\theta}_{ij}^{g}(H)}{\sum_{i,j=1}^{N} \tilde{\theta}_{ij}^{g}(H)} \cdot 100 = \frac{\sum_{\substack{j=1 \\ j\neq i}}^{N} \tilde{\theta}_{ij}^{g}(H)}{N} \cdot 100 \tag{3.4}$$

In addition, the *directional* volatility spillover of market i to all other markets j can be measured as:

$$S_{\cdot i}^{g}(H) = \frac{\sum_{\substack{j=1 \\ j\neq i}}^{N} \tilde{\theta}_{ji}^{g}(H)}{\sum_{i,j=1}^{N} \tilde{\theta}_{ji}^{g}(H)} \cdot 100 = \frac{\sum_{\substack{j=1 \\ j\neq i}}^{N} \tilde{\theta}_{ji}^{g}(H)}{N} \cdot 100 \tag{3.5}$$

By netting the *directional* volatility spillovers related to market i, the *net* spillover of market i can be expressed as follows:

$$S_{i}^{g}(H) = S_{\cdot i}^{g}(H) - S_{i\cdot}^{g}(H) \tag{3.6}$$

Lastly, to examine the *net* volatility spillovers between individual markets, the *net pairwise* volatility spillovers is defined as:

$$\begin{aligned} S_{ij}^{g}(H) &= \left(\frac{\tilde{\theta}_{ji}^{g}(H)}{\sum_{i,k=1}^{N} \tilde{\theta}_{ik}^{g}(H)} - \frac{\tilde{\theta}_{ij}^{g}(H)}{\sum_{j,k=1}^{N} \tilde{\theta}_{jk}^{g}(H)} \right) \cdot 100 \\ &= \left(\frac{\tilde{\theta}_{ji}^{g}(H) - \tilde{\theta}_{ij}^{g}(H)}{N} \right) \cdot 100 \end{aligned} \tag{3.7}$$

3.2.3 Benchmark Models

Time-series Models

HAR-RV The heterogeneous autoregressive model of realized volatility (HAR-RV) proposed by Corsi (2009), is an autoregressive volatility model in which realized volatilities over different time horizons are considered. The HAR-RV model is a direct extension of the HARCH model by Müller et al. (1997), which assumes the existence of heterogeneity across market participants, also called the heterogeneous market hypothesis. The HAR-RV model finds the general pattern in the volatility structure within three different time interval sizes: daily (one day), weekly (5 days), and monthly (22 days). Daily, weekly, and monthly volatility reflects the behavior of short-, mid-, and long-term traders, respectively. Even though the HAR-RV model does not formally belong to the class of long-memory models, it is still capable of capturing the long-memory behavior of volatility in a very intuitive way. The original HAR-RV model can be expressed as

$$RV_{i,t} = \beta_{i,0} + \beta_{i,d} RV_{i,t-1} + \beta_{i,w} RV_{i,t-1:t-5} + \beta_{i,m} RV_{i,t-1:t-22} + \epsilon_{i,t} \tag{3.8}$$

where $RV_{i,t}$ is the realized volatility of stock market index i on day t, $RV_{i,t-1:t-n}$ is the simple average across the realized volatilities of stock market index i over time horizon $t-1$ to $t-n$. Therefore, $RV_{i,t-1:t-5}$, and $RV_{i,t-1:t-22}$ can be calculated as

follows:

$$RV_{i,t-1:t-5} = \frac{1}{5}\sum_{T=t-5}^{t-1} RV_{i,T} \tag{3.9}$$

$$RV_{i,t-1:t-22} = \frac{1}{22}\sum_{T=t-22}^{t-1} RV_{i,T} \tag{3.10}$$

However, Andersen et al. (2007) argue that non-linear HAR models in logarithmic forms and standard deviation show better fitting capabilities compared to linear HAR models. As a result, we modify the original HAR-RV model into the standard deviation form, similar to the model suggested in Liang et al. (2020). The standard deviation form of the HAR-RV model can be written as:

$$\begin{aligned}(RV_{i,t})^{1/2} = \beta_{i,0} + \beta_{i,d}(RV_{i,t-1})^{1/2} + \beta_{i,w}(RV_{i,t-1:t-5})^{1/2} \\ + \beta_{i,m}(RV_{i,t-1:t-22})^{1/2} + \epsilon_{i,t}\end{aligned} \tag{3.11}$$

HAR-RV-KS Liang et al. (2020) also consider the impact of global information flows of the realized volatility using a "kitchen sink" (KS) model, which incorporates all of the realized volatilities from other markets to the standard deviation form of the HAR-RV model. Therefore the HAR-RV-KS model can be written as follows:

$$\begin{aligned}(RV_{i,t})^{1/2} &= \beta_{i,0} + \beta_{i,d}(RV_{i,t-1})^{1/2} + \beta_{i,w}(RV_{i,t-1:t-5})^{1/2} \\ &+ \beta_{i,m}(RV_{i,t-1:t-22})^{1/2} + \sum_{k\in\{1,\dots,N\}\{i\}} \beta_{k,d}(RV_{k,t-1})^{1/2} + \epsilon_i, t \\ &= \beta_{i,0} + \sum_{\{k=1,\dots,N\}} \beta_{k,d}(RV_{k,t-1})^{1/2} + \beta_{i,w}(RV_{i,t-1:t-5})^{1/2} \\ &+ \beta_{i,m}(RV_{i,t-1:t-22})^{1/2} + \epsilon_{i,t}\end{aligned} \tag{3.12}$$

Deep-learning-based Models

RNN & LSTM Artificial neural network (ANN) is designed to resemble human brains and show good performance in generalizing and learning from experience. As mentioned in Kaastra and Boyd (1996), neural networks are data-driven self-adaptive methods that do not require many assumptions. For these properties, ANN have been proven very successful in many areas, including the finance domain, especially in forecasting volatile financial data such as interest rates and stocks.

Recurrent neural network (RNN), a variant of neural networks, was first proposed by Rumelhart et al. (1986) to effectively deal with time-and order-dependent data. For this purpose, RNN is used extensively in domains such as machine translation, speech recognition, or financial time series. RNN consists of a hidden state h with input sequence x and an optional output y. Given an input sequence $x = (x_1, x_2, ..., x_T)$, hidden state $h_t = (h_1, h_2, ..., h_T)$ is computed as follows:

$$h_t = \mathcal{H}(W_h h_{t-1} + W_x \sigma(x_t) + b) \tag{3.13}$$

where σ is an activation function and b is a bias. These hidden states are used to predict the outputs. The total error is defined by comparing the actual and desired outputs. To minimize the error, gradients are computed and back-propagated during the training phase. However, gradients of the RNN may quickly vanish or blow up, also known as the gradient vanishing problem or gradient exploding problem, respectively. To mitigate this problem, Hochreiter and Schmidhuber (1997) proposed LSTM using hidden layers containing memory blocks to store information and better exploit long-range context. In most of the proposed LSTM structures, a memory

block consists of three gates: input, output, and forget gates. Input gate controls the flow of input activations to filter the irrelevant inputs. Similarly, the output gate controls the output flow of cell activations to protect other units from perturbations. The forget gate, introduced by Gers et al. (2000), learns to reset memory contents that are no longer required.

Due to their ability to capture the structure of the dynamic data over time, LSTM have been used successfully in the prediction of stock returns and forecasting time-series events.

Spatial-Temporal Graph Neural Networks GNN have also been used in forecasting node values or labels given dynamic node inputs with interdependency. Among several GNN models proposed to handle dynamic input data, RNN-based approaches pass filter inputs and states to a recurrent unit using graph convolutions. This is known as spatial-temporal forecasting, and these approaches include graph convolutional recurrent network (GCRN) by Seo et al. (2018), and diffusion convlutional RNN (DCRNN) by Li et al. (2017). GCRN uses an LSTM network with CehbNet for graph convolutional layers. First introduced for traffic flow prediction, DCRNN attempts to capture the spatial and temporal dependency by using gated recurrent units (GRUs) and newly proposed diffusion convolutions. First proposed by Chung et al. (2014), GRU is a powerful variant of RNN, and it consists of a reset gate and an update gate. Similar to LSTM, GRU gates adaptively capture the dependencies of different time scales. In DCRNN, GRUs use diffusion convolutions instead of only matrix multiplications, so those GRUs are called diffusion

convolutional gated recurrent units. They are computed as follows:

$$r^{(t)} = \sigma(\Theta_r \star_{\mathcal{G}} [X^{(t)}, H^{(t-1)} + b_r)] \tag{3.14}$$

$$u^{(t)} = \sigma(\Theta_u \star_{\mathcal{G}} [X^{(t)}, H^{(t-1)} + b_u)] \tag{3.15}$$

where $\star_{\mathcal{G}}$ denotes the diffusion convolutions to aggregate the spatial dependencies. The diffusion convolution learns the graph representations using the graph node features and edge features that contain relationship data between nodes. The filter parameters of the reset and update gates are denoted as Θ_r and Θ_u, respectively, and these parameters are trained using back propagation during the training. The reset gates and update gates are then used to compute the current hidden states.

$$C^{(t)} = \tanh(\Theta_C \star_{\mathcal{G}} [X^{(t)}, (r^{(t)} \odot H^{(t-1)})] + b_c \tag{3.16}$$

$$H^{(t)} = u^{(t)} \odot H^{(t-1)} + (1 - u^{(t)}) \odot C^{(t)} \tag{3.17}$$

where $\odot$ and $C^{(t)}$ denote the pointwise operation and candidate hidden states, respectively. GRUs build diffusion convolutional recurrent layers, which are used to encode and decode the graph signals to make predictions. The structure of the GRU and the whole process of the DCRNN can be summarized as illustrated in Figure 3.1.

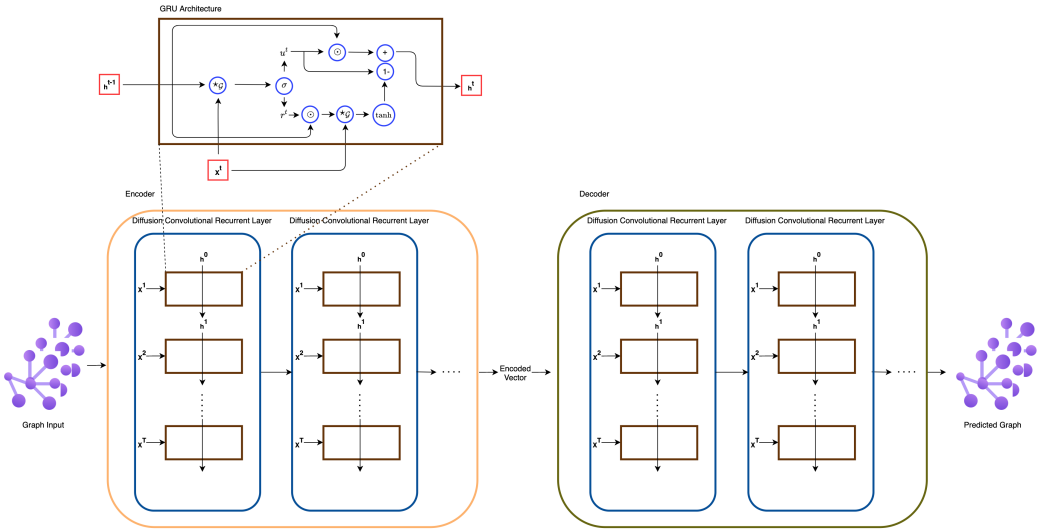

Figure 3.1: Diffusion Convolutional Gated Recurrent Unit and Diffusion Convolutional Recurrent Neural Network

3.3 Empirical Studies

3.3.1 Data

In this paper, we focus on predicting the RV of global market indices. We included both indices that are known to have a significant impact on other markets and those that do not. Overall, eight market indices were selected for analysis: S&P 500 (SPX; United States), DAX (GDAXI; Germany), CAC 40 (FCHI; France), FTSE 100 (FTSE; United Kingdom), OMX Stockholm All Share (OMXSPI; Sweden), Nikkei 225 (N225; Japan), KOSPI (KS11; South Korea), and HANG SENG (HSI; Hong Kong) index as in Liang et al. (2020). Realized variance data were collected from the Oxford-Man Institute's Quantitative Finance Realized Library.[1] The sampling frequency for calculating the realized variance is five minutes. Our sample begins in

[1] https://realized.oxford-man.ox.ac.uk/data

October 2006 and ends in December 2018. We divide the data into training (October 2006-November 2012), validation (November 2012-December 2014), and test samples (December 2014-December 2018) for proper training of our proposed model.

3.3.2 Descriptive Statistics

The descriptive statistics of the RV data are reported in Table 3.1. Augmented Dickey-Fuller (ADF) statistics Cheung and Lai (1995) indicate that the null hypothesis of a unit root can be rejected at the 5% significance level for every realized variance. Therefore, we can guarantee stationarity, which allows us to train the HAR-RV model without taking logs or other transformations of data. The time-series visualization of the data is illustrated in Figure 3.2. Shaded areas are the Global Financial Crisis and European Sovereign Debt Crisis. High volatile moments are mostly located during these two periods.

Table 3.1: Descriptive statistics of realized variance data

Index	Average	Standard deviation	Skewness	Kurtosis	ADF
SPX	0.855	0.687	3.336	18.162	-4.168***
GDAXI	1.011	0.614	3.379	20.629	-4.254***
FCHI	1.009	0.577	3.154	18.805	-4.732***
FTSE	0.941	0.615	3.862	31.846	-3.987***
OMXSPI	0.849	0.634	4.335	35.612	-3.661***
N225	0.856	0.529	3.332	17.326	-5.928***
KS11	0.766	0.529	4.281	31.112	-4.852***
HSI	0.866	0.516	3.546	21.31	-3.831***

Notes. For ADF statistics, the asterisks *** denotes p-values smaller than 1%.

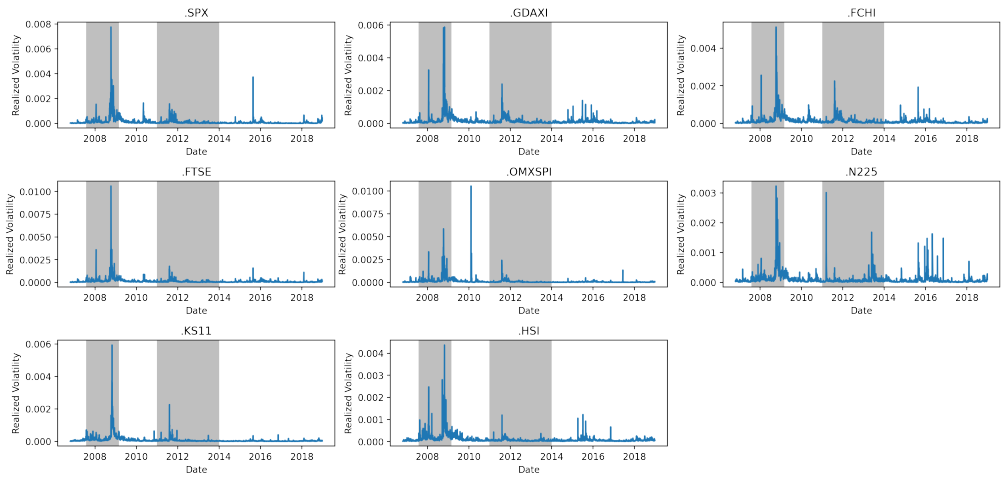

Figure 3.2: Time series of realized volatility data

3.3.3 Proposed Method

We propose a spatial-temporal GNN-based model that constructs graphs of global market indices with correlation or spillover index. Our model borrows the structure of DCRNN. We estimated the graph $\mathcal{G}$ from Equations 3.14 and Equation 3.16 as follows.

$$\mathcal{G}_{ij} = \begin{cases} f(RV_{i,(1:T)}, RV_{j,(1:T)}) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad \text{for} i, j \in 1, ..., N \tag{3.18}$$

where $\mathcal{G}_{ij}$ is the ij-th component of $\mathcal{G}$, and $RV_{i,(1:T)}$ is the length T vector $(RV_{i,1}, ..., RV_{i,T})$. T is the maximum time step in the training sample, and N is the total number of indices. $f(\cdot, \cdot)$ is the graph estimation function. The detailed structure of f is explained in Section 3.3.3. For the remainder of this paper, we refer to the spatial-temporal GNN model with graph estimation by the correlation method and volatility spillover index method as the STG-Correlation model and STG-Spillover model, respectively.

Graph construction

We constructed graphs between components using two types, correlation and spillover index.

Correlation method The correlation method calculates the Pearson correlation coefficients between realized volatility in the training period and uses it as a graph. As a result, the correlation graph is a symmetric matrix with ones on the diagonal, and it becomes an undirected graph. We defined the total correlation index TCI_i of each market i as the sum of the correlation coefficients r between the market and others.

$$TCI_i = \sum_{j \neq i} r_{ij} \tag{3.19}$$

Volatility spillover index method In addition, we constructed total and directed volatility spillover across global market indices Diebold and Yilmaz (2012). The spillover index calculation is described in Section 3.2.2. Volatility spillovers are derived through variance decomposition of errors from the vector autoregressive model. We used $p = 4$ and $H = 5$ for constructing VAR(p) and H-step ahead forecasts to derive error variances.

Since our purpose is to determine the impact between different market volatilities, we used the net pairwise spillover index for graph construction. In addition, the net spillover index is used to determine which market influences other markets and which market is affected by other markets. The constructed correlation graph and spillover graphs are presented in Figure 3.3. Each value on the edge is the Pearson correlation coefficient and the net pairwise spillover index, respectively. In

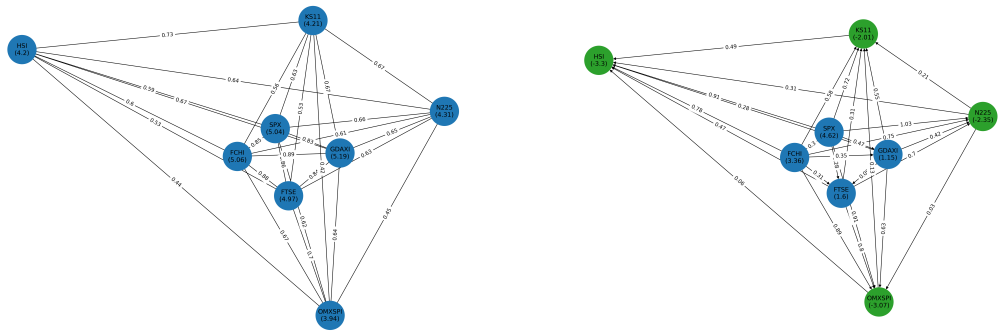

Figure 3.3: Visualization of constructed graphs

Notes. Left figure shows the graph constructed by correlation method and right figure is the one constructed by volatility spillover index method.

addition, the value inside each node represents the net correlation index and the net spillover index. GDAXI has the highest net correlation index of 5.19 and OMXSPI has the smallest net correlation index of 3.94. For the spillover graph, four market indices (SPX, FCHI, FTSE, and GDAXI) have a positive net spillover index, and four (KS11, OMXSPI, N225, and HSI) have a negative net spillover index. SPX has the highest net spillover index of 4.62, and HSI has the smallest value of -3.3.

3.3.4 Empirical Results

We conducted an empirical analysis of eight global market indices. The training dataset is used for training each model, and the validation dataset is used for early stopping and hyperparameter tuning. The optimal hyperparameters obtained through the validation set are as follows. We use number of filters as one, and set number of neurons in first and second layer as 128 and 64, respectively. The dropout rate of 0.5 is used to avoid overfitting and learning rate is 0.001. Finally, our analysis focuses on out-of-sample analysis using a test dataset. Overfitting a training dataset is one of the major problems in deep learning-based models. Therefore, out-of-sample

analysis results provide more meaningful and accurate information about the forecasting model than the in-sample analysis. During the out-of-sample analysis, we performed multi-step ahead forecasting. The forecasting steps are $h = 1, 5, 10, 22$. Each step represents short-term (a day), mid-term (one week and two weeks), and long-term forecasting (a month). We did not construct separate models for each forecasting task. Rather, we predicted h-step ahead forecast of realized volatility in time step t of market i, $(\hat{RV}_{i,t+h})^{1/2}$, by repeatedly using the one-step ahead forecasting model $f(\cdot)$.

$$\begin{aligned} (\hat{RV}_{i,t+h})^{1/2} &= f^h((RV_{i,t})^{1/2}) \\ \forall_{i,k}, (\hat{RV}_{i,k+1})^{1/2} &= f((\hat{RV}_{i,k})^{1/2}) \end{aligned} \tag{3.20}$$

The mean absolute forecast error (MAFE) is used for the evaluation metric. The h-step ahead forecasting MAFE for market i is defined as:

$$MAFE_i^{(h)} = \frac{1}{T-h+1} \sum_{t=0}^{T-h} |(\hat{RV}_{i,t+h})^{1/2} - (RV_{i,t+h})^{1/2}| \tag{3.21}$$

where T is the maximum time step in the test dataset.

This study confirms whether the spatial-temporal graph neural network-based model performs good out-of-sample forecasts on global market index volatility, and the spillover index is adequate for relation graph construction.

Out-of-sample results for multi-step ahead forecasting task for HAR-RV, HAR-RV-KS, RNN, STG-Correlation, and STG-Spillover are reported in Table 3.2. HAR-RV and HAR-RV-KS are used for the time-series based benchmark model, and RNN is used for the deep-learning-based benchmark model. The results show the MAFEs of each market index and the mean of MAFEs across the eight markets.

In the short- and mid-term horizons of $h = 1$, $h = 5$, and $h = 10$, the STG-

Spillover model outperforms other benchmark models with MAFE of 0.162, 0.194, and 0.211, respectively. In addition to the STG-Spillover model, the STG-Correlation model displays the best performance of MAFE 0.164 for short-term horizon $h = 1$. The HAR-RV model is the second-best model for the mid-term horizon $h = 5$ and $h = 10$. However, in the long-term horizon $h = 22$, the RNN model shows the best MAFE of 0.237, followed by the STG-Spillover model with 0.24. During all horizons, the STG-Spillover model always displays better results than the STG-Correlation model. This indicates that the net pairwise spillover index properly captures the relationship between market indices compared to the correlation coefficients.

Table 3.2: Out-of-sample results of five models on four different forecast horizons

Forecast horizon h	1					5				
	HAR-RV	HAR-RV-KS	RNN	STG-Correlation	STG-Spillover	HAR-RV	HAR-RV-KS	RNN	STG-Correlation	STG-Spillover
SPX	0.167	0.167	0.199	**0.166**	0.175	0.228	0.226	0.24	0.231	**0.225**
GDAXI	0.188	0.184	0.203	0.184	**0.181**	0.22	0.224	0.231	0.237	**0.212**
FCHI	0.171	0.176	0.192	0.171	**0.165**	0.208	0.215	0.223	0.227	**0.196**
FTSE	0.201	0.191	**0.187**	0.197	0.188	0.232	0.235	**0.217**	0.247	0.219
OMXSPI	0.157	0.176	0.159	0.133	**0.129**	0.176	0.2	0.179	0.182	**0.153**
N225	0.199	0.183	0.228	**0.198**	0.201	0.247	**0.232**	0.251	0.259	0.239
KS11	0.108	0.114	0.177	0.104	**0.103**	0.127	0.136	0.191	0.14	**0.123**
HSI	0.156	**0.146**	0.164	0.161	0.153	0.186	**0.179**	0.18	0.218	0.183
Mean	0.168	0.167	0.189	0.164	**0.162**	0.203	0.206	0.214	0.218	**0.194**

Forecast horizon h	10					22				
	HAR-RV	HAR-RV-KS	RNN	STG-Correlation	STG-Spillover	HAR-RV	HAR-RV-KS	RNN	STG-Correlation	STG-Spillover
SPX	**0.244**	0.245	0.257	0.264	0.248	0.28	0.278	**0.274**	0.349	0.29
GDAXI	0.241	0.241	0.241	0.277	**0.224**	0.256	0.255	0.259	0.379	**0.246**
FCHI	0.224	0.229	0.24	0.275	**0.211**	0.255	0.254	0.255	0.385	**0.244**
FTSE	0.246	0.242	**0.23**	0.298	0.234	0.277	0.266	**0.245**	0.407	0.268
OMXSPI	0.181	0.206	0.19	0.224	**0.164**	0.196	0.217	0.196	0.343	**0.187**
N225	0.277	**0.257**	0.261	0.312	0.268	0.318	0.298	**0.273**	0.424	0.305
KS11	0.139	0.146	0.192	0.182	**0.137**	**0.156**	0.162	0.191	0.286	0.162
HSI	0.205	0.196	**0.189**	0.275	0.203	0.223	0.222	**0.205**	0.394	0.221
Mean	0.220	0.220	0.225	0.263	**0.211**	0.245	0.244	**0.237**	0.371	0.240

Notes. This table reports out-of-sample MAFE of HAR-RV, HAR-RV-KS, RNN, STG-Correlation, and STG-Spillover model on forecast horizons 1, 5, 10, and 22. For each forecast horizon and market index, the best performing models are shown in bold.

To identify the statistical evidence of the performance difference between the STG-Spillover model and other benchmark models, we conducted the Diebold-Mariano (DM) test introduced by Diebold and Mariano (2002) and modified by

Table 3.3: DM test results of four competing models versus STG-Spillover model

Forecast horizon h	STG-Spillover vs HAR-RV		STG-Spillover vs HAR-RV-KS		STG-Spillover vs RNN		STG-Spillover vs STG-Correlation	
	DM statistic	p-value	DM statistic	p-value	DM statistic	p-value	DM statistic	p-value
1	-17.59	<0.01	-10.04	<0.01	-24.45	<0.01	-13.25	<0.01
5	-10.3	<0.01	-6.55	<0.01	-8.59	<0.01	-18.83	<0.01
10	-7.61	<0.01	-3.59	<0.01	-5.03	<0.01	-17.54	<0.01
22	-0.85	0.398	0.69	0.493	-0.75	0.451	-15.21	<0.01

Notes. This table reports out-of-sample DM statistic and p-value of four pairs models, which are STG-Spillover model versus HAR-RV, HAR-RV-KS, RNN, and STG-Correlation. A negative DM statistic means that STG-Spillover model performs better than the competing model.

Harvey et al. (1997). The DM statistic of two h-step ahead forecasts having errors $(e_{1,t}, e_{2,t}), t = 1, ..., T$ is expressed as follows:

$$DM = [Var(\bar{d})]^{-\frac{1}{2}} \bar{d} \tag{3.22}$$

where $d_t = g(e_{1,t}) - g(e_{2,t})$ while $g(e)$ is some kind of specified function and $\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t$. We selected the absolute function for $g(e)$, which makes $d_t = |e_{1,t}| - |e_{2,t}|$. The following p-value is obtained from the null hypothesis of $d = 0$ from student's t-distribution with $T - 1$ degrees.

The DM-test results of the STG-Spillover model versus others on forecast horizons of $h = 1, 5, 10, 22$ are reported in Table 3.3. The STG-Spillover model has statistically sufficient evidence at the 1% level that it outperforms STG-Correlation in terms of MAFE for every forecast horizon. For HAR-RV, HAR-RV-KS, and RNN, the DM test result confirms that the STG-Spillover model has an improvement over the three models in short- and mid-term forecasts. However, in the long-term, we can check that there is no statistical difference between the STG-Spillover model and the HAR-RV, HAR-RV-KS, and RNN models at the 5% level. Throughout the analysis, we can suggest that the STG-Spillover model successfully captures

the volatility spillover effect. Such effect measures the effect of a shock from a certain index, which decreases naturally over time. Therefore, the predominance of the STG-Spillover model decreases as the forecast horizon becomes longer.

Table 3.4: Out-of-sample results of comparing MAFE between STG-Spillover model trained on leave-one-out dataset and full dataset

Forecast horizon h		1			5			Net spillover index
		MAFE		Difference (%)	MAFE		Difference (%)	
		Leave-one-out	Full		Leave-one-out	Full		
Removed index	SPX	0.167	0.160	**4.38**	0.214	0.189	**13.06**	**4.62**
	GDAXI	0.158	0.159	-0.63	0.198	0.191	3.66	1.15
	FCHI	0.163	0.161	0.97	0.208	0.193	7.53	3.36
	FTSE	0.159	0.158	0.54	0.201	0.190	5.71	1.6
	OMXSPI	0.169	0.167	1.46	0.208	0.200	4.22	-3.07
	N225	0.156	0.156	-0.18	0.19	0.187	1.45	-2.35
	KS11	0.171	0.170	0.42	0.207	0.204	1.54	-2.01
	HSI	0.16	0.163	-1.93	0.192	0.195	-1.68	-3.3

Forecast horizon h		10			22			Net spillover index
		MAFE		Difference (%)	MAFE		Difference (%)	
		Leave-one-out	Full		Leave-one-out	Full		
Removed index	SPX	0.246	0.206	**19.50**	0.304	0.233	**30.31**	**4.62**
	GDAXI	0.222	0.209	6.22	0.266	0.240	10.83	1.15
	FCHI	0.234	0.211	10.83	0.282	0.240	17.57	3.36
	FTSE	0.228	0.208	9.69	0.28	0.236	18.43	1.6
	OMXSPI	0.231	0.218	6.03	0.271	0.248	9.27	-3.07
	N225	0.208	0.203	2.46	0.238	0.231	2.97	-2.35
	KS11	0.228	0.222	2.84	0.268	0.252	6.53	-2.01
	HSI	0.207	0.212	-2.49	0.233	0.243	-4.17	-3.3

Notes. This table reports out-of-sample MAFE of STG-Spillover model trained on the dataset that removed single market index and the model trained on the full dataset. The last column presents the net spillover index of each index that can be also seen in Figure 3.3. A negative difference (%) means that the STG-Spillover model trained on LOO dataset performs better. For each forecast horizon, the removed index with highest difference is shown in bold. In addition, the removed index with highest net spillover index is also shown in bold, too.

We conducted an experiment to determine the exact effect of the global market index spillover graph. Specifically, we created datasets $\mathcal{D}_{-i}$ excluding single market index i from the original dataset $\mathcal{D}$. A total of eight leave-one-out (LOO) datasets are created because we use a total of eight market indices. For each dataset $\mathcal{D}_{-i}$, we trained the STG-Spillover model and calculated the average MAFE loss

$MAFE_{-i}^{(h)}(\mathcal{D}_{-i})$ across seven indices included in dataset $\mathcal{D}_{-i}$. Next, we observed the difference between LOO MAFE loss $MAFE_{-i}^{(h)}(\mathcal{D}_{-i})$ and $MAFE_{-i}^{(h)}(\mathcal{D})$, which is the average MAFE of the same seven indices from the STG-Spillover model trained over the original dataset. The overall procedure is as follows:

$$\begin{aligned} MAFE_{-i}^{(h)}(\mathcal{D}_{-i}) &= \frac{1}{7}\sum_{j\neq i} MAFE_{j}^{(h)}(\mathcal{D}_{-i}) \\ MAFE_{-i}^{(h)}(\mathcal{D}) &= \frac{1}{7}\sum_{j\neq i} MAFE_{j}^{(h)}(\mathcal{D}) \end{aligned} \tag{3.23}$$

where $MAFE_{-i}^{(h)}(\mathcal{D}_{-i})$ denotes the average MAFE of indices included in $\mathcal{D}_{-i}$ and derived from the model using the training set $\mathcal{D}_{-i}$ and $MAFE_{-i}^{(h)}$ is the average MAFE across the same indices from the model trained by training set $\mathcal{D}$.

Out-of-sample results of single index removing analysis are reported in Table 3.4. For short-term forecast horizon, the three indices with the largest MAFE difference are SPX, OMXSPI, and FCHI. SPX, FCHI, and FTSE are the top-three indices with largest MAFE difference for mid- and long-term forecast horizons, which are the indices with a positive net spillover index. It is evidence that indices with high impact on other indices' volatility play an important role in spatial-temporal GNN spillover models. Specifically, SPX having the largest net spillover index of 6.32 mostly effects the prediction of other indices' volatility. Overall, the performance of the STG-Spillover model decreases significantly as the forecast horizon becomes longer. Furthermore, we highlight two significant observations. First, GDAXI and N225 removed model shows better average MAFE than the original model for $h = 1$ but performance worsens when the forecast horizon is over the mid- or long-term. The spillover effect becomes more important for predicting volatility when

the forecast horizon is longer than five days. Second, the HSI removed model is the only model that demonstrates better performance compared to the not removed model in all forecast horizon. HSI is the only market index in which every net pairwise spillover has a negative value. We could check that a market that does not contribute to volatility in other markets does not help predicting volatility of other markets through the STG-Spillover model.

Table 3.5: Out-of-sample results of STG-Spillover model trained on spillover dataset with different KPPS steps

Forecast horizon h	1				5				10				22			
KPPS step H	1	5	10	22	1	5	10	22	1	5	10	22	1	5	10	22
SPX	**0.167**	0.175	0.169	**0.173**	0.223	0.225	**0.221**	0.223	**0.242**	0.248	**0.24**	0.245	0.284	0.29	**0.278**	0.285
GDAXI	0.182	**0.181**	0.182	**0.181**	0.219	**0.212**	0.217	0.214	0.237	**0.224**	0.233	0.227	0.275	**0.246**	0.266	0.251
FCHI	0.166	**0.165**	0.167	0.166	0.203	**0.196**	0.201	0.198	0.223	**0.211**	0.22	0.214	0.273	**0.244**	0.265	0.25
FTSE	0.189	**0.188**	**0.189**	0.189	0.225	**0.219**	**0.222**	0.22	0.244	**0.234**	**0.241**	0.237	0.29	**0.268**	**0.284**	0.273
OMXSPI	**0.127**	0.129	**0.127**	0.128	0.156	**0.153**	**0.153**	**0.153**	0.165	0.164	**0.162**	**0.162**	0.197	0.187	0.188	**0.185**
N225	**0.198**	0.201	0.2	**0.2**	0.243	**0.239**	0.241	0.24	0.278	**0.268**	0.274	0.27	0.332	**0.305**	**0.323**	0.31
KS11	**0.1**	0.103	**0.1**	0.102	0.12	0.123	**0.118**	0.12	0.133	0.137	**0.13**	0.133	**0.158**	0.162	**0.15**	0.155
HSI	**0.153**	**0.153**	**0.153**	**0.153**	0.188	**0.183**	0.186	0.184	0.214	**0.203**	**0.21**	0.206	0.244	**0.221**	**0.235**	0.224
Mean	**0.160**	0.162	0.161	0.162	0.197	**0.194**	0.195	**0.194**	0.217	**0.211**	0.214	0.212	0.257	**0.240**	0.249	0.242

Notes. This table reports the out-of-sample MAFE of STG-Spillover models trained on the dataset constructed with different KPPS step H. For each forecast horizon and market index, the best performing models are shown in bold.

In addition, we show the relationship between H in KPPS H-step ahead forecast and the performance of the STG-Spillover model. H can be seen as the maximum time step for the effect of the shock to maintain. Therefore, the performance of the STG-Spillover model on short-, mid-, and long-term forecast horizons can be affected by setting H as short-, mid-, or long-term. The net pairwise spillover index and net spillover index calculated from each H are presented in Table 3.6. The out-of-sample results of STG-Spillover for different forecast horizons, trained by the dataset in which the following graph is constructed by the spillover index with different H for the KPPS H-step ahead VAR are presented in Table 3.5. In the short-term forecast horizon, the STG-Spillover model with KPPS step $H = 1$ showed the best

performance. In the mid- and long-term forecast horizons, the $H = 5$ model is one of the best performing, while the $H = 22$ model showed similar MAFE in the mid-term of $h = 5$. From the analysis, we can conclude that short KPPS step H helps to train the STG-Spillover model for predicting short-term ahead volatility, while long-term H does not contribute much to long-term volatility forecasting. Furthermore, we find that mid-term KPPS step H performs best when the forecast horizon is longer than five. It is evident that the influence of shock for one market to others mainly works until the mid-term.

3.4 Chapter Summary

We proposed a spatial-temporal GNN based model with a net pairwise spillover index graph. The out-of-sample results were analyzed on eight representative global market indices. We suggest several key findings regarding the direct application of the spillover effect in forecasting models. First, the STG-Spillover model demonstrates the best out-of-sample prediction performance in the short- and mid-term forecast horizons. Although the STG-Spillover model is not the best model for a long-term forecast horizon, it still outperforms the STG-Correlation model. This indicates that the net pairwise spillover index successfully reflects the volatility spillover effect in the spatial-temporal GNN model in comparison to the Pearson correlation coefficients. In addition, for the long-term forecast horizon, it was confirmed that the impact of the spillover effect decreased over time, and eventually made the performance gap between the RNN model and the STG-Spillover model almost non-existent. Second, the STG-Spillover model performs better when assets with high volatility spillover effects are included in the dataset. When SPX with the

largest net spillover is removed from the dataset, the short-term forecast accuracy decreases by 4.38% and the long-term forecast accuracy decreases by 30.31%. In contrast, when HSI, which has the smallest net spillover index and is only affected by other indices, is removed from the dataset, the overall performance increases. We suggest that this is because high-impact indices help predict the volatility of other indices. We can confirm that the STG-Spillover model captures the volatility spillover of each index well, and the model uses it to predict the volatility of other indices. Finally, we empirically show that setting a short KPPS H-step for constructing a net pairwise spillover index graph performs the best for short-term ahead forecasting tasks, while the mid-term KPPS step is the best for mid-term ahead forecasting. Interestingly, the long-term forecasts reveal distinct results: a graph using the mid-term not the long-term KPPS step is optimal. Thus, we conclude that the volatility spillover effect persists up to the mid-term horizon. Thus, constructing a graph of market indices with mid-term KPPS step H will provide the best STG-Spillover model fit for forecasting tasks with forecast horizon longer than five days.

In this chapter, we used a pre-defined volatility spillover measure in a spatial-temporal GNN-based model. Because constructing a proper graph that captures the relationship between input variables is an essential problem, some other variations of the GNN-based model try to estimate the graph itself using the trained model. In future research, measuring the spillover effect directly through the spatial-temporal GNN would be a topic worth considering, similar to the neural relational inference model introduced by Kipf et al. (2018).

Table 3.6: Net pairwise spillover index and net spillover index of global market indices with different KPPS steps

		Net pairwise spillover index								Net spillover index
Panel A : KPPS step H=1										
		To								
		SPX	GDAXI	FCHI	FTSE	OMXSPI	N225	KS11	HSI	
From	SPX	-	-	-	-	0.25	0.19	0.19	0.12	-0.3
	GDAXI	0.28	-	-	-	0.54	0.36	0.32	0.52	1.92
	FCHI	0.36	0.08	-	0.05	0.62	0.37	0.34	0.47	2.29
	FTSE	0.41	0.02	-	-	0.75	0.41	0.3	0.18	2.02
	OMXSPI	-	-	-	-	-	-	0.01	0.01	-2.22
	N225	-	-	-	-	0.08	-	0.34	0.27	0.64
	KS11	-	-	-	-	-	-	-	0.01	-1.49
	HSI	-	-	-	-	-	-	-	-	-1.58
Panel B : KPPS step H=5										
		To								
		SPX	GDAXI	FCHI	FTSE	OMXSPI	N225	KS11	HSI	
From	SPX	-	0.47	0.3	0.28	0.91	1.03	0.72	0.91	4.62
	GDAXI	-	-	-	0.09	0.63	0.42	0.55	0.28	1.15
	FCHI	-	0.35	-	0.31	0.89	0.75	0.58	0.78	3.36
	FTSE	-	-	-	-	0.8	0.7	0.31	0.47	1.6
	OMXSPI	-	-	-	-	-	-	0.13	0.06	-3.07
	N225	-	-	-	-	0.03	-	0.21	0.31	-2.35
	KS11	-	-	-	-	-	-	-	0.49	-2.01
	HSI	-	-	-	-	-	-	-	-	-3.3
Panel C : KPPS step H=10										
		To								
		SPX	GDAXI	FCHI	FTSE	OMXSPI	N225	KS11	HSI	
From	SPX	-	0.86	0.44	0.32	1	1.17	1.23	1.3	6.32
	GDAXI	-	-	-	-	0.54	0.33	0.31	0.22	-0.02
	FCHI	-	0.47	-	0.25	0.88	0.79	0.63	0.82	3.4
	FTSE	-	0.09	-	-	0.82	0.82	0.5	0.66	2.32
	OMXSPI	-	-	-	-	-	0.04	0.18	0.14	-2.88
	N225	-	-	-	-	-	-	-	0.33	-2.97
	KS11	-	-	-	-	-	0.15	-	0.86	-1.84
	HSI	-	-	-	-	-	-	-	-	-4.33
Panel D : KPPS step H=22										
		To								
		SPX	GDAXI	FCHI	FTSE	OMXSPI	N225	KS11	HSI	
From	SPX	-	1.07	0.53	0.34	1.01	1.31	1.51	1.5	7.27
	GDAXI	-	-	-	-	0.44	0.22	0.2	0.21	-0.76
	FCHI	-	0.52	-	0.19	0.85	0.82	0.82	0.91	3.58
	FTSE	-	0.24	-	-	0.83	0.94	0.86	0.9	3.24
	OMXSPI	-	-	-	-	-	0.12	0.34	0.25	-2.42
	N225	-	-	-	-	-	-	-	0.37	-3.29
	KS11	-	-	-	-	-	0.25	-	0.94	-2.54
	HSI	-	-	-	-	-	-	-	-	-5.08

Notes. This table reports the net pairwise spillover index and net spillover index. Four panels each show indices created with the parameter KPPS step $H = 1, 5, 10$, and 22. The number in the cell is the net pairwise spillover index 'from' the market index in row and 'to' the market index in column. For each panel, the last column represents net spillover index.

Chapter 4

Graph-based multi-factor model with time-varying volatility

4.1 Chapter overview

Asset pricing models focus on finding out the beta coefficient and factor that explains the excess return of underlying assets. The development of the asset pricing model over the last half century has many aspects, but we would like to view this as a relaxation of the time-varying property conditions of the coefficients. The first attempt at the asset pricing model, which can be seen as the capital asset pricing model (CAPM) proposed by Sharpe (1964) and Lintner (1975), tries to construct a factor as market return. In this case, they assumed that the beta coefficient does not vary over time. Bollerslev et al. (1988) pointed out that the relationship between market return and asset return may vary over time because of the restriction of the market and proposed the CAPM model with time-varying covariances, which typically denotes the time-varying beta coefficients.

Since most of the relevant researches adopted the idea of time-varying beta, only a few research had focused on the time-variation of the volatility. Unlike asset returns, which are directly observed and calculated from the rate of return derived from the price, volatility is an intrinsic concept and has a characteristic that cannot be directly observed. Therefore, the problem of discussing the volatility of the asset

must first begin with how to define the volatility. Referring to the GARCH process, the most representative method of estimating volatility, volatility can be thought of as the distribution of error terms of the model estimating return. The time-varying volatility concept of the model has empirically important meaning because the volatility of real data changes over time. Nelson (1991) found that a negative correlation exists between the current return and future volatility. Furthermore, Jin (2017) pointed out that the negative return-volatility relationship varies over time. The nature of this volatility is not reflected in linear regression-based models such as CAPM, Fama-French three, and five-factor models. Following the idea, Kim and Kim (2016) considered the framework of varying volatility based on the volatility of error term in CAPM. The main idea is that supposing the error term of CAPM is to follow a distribution with zero mean and time-varying volatility.

One of the major problems of relaxing the time-invariant volatility constraint for the traditional asset pricing model is that it makes the estimation hard because the terms to be estimated are added to beta, factor, and volatility. Kim and Kim (2016) employed the local-linear regression (Cleveland (1979); Stone (1977)). This methodology first estimates beta coefficients, and then obtains an estimated value of volatility based on the estimated beta. Similar research is also done by Ang et al. (2006). They measured the volatility of assets as the standard deviation of residuals from the Fama-French three-factor model.

All of these studies so far limit time-varying volatility to residuals within the asset pricing model. This may be seen as an advantage of non-parametric volatility prediction, but since volatility is not predicted based on the parametric model, there is a problem that the prediction performance itself may be degraded. This

study estimates volatility as realized volatility, estimates it as a separate parametric model from the asset pricing model, and suggests a method of reflecting it in the asset pricing model. We adopted neural network architecture for realized volatility prediction and asset pricing model. Deep-learning-based model architecture enables effective handling of feature space when the size of the feature space is large. Furthermore, compared to the linear regression model used in traditional asset pricing models, deep learning-based models can be used to estimate beta and factor by reflecting externally predicted volatility through modification of loss function. In addition, the neural network model solves the problem of increasing the number of parameters that the model must estimate by adding volatility as a variable, since it can always be learned through the stochastic gradient descent methodology if only the loss function is properly defined.

In this chapter, we propose the multi-factor asset pricing model with time-varying volatility prediction. The proposed model is constructed of two parts, the factor model part and the volatility prediction model part. The objective function of the proposed model becomes the mean squared error between true excess return and estimated excess return divided by estimated volatility plus the mean squared error between true volatility and estimated volatility. The volatility prediction model part uses LSTM as an estimation function and the asset pricing model part utilizes conditional autoencoder and graph neural network as estimation function.

The empirical analysis is done on monthly returns of 119 U.S individual stocks. From the out-of-sample results, we could suggest important findings. First, the increase in statistical performance of proposed time-varying volatility models during a low-vol period is more than twice as large as time-unvarying volatility models. Sec-

ond, the factors from five-factor graph-based factor model with time-varying volatility generate a tangency portfolio with the highest Sharpe ratio. It reveals that the graph-based factor model with time-varying volatility can estimate the most efficient stochastic discount factor.

The remainder of this chapter is organized as follows. Section 4.2 illustrates the background of time-varying volatility estimation. In Section 4.3, we present the structure and loss function of our proposed method. The data and benchmark models for empirical analysis, and empirical results are shown in Section 4.4. Concluding remarks are represented in Section 4.5.

4.2 Preliminaries

4.2.1 Local-linear regression for time-varying parameter estimation

Unlike the traditional time-unvarying regression model, the time-varying regression model can be expressed as the following general form:

$$y_\tau = x_\tau^T \beta_\tau + \sigma_\tau \epsilon_\tau \tag{4.1}$$

where $\tau = 1, 2, ..., T$ is the discrete time-step, β_τ is the coefficient and σ_τ denotes the time-varying volatility. ϵ_τ is the residual that satisfies $\mathbb{E}[\epsilon_\tau | x_\tau] = 0$ and $\mathbb{E}[\epsilon_\tau^2 | x_\tau] = 1$. To estimate β and σ as functions, β_τ and σ_τ can be treated as the smoothly-varying function by restricting domain in $[0, 1]$. The mathematical formulation can be expressed as follows:

$$\begin{aligned} \sigma_\tau &= \sigma(\tau/T), \tau = 1, ..., T \\ \beta_\tau &= \beta(\tau/T), \tau = 1, ..., T \end{aligned} \tag{4.2}$$

where $\sigma(\cdot) : [0, 1] \to \mathbb{R}^+$ and $\beta(\cdot) : [0, 1] \to \mathbb{R}^N$.

The estimation of parameters in Equation 4.1 requires estimation of both $\beta(\cdot)$

and $\sigma(\cdot)$. Therefore, local-linear regression is used (Cleveland (1979)). When $\beta(\cdot)$ has sufficient smoothness, $\beta(\cdot)$ and $\beta'(\cdot)$ can be local-linear estimated as:

$$(\hat{\beta}(t), \hat{\beta}'(t)) = \text{argmin}_{\mu_0,\mu_1} \sum_{\tau=1}^{T} (y_\tau - x_\tau^T \mu_0 - x_\tau^T \mu_1 (t - \tau/T))^2 K\left(\frac{t-\tau/T}{b_\tau}\right) \quad (4.3)$$

where $K(\cdot)$ is a kernel function with finite support, even symmetry, and positive value properties and b_τ is the bandwidth. After estimating $\beta(\cdot)$, $\sigma(\cdot)$ can be local-linear estimated as follows:

$$\hat{\sigma}^2(t) = \sum_{\tau=1}^{T} w_T(t,\tau)(y_\tau - x_\tau^T \hat{\beta}_\tau)^2 \quad (4.4)$$

where $w_T(t,\tau) = K\left(\frac{t-\tau/T}{h_T}\right) \frac{S_2(t)-(t-\tau/T)S_1(t)}{S_2(t)S_0(t)-S_1^2(t)}$ and $S_j(t) = \sum_{\tau=1}^{T} (t-\tau/T)^j K\left(\frac{t-\tau/T}{h_T}\right)$. h_T denotes bandwidth.

4.3 Methodology

4.3.1 Time-varying volatility implied loss function

In this section, we define the loss function for training the asset pricing model with time-varying volatility. Our proposed method is based on the multi-factor model formation of an unconditional asset pricing equation:

$$R_{i,t+1} = \beta_{i,t}^T F_{t+1} + \epsilon_{i,t+1} \quad (4.5)$$

where $R_{i,t+1}$ denotes the excess return of asset i in time horizon $t+1$, β is the beta coefficient as known as risk exposure, F_{t+1} is the factor, and $\epsilon_{i,t+1}$ denotes the residual. The residual $\epsilon_{i,t+1}$ satisfies the following condition.

$$\mathbb{E}_t[\epsilon_{i,t+1}] = \mathbb{E}_t[\epsilon_{i,t+1} F_{t+1}] = 0 \quad (4.6)$$

However, when estimating Equation 4.5, traditional approaches mostly use linear regression, which assumes constant variance for residual that $\mathbb{E}_t[\epsilon_{i,t+1}^2] = \sigma^2$. We relax the constant variance condition by allowing residual to have time-varying variance as follows:

$$\mathbb{E}_t[\epsilon_{i,t+1}^2] = \sigma_{i,t+1}^2 \tag{4.7}$$

where $\sigma_{i,t+1}$ becomes the conditional volatility of asset i in time horizon $t+1$. Then the multi-factor model can be written as the following equation.

$$R_{i,t+1} = \beta_{i,t}^T F_{t+1} + \sigma_{i,t+1}\epsilon'_{i,t+1} \tag{4.8}$$

where $\mathbb{E}_t[\epsilon'_{i,t+1}] = 0$ and $\mathbb{E}_t[{\epsilon'_{i,t+1}}^2] = 1$.

In chapter 2, we showed that when estimating Equation 4.5 with a deep neural network, the MSE loss $l = (R_{i,t+1} - \beta_{i,t}^T F_{t+1})^2$ can be used. Since the MSE loss minimizes the square of the residuals, it is the same target to be conceptually optimized as linear regression, except that it is minimized using the stochastic gradient descent methodology. However, the objective we need to minimize for Equation 4.8 is $\epsilon_{i,t+1}^2 = \sigma_{i,t+1}^2 {\epsilon'_{i,t+1}}^2$. Therefore, we define the modified MSE loss l^r to minimize the residual with time-varying volatility as follows:

$$l^r = \frac{1}{NT} \sum_{i,t} \frac{(R_{i,t+1} - \hat{\beta}_{i,t}^T \hat{F}_{t+1})^2}{\hat{\sigma}_{i,t+1}^2} \tag{4.9}$$

where $\hat{\beta}_{i,t}, \hat{F}_{t+1}$, and $\hat{\sigma}_{i,t+1}$ are the estimated values of $\beta_{i,t+1}, F_{t+1}$, and $\sigma_{i,t+1}$, respectively. N and T denote the number of assets and total time length. The l^r loss implies that when asset volatility is expected to be high, the asset pricing model has a more relaxed error criterion.

The problem of minimizing l^r loss is estimating $\sigma_{i,t+1}$. While $\beta_{i,t}$ and F_{t+1} can be

estimated by using a multi-factor model structure, $\sigma_{i,t+1}$ basically do not exists in the traditional multi-factor asset pricing model. We regarded this volatility prediction task as a separate task, and used it after estimating volatility from the outside of the asset pricing model through the separate model. The independent volatility prediction problem minimizes the l^v loss as follows:

$$l^v = \frac{1}{NT} \sum_{i,t} \left(1 - \frac{\hat{\sigma}_{i,t+1}}{\sigma_{i,t+1}}\right)^2 \tag{4.10}$$

The l^v loss has the structure of ratio formulation of MSE loss. It aims to make l^v loss has the same scale as the l^r loss.

By integrating the l^r loss and the l^v loss, we finally define the l^t which will be used for training the model.

$$\begin{aligned} l^t &= w_1 l^r + w_2 l^t \\ &= w_1 \frac{1}{NT} \sum_{i,t} \frac{(R_{i,t+1} - \hat{\beta}_{i,t}^T \hat{F}_{t+1})^2}{\hat{\sigma}_{i,t+1}^2} + w_2 \frac{1}{NT} \sum_{i,t} \left(1 - \frac{\hat{\sigma}_{i,t+1}}{\sigma_{i,t+1}}\right)^2, \end{aligned} \tag{4.11}$$

where w_1, w_2 are weight parameters for l^r and l^t, respectively.

4.3.2 Proposed model architecture

The time-varying volatility asset pricing model has three parameters to estimate, which are $\beta_{i,t}, F_{t+1}$, and $\sigma_{i,t+1}$. When estimating $\beta_{i,t}$ and F_{t+1}, we follow the multi-factor asset pricing approach used in chapter 2. $\beta_{i,t}$ is estimated from firm characteristics $Z_{i,t}$, and F_{t+1} is estimated by the portfolio of future returns, which is the linear span of R_{t+1}. We refer the estimation functions for $\beta_{i,t}$ and F_{t+1} as $g(\cdot)$ and $h(\cdot)$. Then the mathematical formulation of multi-factor model becomes as follows:

$$R_{i,t+1} = g(Z_{i,t})h(R_{t+1}) + \epsilon_{i,t+1} \tag{4.12}$$

For estimating future volatility, we use the time-series approach. $\sigma_{i,t+1}$ is estimated based on the historical volatility $\sigma_{i,1:t}$. We refer to the estimation function as $f(\cdot)$. Using $g(\cdot), h(\cdot)$, and $f(\cdot)$ notation, the time-varying volatility asset pricing model can be written as follows:

$$R_{i,t+1} = g(Z_{i,t})h(R_{t+1}) + f(\epsilon_{i,1:t})\epsilon'_{i,t+1} \tag{4.13}$$

Estimation of beta and factor

We used two approaches to estimate $g(\cdot)$ and $h(\cdot)$. The first approach is the conditional autoencoder model proposed by Gu et al. (2020a). It estimates $g(\cdot)$ with multi-layer perceptron with input $Z_{i,t}$ and estimates $h(\cdot)$ with a single layer network with input R_{t+1} to make F_{t+1} as the linear combination of future return.

The second approach is the graph-based multi-factor model we proposed in chapter 2. The graph-based structure aims to construct a graph between assets to capture the relationship between assets and adopt it in the model. In this chapter, we used Pearson correlation coefficients of asset returns as the proxy for the adjacency matrix of assets. The cutoff value of Pearson correlation coefficients to construct binary adjacency matrix is set as 0.1, which is the best parameter found in chapter 2.

Estimation of volatility

We used the LSTM approach to estimate $f(\cdot)$ for volatility prediction. It is well known that RNN-based approaches work well for time-series forecasting task. We also showed that it works well in long-term volatility forecasting in chapter 3. Although the objective of the volatility prediction task in this chapter is one-step ahead forecast, the used realized volatility is monthly realized volatility. Therefore, in term of the forecast horizon, the task can be assumed as long-term volatility forecasting.

4.4 Empirical Studies

4.4.1 Data

We used the monthly return data of 119 firms on NYSE, AMEX, and NASDAQ from March 1957 to December 2021. The targeted firms are those that have full data during the above period. The 94 firm characteristics are from Gu et al. (2020a), which is the same dataset used in chapter 2. The list of firm characteristics is represented in Tables 2.1, 2.2, and 2.3. The monthly return data and three-month Treasury bill rate are collected from CRSP. We divided the whole period into training, validation, and test set. The training set is from 1957 to 1974, the validation set is from 1975 to 1986, and the test set is from 1987 to 2021.

4.4.2 Benchmark Models

For a fair comparison, we selected four models as benchmark models. As the benchmark for the fundamental factor model, we selected the Fama-French factor model (Fama and French (1992, 2015)). The K-factor Fama-French factor model consists of the first K elements of (Excess market return, SMB, HML, RMW, CMA). The varying volatility CAPM is used as the benchmark for the varying volatility macroeconomic factor model with the local-linear regression introduced in chapter 4.2.1. Since the CAPM is one-factor model, varying volatility CAPM is only applied in one-factor model.

As the benchmark for latent time-unvarying volatility factor models, the conditional autoencoder model and graph-based multi-factor model is used. Each model is trained using the best parameters as mentioned in its originally proposed papers.

During the rest of this chapter, FF stands for Fama-French factor model, LLR

stands for varying volatility CAPM with local-linear regression estimation, CA-UV stands for conditional autoencoder with time-unvarying volatility model, GF-UV stands for the graph-based multi-factor model with time-unvarying volatility model, CA-VV stands for conditional autoencoder with time-varying volatility model, and GF-VV stands for the graph-based multi-factor model with time-varying volatility model.

4.4.3 Empirical Results

During the out-of-sample period, we compared the statistical performance of benchmark models and proposed models. Statistical performance includes both explanation power and prediction power of excess return. Asset pricing models are basically explanation models because the objective variables and independent variables are in the same time step. Therefore, explanation power comparison is the proper evaluation index for model comparison. Although the asset pricing model takes the form of an explanatory model, it is also possible to change it to the form of a predictive model through a time-series mean. Out-of-sample R^2_{total} and R^2_{pred} each are used for explanatory and prediction performance indicators:

$$\begin{aligned} R^2_{total} &= 1 - \frac{\sum_{(i,t)\in OOS}(R_{i,t} - \hat{\beta}_{i,t-1}\hat{f}_t)^2}{\sum_{(i,t)\in OOS} R^2_{i,t}} \\ R^2_{pred} &= 1 - \frac{\sum_{(i,t)\in OOS}(R_{i,t} - \hat{\beta}_{i,t-1}\bar{f}_{1:t-1})^2}{\sum_{(i,t)\in OOS} R^2_{i,t}} \end{aligned} \tag{4.14}$$

Table 4.1 shows the empirical results. CA-VV model shows the best out-of-sample total R-square in the one-factor model, while the GF-UV model shows the best out-of-sample total R-square in the two-, three-, and five-factor model. The GB-VV four-factor model represents the highest out-of-sample total R-square of 32.28%. For the

out-of-sample prediction R-square, the GF-UV model shows the best performance in the two-, three-, and the five-factor model. The GF-VV model represents the best performance in the one- and the four-factor model, where the 7.26% of the four-factor model is the highest R-square. We can confirm that the graph-based factor models, both unvarying volatility and varying volatility models, show better performance than time-series models and conditional autoencoder models in terms of prediction power.

Revisiting Equation 4.11, the proposed training loss gets smaller when the predicted volatility decreases. It implies that the CA-VV and GF-VV model will be trained more tightly in the low-vol period. To check the quantitative effect based on volatility level, we compared the out-of-sample total R-square and prediction R-square during the low-vol period and high-vol period. The low-vol and high-vol period are defined by the time-series quantile of the average volatility of target assets. We used five-quantile and the k-th quantile set is defined as follows:

$$\begin{aligned} S &= \{\bar{\sigma}_1, ..., \bar{\sigma}_T\}, \text{where } \bar{\sigma}_t = \frac{1}{N}\sum_{i=1}^{N}\sigma_{i,t} \\ U_k &= \{\bar{\sigma}_t | \bar{\sigma}_t \leq x_k, Pr[S \leq x_k] \leq \frac{k}{5}\} \end{aligned} \tag{4.15}$$

Tables 4.2 and 4.3 each report the out-of-sample total R-square and prediction R-square for U_1 and U_4. For comparison of performance difference between the whole period and U_1 period, Figure 4.1 shows the difference in percentage for CA-UV, GF-UV, CA-VV, and GF-VV models. The results show that both total R-square and prediction R-square for varying volatility models, CA-VV and GF-VV, increases far more than unvarying volatility models in the low-vol period. In the low-vol period of the four models, the average increase in total R-square for each factor is 1.22%,

1.86%, 6.84%, and 7.19%, respectively. The average increase in prediction R-square for each factor is 5.04%, 3.23%, 11.19%, and 9.54%, respectively. It shows that in the low-vol period, the R-square of varying volatility models increases more than double compared to unvarying volatility models. Looking at the absolute amount, not the increase in performance, the model that produces the best performance for each number of factors is mostly similar to the result in the total period, but is different in the two-factor model. For the two-factor models in the low-vol period, the CA-VV model shows the best performance in terms of out-of-sample total R-square. During the high-vol period, out-of-sample total R-square and prediction R-square decreases for every model. However, the rank of each metric between models does not change compared to the whole period.

Alongside R-square metrics, the Sharpe ratio of factor tangency portfolio is also used. While out-of-sample total R-square and prediction R-square evaluate the statistical performance of models, the Sharpe ratio of factor tangency portfolio is the measure of the economic value of constructed factors. Following Hansen and Jagannathan (1991), the stochastic discount factor should be well estimated by the linear span of factors. Therefore, the optimal mean-variance portfolio of well-constructed factors should replicate the stochastic discount factor. Since the higher out-of-sample Sharpe ratio of stochastic discount factor infers a more efficient stochastic discount factor, the Sharpe ratio test of factor portfolios becomes the test of efficient stochastic discount factor construction, which is the essential goal of asset pricing. Table 4.4 represents the results for the Sharpe ratio of tangency factor portfolios of each model. For the two- and three-factor models, the GF-UV model shows the highest Sharpe ratio of 2.07. GF-VV model shows the highest Sharpe ratio in one-, four-

Table 4.1: Comparison of Out-of-sample R^2_{total} and R^2_{pred}

	# of Factors	1	2	3	4	5
Out-of-sample total R-square (%)	FF	0.25	0.32	0.4	0.42	0.62
	LLR	2.08	-	-	-	-
	CA-UV	23.25	27.69	27.48	28.08	25.88
	GF-UV	24.97	**27.75**	**28.91**	28.21	**28.6**
	CA-VV	**27.01**	27.12	26.33	26.31	24.38
	GF-VV	26.06	25.33	27.35	**32.28**	24.61
Out-of-sample prediction R-square (%)	FF	<0	<0	<0	<0	<0
	LLR	<0	-	-	-	-
	CA-UV	1.17	2.67	2.97	3.3	2.64
	GF-UV	3.75	**4.97**	**5.15**	5.62	**5.59**
	CA-VV	2.01	2.69	2.55	3.08	2.67
	GF-VV	**4.24**	4.18	3.25	**7.26**	5.1

Notes. The upper part of table represents the empirical results of out-of-sample total R^2 and lower part represents the out-of-sample prediction R^2. In the second column, FF, LLR, CA, and GF each denotes Fama-French model, time-varying volatility CAPM with local-linear regression estimation, conditional autoencoder, and graph factor model. UV and VV each denotes the time-unvarying volatility and time-varying volatility. The best performing models along the fixed number of factors are shown in bold.

, and five-factor models. It is the similar result with the out-of-sample prediction R-square result, except for the five-factor model.

From the empirical analysis of statistical performance and economic valuation of benchmark models and proposed models, we could confirm the following points. First, graph-based factor models show the best out-of-sample prediction R-square regardless of the number of factors. Both the time-unvarying volatility GF model and the time-varying volatility GF model show better statistical performance than CA models. Second, time-varying volatility models have more than doubled the statistical performance in the low-vol period compared to time-unvarying volatility models. Finally, the Sharpe ratio of tangency factor portfolio shows that graph-based models construct efficient factors and the factors constructed by the five-factor GF-VV model can span the most efficient stochastic discount factor along all models.

Table 4.2: Comparison of Out-of-sample R^2_{total} and R^2_{pred} during low 20% quantile volatility period

	# of Factors	1	2	3	4	5
Out-of-sample total R-square (%)	FF	0.26	0.35	0.41	0.42	0.65
	LLR	2.2	-	-	-	-
	CA-UV	23.46	28	27.86	28.64	26.03
	GF-UV	25.42	28.58	**29.32**	28.44	**29.17**
	CA-VV	**28.69**	**28.86**	28.33	27.42	26.76
	GF-VV	28.06	27.48	28.04	**36.21**	25.86
Out-of-sample prediction R-square (%)	FF	<0	<0	<0	<0	<0
	LLR	<0	-	-	-	-
	CA-UV	1.21	2.86	3.13	3.39	2.75
	GF-UV	3.87	**5.14**	**5.45**	5.77	**5.8**
	CA-VV	2.54	2.87	2.76	3.3	2.87
	GF-VV	**4.76**	4.55	3.69	**7.71**	5.45

Notes. This table shows the empirical results of out-of-sample total R^2 and prediction R^2 during the period that average observed realized volatility is in low 20% among out-of-sample periods. The upper part of table represents the out-of-sample total R^2 and lower part represents the out-of-sample prediction R^2. The best performing models along the fixed number of factors are shown in bold.

Table 4.3: Comparison of Out-of-sample R^2_{total} and R^2_{pred} during high 20% quantile volatility period

	# of Factors	1	2	3	4	5
Out-of-sample total R-square (%)	FF	0.23	0.32	0.39	0.4	0.57
	LLR	2.01	-	-	-	-
	CA-UV	23.05	27.51	27.29	27.96	25.64
	GF-UV	24.79	**27.61**	**28.84**	27.9	**28.32**
	CA-VV	**26.8**	26.9	26.08	26.17	24.07
	GF-VV	25.8	25.06	27.26	**31.78**	24.45
Out-of-sample prediction R-square (%)	FF	<0	<0	<0	<0	<0
	LLR	<0	-	-	-	-
	CA-UV	1.01	2.44	2.78	3.84	2.41
	GF-UV	3.71	**4.73**	**5.82**	5.08	**5.31**
	CA-VV	1.82	2.34	2.32	2.84	2.49
	GF-VV	**3.97**	3.94	2.9	**6.75**	4.82

Notes. This table shows the empirical results of out-of-sample total R^2 and prediction R^2 during the period that average observed realized volatility is in high 20% among out-of-sample periods. The upper part of table represents the out-of-sample total R^2 and lower part represents the out-of-sample prediction R^2. The best performing models along the fixed number of factors are shown in bold.

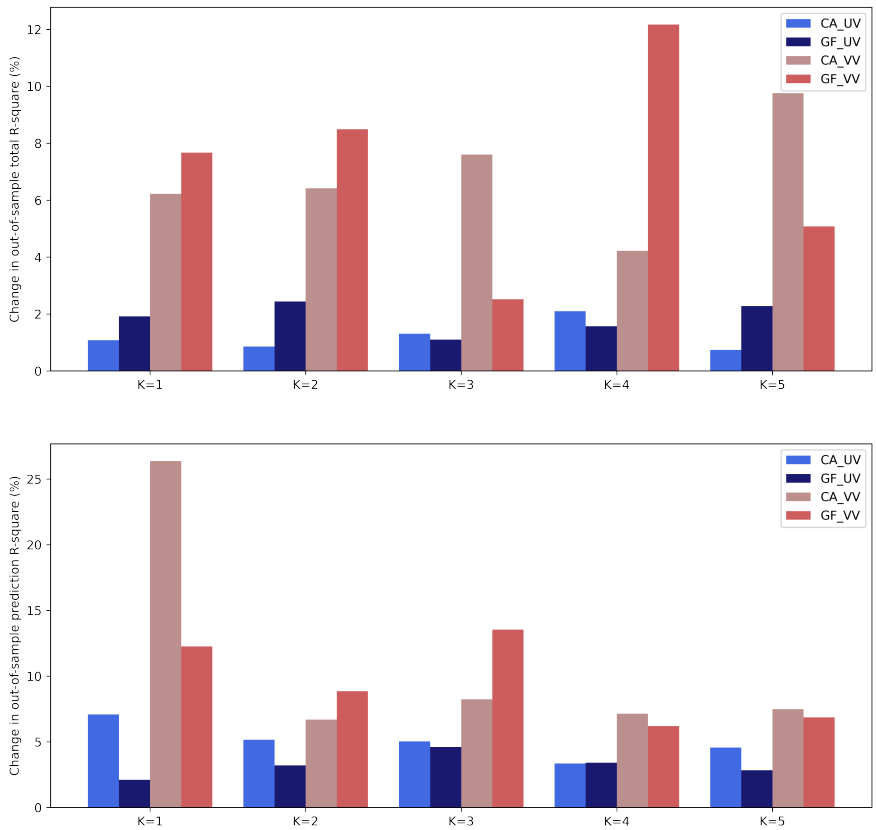

Figure 4.1: Change in R-square between total period and low-vol period

In addition, we have done an empirical analysis on the period from January 2017 to December 2021 to figure out the robustness of the proposed model. Since the global markets have been highly volatile after COVID-19, which has been started in 2021, restricting the test period to recent years can give value for the check of efficiency of proposed model in recent stock market. Tables 4.5 and 4.6 shows the

Table 4.4: Comparison of tangency portfolio Sharpe ratio

# of Factors	1	2	3	4	5
FF	0.52	0.41	0.54	0.69	0.69
LLR	0.48	-	-	-	-
CA-UV	0.53	1.15	0.99	1.62	1.77
GF-UV	0.52	**2.07**	**2.07**	2.09	2.09
CA-VV	0.76	0.78	0.89	1.51	1.82
GB-VV	**0.82**	1.95	1.99	**2.12**	**2.12**

Notes. Results in table denotes the annual Sharpe ratio of tangency portfolio of factors. The best performing models along the fixed number of factors are shown in bold.

statistical performance and Sharpe ratio of factor tangency portfolio, respectively. The results show that the overall result is robust with the result from whole test period, while GF-VV showed the highest Sharpe ratio for every number of factors.

4.5 Chapter Summary

We proposed a multi-factor asset pricing model with time-varying volatility condition. The proposed model is based on a neural network. The model considers volatility as realized volatility and proceeds with the volatility prediction, then utilizes predicted volatility as the regularization term of asset pricing loss. As a result, we proposed a training loss constructed of two parts, volatility estimation loss and residual divided by time-varying volatility. For the asset pricing part of the model, the conditional autoencoder model and the graph-based factor model are used as estimation methods. For the volatility prediction part, LSTM is used for estimation because it is known to work well on long-term forecasting in chapter 3.

The empirical results on U.S individual stocks show the following findings. First, in terms of out-of-sample prediction R-square, the graph-based factor model with time-unvarying volatility and time-varying volatility showed the best performance.

Table 4.5: Comparison of Out-of-sample R^2_{total} and R^2_{pred} during 2017-2021

	# of Factors	1	2	3	4	5
Out-of-sample total R-square (%)	FF	0.23	0.33	0.36	0.43	0.6
	LLR	2.12	-	-	-	-
	CA-UV	23.31	27.7	27.35	28.22	25.76
	GF-UV	25.02	**28**	**28.85**	28.57	**28.73**
	CA-VV	**27.08**	27.23	26.31	26.54	24.72
	GF-VV	26.2	25.45	27.49	**32.17**	24.81
Out-of-sample prediction R-square (%)	FF	<0	<0	<0	<0	<0
	LLR	<0	-	-	-	-
	CA-UV	1.22	2.65	2.91	3.38	2.63
	GF-UV	3.79	**5.03**	**5.13**	5.73	**5.55**
	CA-VV	2.06	2.74	2.55	3.2	2.78
	GF-VV	**4.25**	4.24	3.4	**7.22**	5.29

Notes. This table shows the empirical results of out-of-sample total R^2 and prediction R^2 during January 2017 to December 2021. The upper part of table represents the out-of-sample total R^2 and lower part represents the out-of-sample prediction R^2. The best performing models along the fixed number of factors are shown in bold.

Table 4.6: Comparison of tangency portfolio Sharpe ratio during 2017-2021

# of Factors	1	2	3	4	5
FF	0.5	0.45	0.57	0.73	0.73
LLR	0.49	-	-	-	-
CA-UV	0.58	1.16	1.07	1.61	1.8
GF-UV	0.61	2.01	2.01	2.05	2.05
CA-VV	0.83	0.86	1.03	1.58	1.99
GF-VV	**0.88**	**2.02**	**2.05**	**2.17**	**2.17**

Notes. Results in table denotes the annual Sharpe ratio of tangency portfolio of factors during January 2017 to December 2021. The best performing models along the fixed number of factors are shown in bold.

It is the consistent result with the findings from chapter 2, which showed the graph-based factor models have more explanation and prediction power than the Fama-French factor model and conditional autoencoder model. Second, time-varying volatility models show a large performance increase in total R-square and prediction R-square during the low-vol period. It shows that the expected volatility term in l^r loss effectively worked as a regularization term, and therefore fitted more strictly when expected volatility is low. Third, the graph-based factor models show the highest Sharpe ratio of tangency factor portfolio and the graph-based factor model time-varying volatility on the five-factor model shows the highest Sharpe ratio among all benchmark models. The superiority of the graph-based factor models for the Fama-French factor models and conditional autoencoder models are consistent with the findings from chapter 2. Moreover, even though the GF-UV model shows the highest Sharpe ratio in the two- and three-factor model, the fact that the Sharpe ratio of tangency factor portfolio of the GF-VV model in the four- and five-factor model is the highest for all shows that the GF-VV model consequently estimates the most efficient stochastic discount factor through its factors.

Chapter 5

Macroeconomic factor model and spillover-based volatility prediction for ERC-20 tokens

5.1 Chapter Overview

Blockchain, a distributed ledger, has grown rapidly, starting with Bitcoin and Ethereum. The second-largest cryptocurrency, Ethereum, made it possible to implement smart contracts on the chain, allowing other cryptocurrencies called tokens to be issued on the Ethereum blockchain. Therefore, the Ethereum blockchain plays an important role in the blockchain ecosystem, in which not only Ethereum itself but also numerous tokens are traded. This characteristic of the Ethereum blockchain makes Ethereum and Ethereum-based tokens have a structural association. Thus, it is expected that Ethereum and Ethereum-based tokens are highly likely to have certain common characteristics.

Many previous studies have explored factors that can explain the return of cryptocurrencies. The approaches used in these studies can be broadly classified into two categories: internal and external. The first approach tries to find internal factors that can be derived from the price of the cryptocurrency or technical specifications of blockchain, such as mining cost and reward. Momtaz (2021) showed that liquidity, market capitalization, and high-low price ratios can predict cryptocurrency returns, while Shen et al. (2020) showed that market, size, and reversal factors affect

cryptocurrency pricing. Studies have also been conducted on chain-related factors. Fantazzini and Kolodin (2020) analyzed the causal relationship between Bitcoin and its hashrate, and Meynkhard (2019) addressed the fair market value of Bitcoin in terms of the halving effect. The second approach attempts to find external factors outside the blockchain system that provide predictive power for cryptocurrency prices. Chen et al. (2021b) and Abraham et al. (2018) suggested that global economic factors and Twitter and Google Trends data are important for predicting the cryptocurrency price using a machine learning model. Aysan et al. (2019) showed that global geopolitical risks is related to the returns and volatility of Bitcoin.

Our study followed the previous approach. To the best of our knowledge, no previous study has found common characteristics of only Ethereum-based tokens in pricing. Urquhart (2021) offered a detailed analysis of the, Ethereum blockchain including gas value, but they did not provide a relationship between gas price and token price. In this study, we show that Ethereum gas price is related to the returns and volatility of Ethereum-based tokens. Our study contributes to existing studies by suggesting a new aspect: among many cryptocurrencies, cryptocurrencies that exist on the same blockchain (Ethereum) can have their own common characteristics, and gas can be the characteristic.

Gas is the cost that the transaction issuer must pay to publish a transaction on Ethereum. Since issuers have to pay gas to trade Ethereum or tokens in the chain, gas can be considered similar to the transaction fee paid when market participants trade traditional assets such as equity. However, gas has different characteristics from transaction fees in the following respects: First, transaction fees are determined at a fixed rate with respect to the transaction amount, whereas gas is proportional to the

size of the transaction code, regardless of the transaction amount. In addition, one of the most important characteristics of gas is that it is affected by the number of nodes trying to publish a transaction in the current Ethereum network, that is, the network congestion level, because it is determined in the form of a first-price auction. Therefore, the prices of assets traded in the Ethereum chain are structurally related to gas. However, the first price auction system makes it difficult for transaction issuers to decide the appropriate gas price bid and causes overpayment in the gas price. Recently, the EIP-1559 proposal was adopted in Ethereum to address this issue Buterin et al. (2019). It is expected that gas prices will be more stable and more precisely represent network congestion.

In this chapter, we empirically analyzed the relationship between the return and volatility of Ethereum-based tokens and gas returns using four different measures: a Pearson correlation analysis, the autoregressive regression (AR) and heterogeneous AR (HAR) based analysis, and the Granger causality test. The analysis was performed on two periods: pre-EIP-1559 and post-EIP-1559. The empirical results revealed four important findings. First, a strong correlation exists between the returns and volatility of Ethereum and gas returns in the same time horizon. Second, the returns and volatility of the Ethereum and ERC-20 tokens caused gas returns before EIP-1559 was adopted. Third, gas returns caused ERC-20 tokens' return in the post-EIP-1559 period. In addition, the predictive ability of Ethereum volatility on gas returns remained after EIP-1559. Lastly, ERC-721 token returns and volatility both do not show a clear pattern of relationship with gas returns over the entire period. Our results provide evidence that the price of ERC-20 tokens is now affected by the gas price because the gas price reflects the network congestion level well after

EIP-1559 was adopted, while the volatility of Ethereum affects the gas price. Along with the relationship analysis of tokens' volatility and gas return, we constructed the volatility prediction model for ETH and ERC-20 tokens with gas return using STG-Spillover model proposed in Chapter 3. The empirical result shows that including spillover effect of gas return can increase the volatility prediction performance of ETH and ERC-20 tokens affected by gas return.

From the analysis of Ethereum-based tokens' return and gas return, we confirmed that gas return has statistically significant effect on Ethereum-based tokens' return. Therefore, we also constructed the two factor model for Ethereum-based tokens. The two factors are cryptocurrency market return and Ethereum gas return. Cryptocurrency market return is derived as the size weighted return of whole cryptocurrencies. The empirical analysis of two factor model shows that starting with EIP-1559, the r-square of factor model increases for ERC-20 tokens and metaverse ERC-721 tokens. The decrease of intercept also confirms that Ethereum and ERC-20 tokens return are better explained by gas return after EIP-1559 adoption.

The remainder of this chapter is organized as follows. In Section 5.2, we explain the ERC standards and provide more details on the EIP-1559. Section 5.3 describes the measures and macroeconomic factor model used in this study. Section 5.4 presents the data used for the empirical analysis and following empirical results. Finally, concluding remarks are presented in Section 5.5.

5.2 Preliminaries

Ethereum-based tokens can be classified according to composition standards, and representative classifications include ERC-20 and ERC-721. ERC-20 has been the

standard for fungible tokens on the Ethereum blockchain owing to its appearance Vogelsteller and Buterin (2015). After the boom in ERC-20 tokens, the concept of ERC-721 tokens was proposed for non-fungible tokens (NFT) Entriken et al. (2018). Owing to the spike in demand for Ethereum-based tokens, gas fees have become a factor for Ethereum users, but the original fee mechanism made it extremely volatile. To address this problem, Buterin et al. (2019) proposed EIP-1559 introducing a new concept of base fee (b_t) which reflects the network condition at time t. Under this system, the user submits two parameters: maximum fee (f) and maximum priority fee (p). The maximum fee is the maximum amount of gas that the user would pay for their transaction, and the maximum priority fee is the maximum amount of tip for miners that the user would pay to include in their transaction. Under this protocol, the amount of gas paid by the user was $\max\{b_t + p, f\}$. For every block, the base fee, which is indexed by the block height t, is updated using the following equation:

$$b_{t+1} = b_t \left(1 + d \cdot \frac{G_t - T/2}{T/2}\right) \tag{5.1}$$

Here, d denotes an adjustment facto,r which is currently set to 0.125. G_t is the total amount of gas used for transactions included in block t and T is the block size.

5.3 Methodology

5.3.1 Relation analysis

For further analysis, we used the logarithmic return of asset price and logarithmic return of gas price at time t as r_t^a and r_t^g. We approximated asset volatility as the squared asset return at time t as $(\sigma_t^a)^2 = (r_t^a)^2$.

We used four methodologies to identify the relationship and causality between

gas and asset prices, following Baur and Dimpfl (2018). Each methodology was used twice to investigate the relationship between asset returns and gas returns and between asset price volatility and gas returns. Since there exists little consensus about the causal relationship between asset and gas prices, we investigated the relationship by changing the time sequence between the two. Each measure is calculated from the methodologies as follows. First, we calculated the Pearson correlation coefficient between asset return, volatility and gas return as follows:

$$\begin{aligned} \rho_t^{r,1} &= cor\,(r_{t+1}^a, r_t^g), \quad \rho_t^{r,2} = cor\,(r_t^a, r_{t+1}^g), \quad \rho_t^{r,3} = cor\,(r_t^a, r_t^g) \\ \rho_t^{v,1} &= cor\,({\sigma_{t+1}^a}^2, r_t^g), \ \rho_t^{v,2} = cor\,(\sigma_t^{a2}, r_{t+1}^g), \ \rho_t^{v,3} = cor\,(\sigma_t^{a2}, r_t^g) \end{aligned} \tag{5.2}$$

Among the four methodologies, only the Pearson correlation coefficient method observes the relationship between assets and gas in the same time horizon (see $\rho_t^{r,3}$ and $\rho_t^{v,3}$ in Equation 5.2).

Second, we implemented an AR analysis by adding an additional term to the independent variable. When predicting asset returns using the AR(1) model, an additional regression term of gas returns is provided to figure out if gas returns can provide additional predictive ability for asset returns. The implemented model is expressed as follows:

$$\begin{aligned} r_{t+1}^a &= \alpha_0^{a,1} + \alpha_1^{a,1} r_t^a + \beta^{a,1} r_t^g \\ {\sigma_{t+1}^a}^2 &= \alpha_0^{a,2} + \alpha_1^{a,2} \sigma_t^{a2} + \beta^{a,2} r_t^g \end{aligned} \tag{5.3}$$

$$\begin{aligned} r_{t+1}^g &= \alpha_0^{g,1} + \alpha_1^{g,1} r_t^g + \beta^{g,1} r_t^a \\ r_{t+1}^g &= \alpha_0^{g,2} + \alpha_1^{g,2} r_t^g + \beta^{g,2} \sigma_t^{a2} \end{aligned} \tag{5.4}$$

The corresponding coefficient β documents the effect of the added independent vari-

able on predicting asset returns, asset volatility, or gas returns.

Third, in the context of expanding the AR analysis, a HAR analysis was performed Corsi (2009). The HAR model, proposed by Corsi (2009), showed good performance in predicting realized volatility. We followed this idea and implemented the HAR analysis by affixing an additional term of the independent variable as follows:

$$\begin{aligned} r^a_{t+1} &= \alpha^{a,1}_0 + \alpha^{a,1}_1 r^a_t + \alpha^{a,1}_2 r^a_{t:t-4} + \alpha^{a,1}_3 r^a_{t:t-21} + \beta^{a,1} r^g_t \\ {\sigma^a_{t+1}}^2 &= \alpha^{a,2}_0 + \alpha^{a,2}_1 \sigma^{a2}_t + \alpha^{a,2}_2 {\sigma^a_{t:t-4}}^2 + \alpha^{a,2}_3 {\sigma^a_{t:t-21}}^2 + \beta^{a,2} r^g_t \end{aligned} \tag{5.5}$$

$$\begin{aligned} r^g_{t+1} &= \alpha^{g,1}_0 + \alpha^{g,1}_1 r^g_t + \alpha^{g,1}_2 r^g_{t:t-4} + \alpha^{g,1}_3 r^g_{t:t-21} + \beta^{g,1} r^a_t \\ r^g_{t+1} &= \alpha^{g,2}_0 + \alpha^{g,2}_1 r^g_t + \alpha^{g,2}_2 r^g_{t:t-4} + \alpha^{g,2}_3 r^g_{t:t-21} + \beta^{g,2} \sigma^{a2}_t \end{aligned} \tag{5.6}$$

where $r_{t:t-k} = \frac{1}{k}\sum_{i=0}^{k} r_i$ and ${\sigma^a_{t:t-k}}^2 = \frac{1}{k}\sum_{i=0}^{k} {\sigma^a_{t:t-i}}^2$. Therefore, $r^a_{t:t-4}$, ${\sigma^a_{t:t-4}}^2$, $r^a_{t:t-21}$, and ${\sigma^a_{t:t-21}}^2$ denote the weekly average of asset returns and volatility and the monthly average of asset returns and volatility, respectively. The corresponding coefficient β documents the effect of the added independent variable on predicting asset returns, asset volatility, or gas returns.

Lastly, we conducted a Granger causality test to determine the inference structure. To test the null hypothesis that stationary time series $X : x_{1,\dots,T}$ does not Granger-cause other stationary time series $Y : y_{1,\dots,T}$ for lag p or vice versa, the augmented formulation of autoregressive model is used as follows:

$$\begin{aligned} x_t &= \alpha_{x,0} + \sum_{i=1}^{p} \alpha_{x,i} x_{t-i} + \sum_{i=1}^{p} \beta_{x,i} y_{t-i} + \epsilon_{x,t} \\ y_t &= \alpha_{y,0} + \sum_{i=1}^{p} \alpha_{y,i} y_{t-i} + \sum_{i=1}^{p} \beta_{y,i} x_{t-i} + \epsilon_{y,t} \end{aligned} \tag{5.7}$$

The null hypothesis becomes $H_0 : \sum_{i=1}^{p} \beta_{x,i} = 0$ when Y Granger-causes X. The

hypothesis was tested using an F-test.

5.3.2 Factor model analysis

We used fundamental-based factor model for constructing asset pricing model. The fundamental factor model assumes that some macroeconomic variable affects the assets as a whole. The most well-known example is the one factor model CAPM proposed by Sharpe (1964) and Lintner (1965). The mathematical formulation of CAPM is expressed as follows:

$$R_{i,t} = \alpha_{i,t} + \beta_{i,t} R_t^m + \epsilon_{i,t} \tag{5.8}$$

where $R_{i,t}$ is the excess return of asset i at time-step t, $\alpha_{i,t}$ is the intercept, $\beta_{i,t}$ is the beta coefficient also known as risk exposure, R_t^m is the excess market return at time-step t, and $\epsilon_{i,t}$ is the residual.

While CAPM assumes that the excess market return can explain asset return, we proposed two factor model by adding gas return term to CAPM. The mathematical formulation is as follows:

$$R_{i,t} = \alpha_{i,t} + \beta_{i,t}^1 R_t^m + \beta_{i,t}^2 R_t^g + \epsilon_{i,t} \tag{5.9}$$

Since the purpose of the factor model is to analyze whether the fundamental factor can account for excess returns, R^2 is used as an evaluation index. Along with R^2, the absolute value of intercept is also important. If the market is perfectly efficient, the well designed factor model can fully explain the excess return and it means that there exists no anomaly. Therefore, no anomaly gets the same meaning as the intercept of factor model becomes zero.

5.3.3 Volatility prediction with volatility spillover index

We used the developed volatility prediction model from Chapter 3. The developed model is based on spatial-temporal graph neural network structure, and estimates graph structure using the volatility spillover index proposed by Diebold and Yilmaz (2012). For the rest of the chapter, the used volatility prediction model is referred as STG-Spillover model.

We compared the volatility forecasting performance of two STG-Spillover models, where one is trained with only volatility data of ETH and ERC-20 tokens, and another is trained with volatility data and gas return. The analysis is aimed to identify the effect of gas return for predicting volatility of ERC-20 tokens.

5.4 Empirical Studies

5.4.1 Data

To examine the relationship between the Ethereum gas and Ethereum-based tokens, this study included ETH, ERC-20, and ERC-721 based tokens. Among the ERC-20 tokens, we selected MANA, AXS, SAND, ENJ, and CHZ, considering their use in NFT trades and the awareness of users. We obtained the hourly prices of these currencies from Coinmarketcap[1] and analyzed them on a daily basis. Prices were collected between November 2020 and January 2022.

We divided ERC-721 based tokens into two subcategories, metaverse and collectibles. Among each domain, we selected the top three representative tokens regarding the market capitalization, which is defined as the sum of the last price of the NFT collection in the project and its awareness. For metaverse-related ERC-721

[1] `https://coinmarketcap.com`

tokens, we obtained the prices of secondary market trades in Decentraland LAND tokens, Sandbox LAND tokens and every recorded sales price in the Axie Infinity marketplace, following approaches from Dowling (2022b) and Dowling (2022a). The recorded price of each transaction was sourced from NonFungible.com[2] and then aggregated into daily frequency. For ERC-721 based collectibles, we collected daily average sales prices in USD for CryptoPunk images, Art Blocks, and Bored Ape Yacht Club images from Cryptoslam[3]. The average gas price was downloaded from EtherScan[4]. The gas price has its own unit called Gwei, which is 10^{-9} ETH.

EIP-1559 was implemented on August 5, 2021. To analyze its impact on asset prices, we divided our data into two parts for analysis: pre-EIP-1559 and post-EIP-1559. The descriptive statistics of daily asset returns throughout this period are presented in Table 5.1. From this table, we verify the fact that EIP-1559 stabilized the Ethereum gas price. To check the stationarity of returns, we conducted Augmented Dickey-Fuller test and the results are shown in the ADF column in Table 5.1. Every asset shows a stationary return, which justifies our use of the AR, HAR-based method.

For the factor model construction, we used CRIX index, introduced by Trimborn and Härdle (2018), as the cryptocurrency market return. It can be downloaded from CRIX[5]. The CRIX index is value-weighted return of representative cryptocurrencies on market, which is calculated as follows:

$$R_t^m = \frac{\sum_{i=1}^{N} R_{i,t} P_{i,t} D_{i,t}}{\sum_{i=1}^{N} P_{i,t} D_{i,t}} \tag{5.10}$$

[2] `https://nonfungible.com`
[3] `https://cryptoslam.io`
[4] `https://etherscan.io/chart/gasprice`
[5] `https://www.royalton-crix.com/`

where $P_{i,t}$ and $D_{i,t}$ each denotes the price of asset i at time-step t and the quantity, respectively. The time-series visualization of price and return of CRIX index is presented in Figure 5.1.

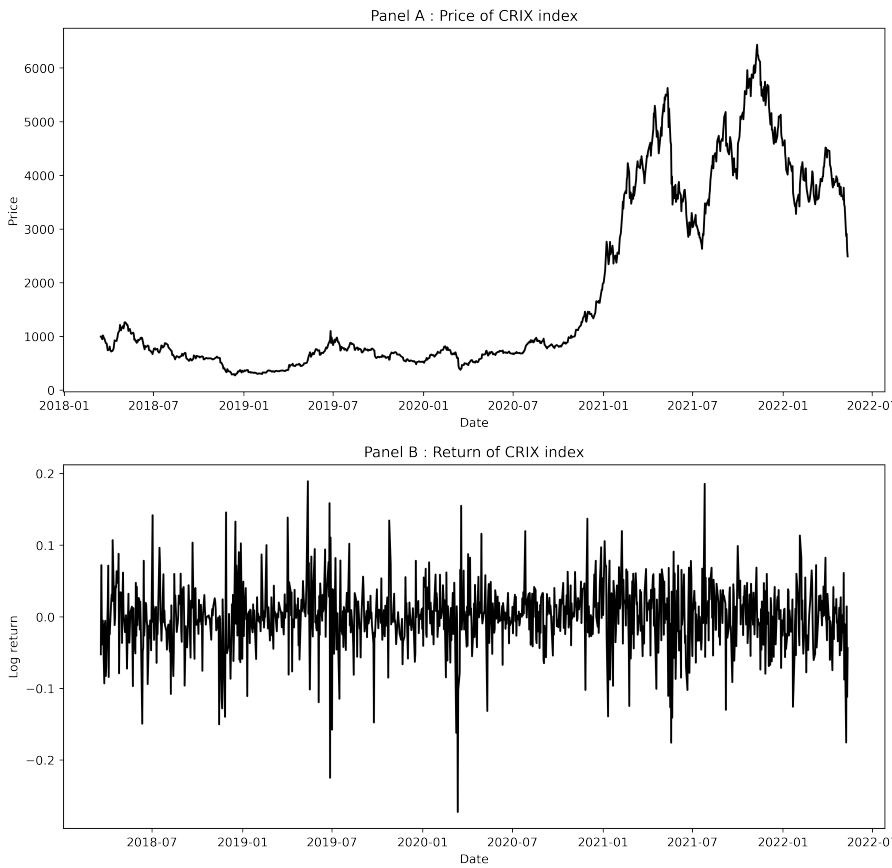

Figure 5.1: Price and return of CRIX index

Table 5.1: Descriptive Statistics

	Mean	Median	Std Dev	Min	Max	Skewness	Kurtosis	ADF
Panel A: Pre-EIP-1559								
ETH	0.0090	0.0072	0.0664	-0.3411	0.3152	0.1409	5.2734	-4.7057***
MANA	0.0144	-0.0020	0.1108	-0.3707	0.7203	1.7531	8.8607	-8.3666***
AXS	0.0295	0.0032	0.1394	-0.4499	0.5920	1.1980	2.8290	-10.8749***
SAND	0.0174	0.0059	0.1216	-0.4306	0.8850	1.9695	11.5045	-7.5020***
ENJ	0.0152	0.0033	0.1184	-0.4275	0.8565	2.1098	11.9351	-5.6011***
CHZ	0.0192	0.0086	0.1308	-0.4211	1.0571	3.3021	21.6978	-4.6598***
Decentraland LAND	0.1537	0.0136	1.3336	-0.9199	27.2818	16.2858	320.4130	-25.5536***
Sandbox LAND	0.0818	0.0131	0.6083	-0.9192	10.1112	9.7400	143.5687	-5.4059***
Axie Infinity	0.2691	0.0021	1.6060	-0.9489	22.3549	8.2915	88.0370	-10.5719***
CryptoPunks	0.0624	0.0277	0.3706	-0.7291	2.0745	1.9543	8.6931	-10.9276***
ArtBlocks	0.2047	0.0166	0.8512	-0.9309	5.8502	3.5911	20.0720	-3.5434***
Bored Ape	0.0878	0.0258	0.3309	-0.4037	2.0874	3.6665	19.1721	-8.6524***
Gas Price (Gwei)	0.0557	-0.0009	0.3842	-0.6831	2.3940	2.2086	8.4587	-9.7199***
Panel B: Post-EIP-1559								
ETH	0.0005	0.0014	0.0413	-0.1386	0.1118	-0.1118	0.8886	-10.2350***
MANA	0.0138	-0.0003	0.1519	-0.2045	1.7354	8.3656	93.2101	-7.8596***
AXS	0.0040	-0.0068	0.0842	-0.1529	0.6182	3.9421	25.2381	-13.6219***
SAND	0.0158	-0.0037	0.1206	-0.1744	0.8513	3.2727	17.1215	-7.1185***
ENJ	0.0036	0.0032	0.0719	-0.1772	0.3475	0.5941	2.4248	-14.6986***
CHZ	0.0004	-0.0025	0.0751	-0.2058	0.5674	2.2705	17.8680	-15.4614***
Decentraland LAND	0.0866	0.0232	0.4477	-0.7706	3.0995	2.3763	11.8646	-9.6752***
Sandbox LAND	0.0199	0.0156	0.1510	-0.4532	0.8728	1.2236	6.4411	-19.5031***
Axie Infinity	-0.0102	-0.0076	0.0472	-0.2008	0.1616	0.0098	2.5328	-12.9624***
Cryptopunks	0.0421	0.0000	0.3274	-0.7015	2.0537	3.1444	16.4858	-11.3029***
ArtBlocks	0.1368	-0.0235	0.6112	-0.7931	2.8745	1.5066	3.2678	-20.7239***
Bored Ape	0.0341	0.0056	0.2514	-0.5846	1.7099	2.5320	13.8854	-4.9299***
Gas Price (Gwei)	0.0255	-0.0145	0.2192	-0.4041	0.9432	0.9033	1.6474	-5.0258***
Panel C: Entire Period								
ETH	0.0058	0.0046	0.0579	-0.3411	0.3152	0.1923	6.0169	-6.1197***
MANA	0.0142	-0.0012	0.1284	-0.3707	1.7354	6.1012	73.7546	-12.7673***
AXS	0.0195	-0.0002	0.1210	-0.4499	0.6182	1.7415	5.8406	-13.8636***
SAND	0.0169	0.0023	0.1209	-0.4306	0.8850	2.4679	13.5080	-9.1195***
ENJ	0.0107	0.0033	0.1025	-0.4275	0.8565	2.0998	13.7466	-7.1094***
CHZ	0.0117	0.0035	0.1122	-0.4211	1.0571	3.5095	26.5599	-6.3937***
Decentraland LAND	0.1364	0.0151	1.1747	-0.9199	27.2818	17.8578	398.3796	-29.7540***
Sandbox LAND	0.0674	0.0134	0.5327	-0.9192	10.1112	10.9081	183.8055	-6.1605***
Axie Infinity	0.1986	-0.0038	1.3937	-0.9489	22.3549	9.6201	118.5374	-11.9709***
Cryptopunks	0.0494	0.0092	0.3421	-0.7291	2.0745	2.6317	12.8169	-13.6599***
ArtBlocks	0.1596	-0.0002	0.7031	-0.9309	5.8502	2.8737	16.3893	-24.5041***
Bored Ape	0.0545	0.0099	0.2831	-0.5846	2.0874	3.2192	18.2648	-5.6472***
Gas Price (Gwei)	0.0445	-0.0045	0.3289	-0.6831	2.3940	2.2704	10.3603	-4.9452***

Note. Asterisks flag levels of statistical significance of result statistic in ADF Test. The significance levels are flagged as follows: *** : p-value < 0.01

Table 5.2: Result Statistics (Return)

Measure	**Pearson Correlation Coefficient**			**AR beta**		**HAR beta**	
	$\rho_t^{r,1}$	$\rho_t^{r,2}$	$\rho_t^{r,3}$	$\beta^{a,1}$	$\beta^{g,1}$	$\beta^{a,1}$	$\beta^{g,1}$
Panel A: Pre-EIP-1559							
ETH	0.1200**	0.1903***	-0.1613***	0.0205	0.8697***	0.0211	0.7496**
MANA	0.1573***	0.1067	-0.2704***	0.0471***	0.1927	0.0581***	0.1570
AXS	0.0624	0.1335**	-0.1932***	0.0218	0.2539	0.0351	0.2658
SAND	0.1157	0.1767***	-0.2104***	0.0345	0.4187**	0.0330	0.3756**
ENJ	0.1391**	0.0954	-0.2534***	0.0404**	0.1504	0.0421**	0.0822
CHZ	0.0514	0.0967	-0.1418**	0.0189	0.2120	0.0163	0.1624
Decentraland LAND	-0.0711	0.0725	-0.0352	-0.1138**	0.0520	-0.0967**	0.0474
Sandbox LAND	-0.0112	0.0230	-0.0559	-0.0373	0.0170	-0.0524	0.0093
Axie Infinity	-0.0063	0.0961**	-0.0176	-0.0221	0.0547**	-0.0044	0.0521**
CryptoPunks	0.0358	0.2147**	-0.1115	-0.0007	0.2102	-0.0456	0.0431
ArtBlocks	-0.2989***	-0.1286	0.4391***	-0.1410	-0.0341	-0.2745	-0.0289
Bored Ape	-0.1257	0.0114	-0.0753	-0.0659	-0.0034	-0.0581	0.2205
Panel B: Post-EIP-1559							
ETH	-0.1219	0.1143	-0.1707**	-0.0261	0.5192	-0.0148	0.6202
MANA	-0.0627	-0.0491	-0.1480**	-0.0227	-0.1177	-0.0140	0.0774
AXS	-0.0253	0.0286	-0.2335***	-0.0047	0.0488	-0.0072	0.1634
SAND	-0.0655	-0.0211	-0.1130	-0.0244	-0.0619	-0.0237	0.0949
ENJ	-0.1078	0.0060	-0.1480**	-0.0395	-0.0057	-0.0298	0.1079
CHZ	-0.0720	0.0571	-0.1889**	-0.0294	0.1478	-0.0229	0.3145
Decentraland LAND	0.0159	-0.0354	0.0252	0.0528	-0.0182	-0.0051	-0.0193
Sandbox LAND	0.0160	0.0365	-0.0902	-0.0115	0.0459	-0.0091	0.0873
Axie Infinity	-0.0547	0.1179	-0.0026	-0.0128	0.5095	-0.0113	0.5970
CryptoPunks	-0.0058	0.0480	-0.0985	-0.0615	0.0331	-0.0865	0.0389
ArtBlocks	-0.0626	0.1479	-0.0124	-0.1781	0.0601	-0.1782	0.0542
Bored Ape	-0.0123	0.2887***	-0.1962***	-0.0994	0.2728***	-0.1062	0.2861***

Note. This table reports result statistics for Pearson's Correlation Coefficient method, AR method and HAR method on asset return and gas return. Greek letters in the second row are from the implemented model we proposed in Methodology section. Asterisks flag levels of statistical significance of result statistic using t-test. The significance levels are flagged as follows: *** : p-value < 0.01, ** : p-value < 0.05

Table 5.3: Granger Causality Test Statistics (Return)

	F-statistics							
	H_0**: Gas return granger causes token return**				H_0**: Token return granger causes gas return**			
	lag = 1	lag = 2	lag = 3	lag = 4	lag = 1	lag = 2	lag = 3	lag = 4
Panel A: Pre-EIP-1559								
ETH	3.2864	1.4906	1.0616	1.1533	7.3141***	4.7362***	4.1741***	4.8584***
MANA	7.2013***	3.9163**	3.5035**	2.2526	0.7734	0.7057	2.0644	1.7440
AXS	0.8063	0.6682	1.8314	1.7306	2.5756	1.5950	2.4918	2.1100
SAND	2.8499	1.1757	2.1604	1.4794	5.3136**	3.9254**	2.3072	1.8323
ENJ	4.6386**	2.5752	2.5273	2.0484	0.5457	0.4814	1.2011	2.3961
CHZ	0.9146	0.4652	0.4316	0.2794	1.3053	0.7609	1.0096	0.8187
Decentraland LAND	5.4921**	2.4280	1.4251	1.1870	2.9266	0.8503	0.4807	0.5193
Sandbox LAND	0.9861	0.8278	0.6466	0.8349	0.1875	0.0545	0.0974	1.5137
Axie Infinity	0.1218	0.9915	0.3752	0.2305	5.4913**	3.0069	2.0569	1.6517
CryptoPunks	0.0001	3.5330**	2.3229	1.4425	3.7734	2.8582	1.2561	1.7076
ArtBlocks	0.7247	0.7995	0.7296	0.7703	0.2509	1.4578	0.7376	1.0362
Bored Ape	1.2131	1.4821	0.9498	1.0111	0.0005	1.1707	0.4395	0.5880
Panel B: Post-EIP-1559								
ETH	2.7279	4.4750**	2.9754**	2.1329	2.0177	0.6218	1.3020	2.6178**
MANA	0.3687	4.7892***	4.1101***	3.7826***	0.5880	0.5910	0.3250	0.7645
AXS	0.0320	2.2650	1.5651	1.5342	0.0471	0.0227	0.3532	0.8216
SAND	0.5269	5.5260***	3.9263***	3.0294**	0.1321	0.2145	0.4591	1.1996
ENJ	2.3342	6.6778***	4.3324***	3.0357**	0.0007	0.7317	0.4269	2.0582
CHZ	1.3889	4.1548**	3.1477**	2.3464	0.4009	0.1455	0.2482	1.0547
Decentraland LAND	0.1627	0.3586	0.3744	0.0736	0.1981	0.8405	0.7806	0.6125
Sandbox LAND	0.0535	0.1072	0.1508	0.1058	0.1805	0.5705	0.8263	1.3843
Axie Infinity	0.5256	0.4339	0.6386	0.3880	2.4568	1.6976	1.1612	1.3336
CryptoPunks	0.4628	0.3920	0.5437	0.7981	0.3271	0.1713	0.1093	1.2697
ArtBlocks	1.2033	0.3336	0.1921	0.8865	3.8281	1.5672	1.0156	0.8647
Bored Ape	1.7617	0.1294	0.3558	0.7808	15.2563***	7.9090***	5.2779***	4.5278***

Note. This table reports result statistics for Granger Causality test from lag 1 to lag 4 on asset return and gas return. Asterisks flag levels of statistical significance of result statistic using t-test. The significance levels are flagged as follows: *** : p-value < 0.01, ** : p-value < 0.05

Table 5.4: Result Statistics (Volatility)

Measure	**Pearson Correlation Coefficient**			**AR beta**		**HAR beta**	
	$\rho_t^{v,1}$	$\rho_t^{v,2}$	$\rho_t^{v,3}$	$\beta^{a,2}$	$\beta^{g,2}$	$\beta^{a,2}$	$\beta^{g,2}$
Panel A: Pre-EIP-1559							
ETH	0.0681	-0.3516***	0.2150***	-0.0007	-18.3156***	0.0007	-17.5412***
MANA	0.0665	-0.2597***	0.1812***	0.0021	-3.4358***	0.0036	-3.0945***
AXS	0.0187	-0.1898	0.2197***	-0.0013	-1.8330**	0.0006	-2.2128***
SAND	-0.0119	-0.0893	0.1299**	-0.0030	-0.4963	-0.0025	-0.3383
ENJ	0.0138	-0.2407***	0.1525**	-0.0007	-2.7949***	0.0004	-2.7245***
CHZ	0.0468	-0.2182***	0.1655***	-0.0003	-2.6788***	0.0022	-2.2467***
Decentraland LAND	-0.0151	0.0158	0.0031	-0.0354	0.0075	-0.0418	0.0020
Sandbox LAND	0.0141	0.0500	-0.0076	0.0241	0.0361	0.0296	0.0396
Axie Infinity	0.0303	-0.0163	0.0064	0.0856	-0.0050	0.0566	-0.0061
CryptoPunks	-0.0597	-0.0535	0.2631***	-0.1017	-0.0074	-0.0495	-0.5432
ArtBlocks	0.0084	-0.0513	-0.0378	0.0376	-0.0238	0.0600	-0.1578
Bored Ape	-0.0567	-0.0009	-0.0195	-0.0140	-0.0091	-0.0314	0.0079
Panel B: Post-EIP-1559							
ETH	0.1111	-0.2378***	0.2080***	0.0003	-32.2404***	0.0009	-28.9441***
MANA	-0.0033	-0.1088	-0.0205	0.0008	-1.2861	0.0000	-0.1996
AXS	0.1773**	-0.1248	0.0089	0.0051**	-4.1831	0.0048***	-2.5558
SAND	0.0532	-0.1224	0.0424	0.0015	-2.0330	0.0001	-1.2005
ENJ	0.0503	-0.1433	0.1258	0.0003	-4.3704	0.0001	-3.3768
CHZ	0.0955	-0.1871**	0.0790	0.0016	-6.6648**	0.0001	-4.8698
Decentraland LAND	0.0242	0.0178	0.0286	0.0180	0.0132	0.0237	0.0380
Sandbox LAND	-0.0199	-0.0768	0.0938	-0.0092	-0.3121	-0.0004	0.0588
Axie Infinity	0.0091	-0.1259	-0.0171	0.0003	-4.9585	0.0002	-4.8751
CryptoPunks	-0.1021	0.0265	-0.0513	-0.0804	0.0236	-0.0943	-0.0261
ArtBlocks	-0.0733	-0.0684	-0.1552**	-0.0830	-0.0452	-0.0275	-0.0262
Bored Ape	-0.0686	0.0699	0.0494	-0.0554	0.1179	-0.0589	-0.0351

Note. This table reports result statistics for Pearson's Correlation Coefficient method, AR method and HAR method on asset volatility and gas return. Greek letters in the second row are from the implemented model we proposed in Methodology section. Asterisks flag levels of statistical significance of result statistic using t-test. The significance levels are flagged as follows: *** : p-value < 0.01, ** : p-value < 0.05

Table 5.5: Granger Causality Test Statistics (Volatility)

	F-statistics							
	H_0: **Gas return granger causes token volatility**				H_0: **Token volatility granger causes gas return**			
	lag = 1	lag = 2	lag = 3	lag = 4	lag = 1	lag = 2	lag = 3	lag = 4
Panel A: Pre-EIP-1559								
ETH	0.5500	1.5445	3.7444**	1.5585	31.0551***	14.2211***	8.7081***	6.5410***
MANA	0.2944	0.4903	0.6007	0.3055	15.1906***	6.7749***	4.3214***	3.1916**
AXS	0.0707	0.0538	0.1285	0.2706	6.2948**	3.0249	2.8620**	2.1605
SAND	0.1532	0.6246	0.7495	0.8168	1.1156	3.1726**	2.6627**	2.1524
ENJ	0.0253	0.0178	0.3648	0.3410	13.2342***	6.5321***	4.6766***	3.5120***
CHZ	0.0087	0.3573	1.2611	0.4208	10.1692***	4.7636***	2.4910	2.2725
Decentraland LAND	0.1413	0.3133	0.6318	0.7942	0.1413	0.3133	0.6318	0.7942
Sandbox LAND	0.2162	1.1384	0.5649	0.3885	1.3227	0.8583	0.7350	3.6101***
Axie Infinity	0.5059	0.4006	0.3106	0.2532	0.1352	0.0244	0.0136	0.2711
CryptoPunks	2.3908	2.1448	2.1777**	1.6787	0.0021	0.4851	0.4129	0.4450
ArtBlocks	0.0225	0.7480	1.3237	0.8653	0.3270	0.9816	1.1302	0.8259
Bored Ape	0.3251	2.0443	1.3579	1.6907	0.0017	0.9829	0.7525	0.6080
Panel B: Post-EIP-1559								
ETH	0.4169	3.0859**	2.1110	1.7478	9.9522***	3.8828**	2.5461	3.3383**
MANA	0.0310	0.3183	1.4739	1.6791	2.1480	1.5106	0.9844	1.0210
AXS	6.3314**	3.9496**	2.9983**	2.4774**	2.7617	1.1092	0.6864	1.1431
SAND	0.2123	0.2082	0.3790	0.0784	2.5757	1.0471	1.4987	1.3140
ENJ	0.0187	0.5575	0.5977	0.3994	3.3901	1.4370	1.2768	1.2839
CHZ	0.8090	0.8683	0.9956	1.1226	6.1149**	2.3388	2.0209	3.1392**
Decentraland LAND	0.0309	0.2186	0.3631	0.4678	0.0667	1.4177	1.0668	0.8459
Sandbox LAND	0.2878	1.1393	0.9102	1.7045	0.9578	0.9155	0.7517	0.7374
Axie Infinity	0.0335	0.2577	0.2069	0.2547	2.9057	1.5283	1.0732	0.8763
CryptoPunks	1.3598	0.6032	0.8968	0.7467	0.0993	0.0290	0.6466	0.4117
ArtBlocks	0.4032	3.7403**	2.7664**	2.3097	1.0341	0.9420	0.6445	0.7283
Bored Ape	1.7193	0.7301	1.6206	1.5209	0.9187	1.6289	1.1450	1.8266

Note. This table reports result statistics for Granger Causality test from lag 1 to lag 4 on asset volatility and gas return. Asterisks flag levels of statistical significance of result statistic using t-test. The significance levels are flagged as follows: *** : p-value < 0.01, ** : p-value < 0.05

5.4.2 Empirical Results

Tables 5.2 and 5.3 show the empirical results for the four proposed measures of asset returns and gas returns. There is a strong negative correlation between the returns and volatility of Ethereum and ERC-20 tokens and gas returns over the entire period. Negative correlation represents the phenomenon in which gas prices rise rapidly when the Ethereum token market plummets, such as Black Thursday. ETH, SAND, and Axie Infinity return Granger-causes and have predictive power for gas returns in the pre-EIP-1559 period. However, after EIP-1559 adoption, the trend is reversed for ETH, while SAND and Axie Infinity returns no longer have a relationship with gas returns. Gas returns Granger-causes ETH returns and ERC-20 tokens' returns when lag is longer than two. Since EIP-1559 aims to structure the gas price valuation system systematically, it stabilizes the gas price and makes it reflect network transaction congestion more precisely. Therefore we can conclude that EIP-1559 caused network congestion levels to affect ETH and ERC-20 based tokens' returns. Despite the clear relationship between ERC-20 tokens' returns and gas returns, no correlation exists between ERC-721 tokens' returns and gas returns regardless of EIP-1559 adoption, except for ArtBlocks and Bored Ape. During the post-EIP-1559 period, the Bored Ape returns had a causal and non-causal effect on gas returns.

The empirical analysis of the relationship between asset volatility and gas returns is presented in Tables 5.4 and 5.5. For the Pearson correlation coefficient measurement result, the most representative finding is that a positive correlation of ERC-20 tokens' returns and gas returns on the same time horizon exists in the pre-EIP-1559 period but disappears afterward, while the correlation between ETH and gas returns

remains. The positive correlation is explained by the fact that when the price fluctuation of Ethereum is large, network congestion increases and the gas price rises. Looking at the AR, HAR, and Granger causality test results, before EIP-1559 is applied to the Ethereum blockchain, the volatility of ERC-20 tokens consistently affects and Granger-causes gas returns except for the SAND tokens. After EIP-1559 adoption, every causal effect between ERC-20 token volatility and gas returns disappears except for the CHZ token, while the effect of ETH remains. The predictive power of the ERC20 token volatility remains for some, but many disappear, and the significance level decreases. Both Ethereum and ERC-20 tokens are assets that transfer in transactions on the Ethereum network, but considering that Ethereum is the main asset and the rest are sub-tokens, EIP-1559 has ensured that sub-tokens do not affect the gas price by controlling the gas price more systematically and predictably. Similarly, with asset return analysis, the ERC-721 tokens did not show any clear pattern in relation to gas. This seems to be the opposite result of the fact that OpenSea, the ERC-721 token trading platform, accounts for the majority of total gas consumption. This is due to the unique characteristics of NFTs, which are expensive and infrequently traded. The proportion of the gas price for trading the NFT is low because of the relatively high price of the NFT. In addition, low trading frequency indicates that NFTs have little incentive to react sensitively to temporary gas price fluctuations.

Table 5.6 shows the statistics of constructed two-factor model. Compared to pre-EIP-1559 period, in post-EIP-1559 period, R^2 of ETH, ERC-20 tokens, and ERC-721 metaverse tokens have increased, while R^2 of CryptoPunks and ArtBlocks decreased. The results of factor analysis imply the same meaning as the result that

Ethereum and ERC-20 token returns are more related to gas return after the EIP-1559 adoption from relationship analysis. This tendency shows the same result when observing absolute intercept. The absolute intercept of ERC-20 tokens decreases after the EIP-1559 adoption. We can confirm that after EIP-1559, gas return better explain ERC-20 tokens return since its stability has increased.

Another interesting point is that the gas return coefficient of Ethereum has positive sign during pre-EIP-1559 period. Every gas return coefficients of ERC-20 tokens have negative sign in whole period and even Ethereum has negative sign in post-EIP-1559 period. We believe that it is due to the noise in gas return during the pre-EIP-1559 period. Transaction issuers had overestimated the gas price because there is no guideline as base fee before EIP-1559 is adopted. The elimination of overestimation by EIP-1559 had derived both Ethereum and ERC-20 tokens to have negative gas return coefficient.

Lastly, we compared the volatility prediction performance of STG-Spillover model for ETH and ERC-20 tokens, with and without the gas return. From the previous analysis, we could confirm that gas return has correlation with volatility of certain tokens. In the pre-EIP-1559 period, gas return helps predicting volatility of ETH, MANA, AXS, SAND, and CHZ. After the EIP-1559 adoption, gas return affects volatility of ETH and CHZ. We first constructed the net pairwise volatility spillover index graph of ETH, ERC-20 tokens, and gas return to quantitatively measure the effect between assets. The constructed graph is shown in Figure 5.2. The graph shows the similar structure with the result of Table 5.4.

Using the constructed volatility spillover graph, the empirical results of STG-Spillover models are shown in Tables 5.7 and 5.8. For each forecast horizon, one

Table 5.6: Summary statistics for two factor model analysis

Measure	R square	Intercept	Market return coef	Gas return coef
Panel A : Pre-EIP-1559				
ETH	0.573	0.001	0.869	0.037
MANA	0.35	0.008	0.892	-0.079
AXS	0.129	0.022	0.751	-0.071
SAND	0.12	0.013	0.582	-0.062
ENJ	0.303	0.013	0.924	-0.091
CHZ	0.157	0.018	0.771	-0.058
Decentraland LAND	0.007	0	0.786	-0.002
Sandbox LAND	0.005	0.001	-0.032	-0.065
Axie Infinity	0.003	-0.001	0.673	-0.014
CryptoPunks	0.086	0.006	1.047	-0.198
ArtBlocks	0.146	-0.015	0.973	0.685
Bored Ape	0.032	0.01	0.551	0.027
Panel B : Post-EIP-1559				
ETH	0.727	-0.001	1.009	-0.023
MANA	0.262	0.004	0.953	-0.073
AXS	0.405	0.005	1.083	-0.097
SAND	0.182	0.01	0.929	-0.056
ENJ	0.407	0	1.135	-0.03
CHZ	0.487	0	1.088	-0.055
Decentraland LAND	0.016	0.017	1.202	0.09
Sandbox LAND	0.04	0.016	0.356	-0.136
Axie Infinity	0.189	-0.009	0.505	0.011
CryptoPunks	0.008	-0.002	0.376	-0.092
ArtBlocks	0.008	-0.02	-0.147	0.219
Bored Ape	0.096	0.012	1.131	-0.27

Note. This table reports the R^2, intercept α, beta coefficients of market return and gas return.

model is trained with tokens' volatility and gas, and another model is trained with only tokens' volatility. The empirical results show that gas return can help forecasting volatility of ETH and ERC-20 tokens, which are affected by gas return in terms of both linear time-series and spillover effect.

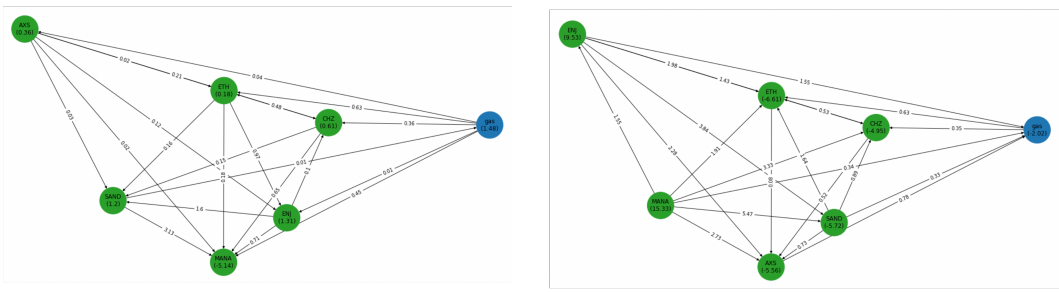

Figure 5.2: Spillover graph of ETH, ERC-20 tokens, and gas
Notes. Left figure shows the graph constructed in pre-EIP-1559 period and right figure is the one constructed in post-EIP-1559 period.

5.5 Chapter Summary

In this chapter, we conducted the first analysis of the relationship between Ethereum gas price and Ethereum-based token price. The analysis was conducted in two periods based on EIP-1559. The empirical results show that the returns and volatility of Ethereum and gas returns have strong correlations. In addition, before EIP-1559, Ethereum and ERC-20 token returns Granger-cause gas returns. However, after EIP-1559 adoption, the pattern is reversed. We also showed that the volatility of ERC-20 tokens has a predictive ability for gas returns in the pre-EIP-1559 period but it disappears in the post-EIP-1559 period, while the predictive ability of Ethereum volatility on gas returns remains. Finally, the ERC-721 token price has no clear relationship with gas price. From these findings, we conclude that after EIP-1559, only Ethereum volatility affects gas returns among various tokens and the price of ERC-

Table 5.7: Out-of-sample results of STG-Spillover models on dataset consisted of ETH and ERC-20 tokens with and without gas return

Forecast horizon h	1		5		10		22	
	With gas	Without gas	With gas	Without gas	With gas	Without gas	With gas	Without gas
Panel A : Pre-EIP-1559								
ETH	0.004	0.005	0.005	0.007	0.007	0.008	0.007	0.008
MANA	0.009	0.011	0.01	0.012	0.011	0.013	0.012	0.015
AXS	0.013	0.017	0.014	0.018	0.016	0.018	0.019	0.02
SAND	0.012	0.011	0.013	0.012	0.013	0.013	0.015	0.018
ENJ	0.009	0.009	0.011	0.012	0.012	0.014	0.015	0.018
CHZ	0.01	0.012	0.01	0.013	0.011	0.015	0.011	0.016
Mean	0.010	0.011	0.011	0.012	0.012	0.014	0.013	0.016

Forecast horizon h	1		5		10		22	
	With gas	Without gas	With gas	Without gas	With gas	Without gas	With gas	Without gas
Panel B : Post-EIP-1559								
ETH	0.002	0.004	0.004	0.006	0.005	0.009	0.007	0.01
MANA	0.007	0.008	0.008	0.008	0.01	0.009	0.012	0.011
AXS	0.01	0.009	0.012	0.011	0.014	0.013	0.015	0.015
SAND	0.009	0.01	0.01	0.011	0.011	0.011	0.014	0.014
ENJ	0.008	0.008	0.009	0.009	0.01	0.009	0.011	0.012
CHZ	0.01	0.012	0.012	0.014	0.013	0.015	0.015	0.017
Mean	0.008	0.009	0.009	0.010	0.011	0.011	0.012	0.013

Note. This table reports out-of-sample MAFE of STG-Correlation model on ETH and ERC-20 tokens. With gas denotes the dataset containing gas return and without gas denotes the dataset without gas return.

Table 5.8: DM test result for STG-Spillover model trained with gas versus without gas

Forecast horizon h	1		5		10		22	
	DM statistic	p-value	DM statistic	p-value	DM statistic	p-value	DM statistic	p-value
Panel A : Pre-EIP-1559								
ETH	-7.83	<0.01	-7.58	<0.01	-3.53	<0.01	-3.71	<0.01
MANA	-5.11	<0.01	-4.32	<0.01	-4.19	<0.01	-6.04	<0.01
AXS	-9.82	<0.01	-9.15	<0.01	-3.88	<0.01	-2.92	<0.01
SAND	1.43	0.47	2.68	0.61	0.15	0.18	-3.07	<0.01
ENJ	-0.37	0.23	-0.79	0.19	-3.54	<0.01	-4.66	<0.01
CHZ	-5.56	<0.01	-6.3	<0.01	-6.21	<0.01	-7.49	<0.01

Forecast horizon h	1		5		10		22	
	DM statistic	p-value	DM statistic	p-value	DM statistic	p-value	DM statistic	p-value
Panel B : Post-EIP-1559								
ETH	-9.38	<0.01	-7.05	<0.01	-8.23	<0.01	-5.38	<0.01
MANA	-1.2	0.09	-0.32	0.11	1.47	0.69	0.53	0.43
AXS	0.84	0.28	0.59	0.58	0.83	0.6	0.08	0.19
SAND	-0.93	0.12	-0.72	0.22	-0.16	0.09	-0.11	0.13
ENJ	-0.11	0.14	0.28	0.41	0.9	0.61	-0.82	0.11
CHZ	-5.37	<0.01	-4.84	<0.01	-4.51	<0.01	-3.06	<0.01

Note. This table reports the DM test result for Table 5.7. The competing models are STG-Spillover models trained with the dataset containing gas return and not containing gas return.

20 tokens is affected by gas returns since EIP-1559 made gas prices more stable and directly affected by Ethereum network congestion. This trend is also confirmed by the volatility forecasting approach using STG-Spillover model. Including gas return into the volatility spillover index graph of ETH and ERC-20 tokens had increased the forecasting performance of the assets that are affected by gas return.

Furthermore, based on the above findings about relationship between Ethereum gas and Ethereum-based tokens, we constructed the two factor models having cryptocurrency market return and Ethereum gas return as factors. In the same line with relationship analysis, we constructed the factor model each in pre-EIP-1559 and post-EIP-1559 period. The empirical result show that the performance of factor model evaluated by the R square and absolute value of intercept increases after EIP-1559 adoption for Ethereum and ERC-20 tokens. From the result, we confirm that the EIP-1559 has made Ethereum gas price more stable and predictable, and thus it can explain the excess return of Ethereum-based tokens better than before EIP-1559 adoption. In addition, the performance of two-factor model for ERC-721 tokens in collectible field decreases after EIP-1559 adoption. Even though there is no clear evidence of decrease in explanation power of gas return on ERC-721 token returns, we guess that it can be explained as there exists no explanation power or dependency between Ethereum gas return and ERC-721 tokens return.

Chapter 6

Conclusion

6.1 Contributions

Research on constructing a model that explains the excess return of assets enables measuring the efficiency of the market. Furthermore, the asset pricing model can explain how the risk of an asset is composed and how much each asset is compensated for the following risk. Among various approaches for asset pricing studies, we concentrated on the factor model. The factor model explains the expected excess return of assets with the factors and beta coefficients. Factor models can be largely divided into macro factor models, fundamental factor models, and latent factor models. Each model is classified according to which variables to estimate the factor and beta, of which this dissertation focuses mainly on latent factor models. While conventional latent factors models utilize a large set of firm characteristics to estimate latent factors, most models cannot reflect the network structure of assets. Connected dynamics of asset returns implies that asset returns affect each other, which implies that pricing of certain asset return should be done with other asset returns in a graph-based manner.

This dissertation aims to develop an AI-based empirical asset pricing model that can reflect the connectedness between assets. First, we developed the graph neu-

ral network-based multi-factor asset pricing model for U.S individual stocks, which estimated the connectedness between assets as the Pearson correlation coefficients of excess returns. Also, we developed the realized volatility prediction model using volatility spillover index and spatial-temporal graph neural network because volatility is one of the major components for analyzing the nature of asset return. Based on previous approaches, we constructed the graph neural network-based asset pricing model with time-varying volatility prediction. Finally, we showed that the connected assets can have common characteristics and they can be used as the observable factor of the macroeconomic factor model as an application in the cryptocurrency market. It is done in cryptocurrency market where certain tokens are issued on the same blockchain, and thus the gas price of that blockchain can become the macroeconomic factor. The detailed contribution of each work is explained as follows:

1. We proposed the graph-based multi-factor asset pricing model for U.S individual stocks. The proposed model directly reflects the connectedness between assets through the graph neural network-based structure. Graph neural network can solve the high dimension problem of firm characteristics. The adjacency matrix of assets, which becomes the definition of connectedness of assets as the input of graph neural network, is estimated as the Pearson correlation coefficients of excess returns with cutoff value. The empirical results show that the proposed model effectively estimated beta coefficients and factors both in terms of statistical performance and economic value of factors.

2. We proposed the volatility prediction model with a spatial-temporal graph neural network and volatility spillover index. Even though there exist many

evidence that the volatility of assets has a close relationship, conventional volatility prediction models cannot directly reflect the measured relationship because of their AR-based structure. We utilized the spatial-temporal graph neural network that can take pre-defined relations as input for volatility prediction of global market indices' volatility. We estimated the relationship between volatility by the pairwise volatility spillover index. The empirical results show that the proposed model outperforms benchmark models in short- and mid-term forecasting. In addition, a market with a high spillover effect on other markets can highly help making a prediction on other markets.

3. We proposed the neural network-based asset pricing model that allows time-varying volatility. To relax the time-unvarying volatility constraint of conventional factor models, we proposed the model that predicts the time-varying realized volatility using LSTM and constructed the training loss as the mean square error between true and estimated excess return divided by predicted volatility. The asset pricing part of the proposed model is constructed as the graph-based multi-factor model to reflect the connected structure of assets. The empirical result on U.S individual stocks shows that the statistical performance of the proposed model highly increases in a low volatility period and constructed factors can estimate efficient stochastic discount factor.

4. We investigated the common factor that connected assets in the cryptocurrency market can have. Ethereum-based tokens can be clearly identified as connected assets since they exist and be traded on the same blockchain. We showed that the Ethereum gas price, which is the transaction cost on the

Ethereum blockchain, Granger-causes the Ethereum and ERC-20 tokens returns after the EIP-1559 adoption, where EIP-1559 is the protocol that aims to make the gas price more structured and predictable. Furthermore, we showed that after EIP-1559, Ethereum gas return can act as a factor of the two-factor model with the market return.

6.2 Future Work

This dissertation constructed an asset pricing model that reflects the graph structure of assets, but it leaves additional research challenges in the estimation of the graph structure of assets. In this dissertation, we used Pearson correlation coefficients and volatility spillover index as the estimation for the adjacency matrix of the graph consisting of assets. The used methods are based on return or volatility, which are both primarily derived from the asset price. However, considering that various firm characteristics are used when constructing an asset pricing model to explain asset returns, there is ample room for additional use of firm characteristics other than price when configuring the graph of assets. The extended research topic can consider the more complex model for graph estimation using multi-dimensional variables. Moreover, deep learning can be used for graph estimation based on the asset pricing model. Kipf et al. (2018) showed the possibility of inferring the graph from the interactive data using a graph neural network. Therefore, assuming that asset returns are data that fully reflect dynamics between them, the development of the model in the form of an autoencoder that estimates graph structure with separate graph neural networks and then estimates it through graph neural networks for asset pricing will also have great value as a future research topic.

Bibliography

Abarbanell, J.S., Bushee, B.J., 1998. Abnormal returns to a fundamental analysis strategy. Accounting Review , 19–45.

Abraham, J., Higdon, D., Nelson, J., Ibarra, J., 2018. Cryptocurrency price prediction using tweet volumes and sentiment analysis. SMU Data Science Review 1, 1.

Ali, A., Hwang, L.S., Trombley, M.A., 2003. Arbitrage risk and the book-to-market anomaly. Journal of Financial Economics 69, 355–373.

Almeida, H., Campello, M., 2007. Financial constraints, asset tangibility, and corporate investment. The Review of Financial Studies 20, 1429–1460.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of financial markets 5, 31–56.

Amihud, Y., Mendelson, H., 1989. The effects of beta, bid-ask spread, residual risk, and size on stock returns. The Journal of Finance 44, 479–486.

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International economic review , 885–905.

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: Including jump

components in the measurement, modeling, and forecasting of return volatility. The review of economics and statistics 89, 701–720.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71, 579–625.

Anderson, C.W., Garcia-Feijoo, L., 2006. Empirical evidence on capital investment, growth options, and security returns. The Journal of Finance 61, 171–194.

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. The journal of finance 61, 259–299.

Asness, C.S., Porter, R.B., Stevens, R.L., 2000. Predicting stock returns using industry-relative firm characteristics. Available at SSRN 213872 .

Aysan, A.F., Demir, E., Gozgor, G., Lau, C.K.M., 2019. Effects of the geopolitical risks on bitcoin returns and volatility. Research in International Business and Finance 47, 511–518.

Back, K., 2010. Asset pricing and portfolio choice theory. Oxford University Press.

Baillie, R.T., Bollerslev, T., 1991. Intra-day and inter-market volatility in foreign exchange rates. The Review of Economic Studies 58, 565–585.

Balakrishnan, K., Bartov, E., Faurel, L., 2010. Post loss/profit announcement drift. Journal of Accounting and Economics 50, 20–41.

Bali, T.G., Cakici, N., Whitelaw, R.F., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. Journal of financial economics 99, 427–446.

Ball, C.A., Roma, A., 1994. Stochastic volatility option pricing. Journal of Financial and Quantitative Analysis 29, 589–607.

Bandyopadhyay, S.P., Huang, A.G., Wirjanto, T.S., 2010. The accrual volatility anomaly. Unpublished Manuscript, University of Waterloo .

Bank, M., Insam, F., 2019. Risk premium contributions of the fama and french mimicking factors. Finance Research Letters 29, 347–356.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. The journal of Finance 59, 1481–1509.

Banz, R.W., 1981. The relationship between return and market value of common stocks. Journal of financial economics 9, 3–18.

Barbee Jr, W.C., Mukherji, S., Raines, G.A., 1996. Do sales–price and debt–equity explain stock returns better than book–market and firm size? Financial Analysts Journal 52, 56–60.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64, 253–280.

Barth, M.E., Elliott, J.A., Finn, M.W., 1999. Market rewards associated with patterns of increasing earnings. Journal of Accounting Research 37, 387–413.

Basu, S., 1977. Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. The journal of Finance 32, 663–682.

Baur, D.G., Dimpfl, T., 2018. The asymmetric return-volatility relationship of commodity prices. Energy Economics 76, 378–387.

Belo, F., Lin, X., Bazdresch, S., 2014. Labor hiring, investment, and stock return predictability in the cross section. Journal of Political Economy 122, 129–177.

Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., 2006. Greedy layer-wise training of deep networks. Advances in neural information processing systems 19.

Beran, J., 1995. Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. Journal of the Royal Statistical Society: Series B (Methodological) 57, 659–672.

Berument, H., Doğan, N., 2011. Stock market return and volatility relationship: Monday effect. International Journal of Economic Perspectives 5, 175–185.

Bhandari, L.C., 1988. Debt/equity ratio and expected common stock returns: Empirical evidence. The journal of finance 43, 507–528.

Bhar, R., Hamori, S., 2003. New evidence of linkages among g7 stock markets. Finance Letters 1.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of econometrics 31, 307–327.

Bollerslev, T., 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. The review of economics and statistics , 542–547.

Bollerslev, T., Engle, R.F., Wooldridge, J.M., 1988. A capital asset pricing model with time-varying covariances. Journal of political Economy 96, 116–131.

Bouchaud, J.P., Potters, M., 2003. Theory of financial risk and derivative pricing: from statistical physics to risk management. Cambridge university press.

Brandt, M.W., Kishore, R., Santa-Clara, P., Venkatachalam, M., 2008. Earnings announcements are full of surprises. SSRN eLibrary .

Brown, D.P., Rowe, B., 2007. The productivity premium in equity returns. Available at SSRN 993467 .

Bubák, V., Kočenda, E., Žikeš, F., 2011. Volatility transmission in emerging european foreign exchange markets. Journal of Banking & Finance 35, 2829–2841.

Buterin, V., Conner, E., Dudley, R., Slipper, M., Norden, I., Bakhta, A., 2019. Eip-1559: Fee market change for eth 1.0 chain. Published online .

Cai, Y., Ge, L., Liu, J., Cai, J., Cham, T.J., Yuan, J., Thalmann, N.M., 2019. Exploiting spatial-temporal relationships for 3d pose estimation via graph convolutional networks, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 2272–2281.

Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of political Economy 107, 205–251.

Carhart, M.M., 1997. On persistence in mutual fund performance. The Journal of finance 52, 57–82.

Carnero, M.A., Peña, D., Ruiz, E., 2004. Persistence and kurtosis in garch and stochastic volatility models. Journal of financial econometrics 2, 319–342.

Chamberlain, G., Rothschild, M., 1982. Arbitrage, factor structure, and mean-variance analysis on large asset markets. Technical Report. National Bureau of Economic Research.

Chandrashekar, T., Muralidhara, M., Kashyap, K., Rao, P.R., 2009. Effect of growth restricting factor on grain refinement of aluminum alloys. The International Journal of Advanced Manufacturing Technology 40, 234–241.

Chen, L., Pelger, M., Zhu, J., 2019. Deep learning in asset pricing. Available at SSRN 3350138 .

Chen, L., Zhang, L., 2010. A better three-factor model that explains more anomalies. Journal of Finance 65, 563–595.

Chen, N.F., Roll, R., Ross, S.A., 1986. Economic forces and the stock market. Journal of business , 383–403.

Chen, W., Jiang, M., Zhang, W.G., Chen, Z., 2021a. A novel graph convolutional feature based convolutional neural network for stock trend prediction. Information Sciences 556, 67–94.

Chen, W., Xu, H., Jia, L., Gao, Y., 2021b. Machine learning model for bitcoin exchange rate prediction using economic and technology determinants. International Journal of Forecasting 37, 28–43.

Chen, X., Ghysels, E., 2011. News—good or bad—and its impact on volatility predictions over multiple horizons. The Review of Financial Studies 24, 46–81.

Chen, Y., Wei, Z., Huang, X., 2018. Incorporating corporation relationship via graph convolutional neural networks for stock price prediction, in: Proceedings of the

27th ACM International Conference on Information and Knowledge Management, pp. 1655–1658.

Cheung, Y.W., Lai, K.S., 1995. Lag order and critical values of the augmented dickey–fuller test. Journal of Business & Economic Statistics 13, 277–280.

Chordia, T., Subrahmanyam, A., Anshuman, V.R., 2001. Trading activity and expected stock returns. Journal of financial Economics 59, 3–32.

Choudhury, S., Ghosh, S., Bhattacharya, A., Fernandes, K.J., Tiwari, M.K., 2014. A real time clustering and svm based price-volatility prediction for optimal trading strategy. Neurocomputing 131, 419–426.

Chung, J., Gulcehre, C., Cho, K., Bengio, Y., 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555 .

Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American statistical association 74, 829–836.

Connor, G., Korajczyk, R.A., 1986. Performance measurement with the arbitrage pricing theory: A new framework for analysis. Journal of financial economics 15, 373–394.

Connor, G., Korajczyk, R.A., 1988. Risk and return in an equilibrium apt: Application of a new test methodology. Journal of Financial Economics (JFE) 21.

Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. the Journal of Finance 63, 1609–1651.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7, 174–196.

Datar, V.T., Naik, N.Y., Radcliffe, R., 1998. Liquidity and stock returns: An alternative test. Journal of financial markets 1, 203–219.

Degiannakis, S., Filis, G., Hassani, H., 2018. Forecasting global stock market implied volatility indices. Journal of Empirical Finance 46, 111–129.

Desai, H., Rajgopal, S., Venkatachalam, M., 2004. Value-glamour and accruals mispricing: One anomaly or two? The Accounting Review 79, 355–385.

Diebold, F.X., Mariano, R.S., 2002. Comparing predictive accuracy. Journal of Business & economic statistics 20, 134–144.

Diebold, F.X., Yilmaz, K., 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. The Economic Journal 119, 158–171.

Diebold, F.X., Yilmaz, K., 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. International Journal of forecasting 28, 57–66.

Doosti, B., Naha, S., Mirbagheri, M., Crandall, D.J., 2020. Hope-net: A graph-based model for hand-object pose estimation, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 6608–6617.

Dowling, M., 2022a. Fertile land: Pricing non-fungible tokens. Finance Research Letters 44, 102096.

Dowling, M., 2022b. Is non-fungible token pricing driven by cryptocurrencies? Finance Research Letters 44, 102097.

Eberhart, A.C., Maxwell, W.F., Siddique, A.R., 2004. An examination of long-term abnormal stock returns and operating performance following r&d increases. The Journal of Finance 59, 623–650.

Eisfeldt, A.L., Papanikolaou, D., 2013. Organization capital and the cross-section of expected returns. The Journal of Finance 68, 1365–1406.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: Journal of the econometric society , 987–1007.

Engle, R.F., 1993. Statistical models for financial volatility. Financial Analysts Journal 49, 72–78.

Engle, R.F., Bollerslev, T., 1986. Modelling the persistence of conditional variances. Econometric reviews 5, 1–50.

Engle, R.F., Kroner, K.F., 1995. Multivariate simultaneous generalized arch. Econometric theory 11, 122–150.

Engle III, R.F., Ito, T., Lin, W.L., 1988. Meteor showers or heat waves? heteroskedastic intra-daily volatility in the foreign exchange market.

Engle III, R.F., Ng, V., 1991. Time-varying volatility and the dynamic behavior of the term structure.

Entriken, W., Shirley, D., Evans, J., Sachs, N., 2018. Eip-721: Erc-721 non-fungible token standard. Ethereum Improvement Proposals .

Fairfield, P.M., Whisenant, J.S., Yohn, T.L., 2003. Accrued earnings and growth: Implications for future profitability and market mispricing. The accounting review 78, 353–371.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. the Journal of Finance 47, 427–465.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics 33, 3–56.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of financial economics 116, 1–22.

Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of political economy 81, 607–636.

Fantazzini, D., Kolodin, N., 2020. Does the hashrate affect the bitcoin price? Journal of Risk and Financial Management 13, 263.

Feng, G., Polson, N., Xu, J., 2020. Deep learning in characteristics-sorted factor models. Available at SSRN 3243683 .

Forbes, K.J., Rigobon, R., 2002. No contagion, only interdependence: measuring stock market comovements. The journal of Finance 57, 2223–2261.

Francis, J., LaFond, R., Olsson, P.M., Schipper, K., 2004. Costs of equity and earnings attributes. The accounting review 79, 967–1010.

Garvey, J.F., Gallagher, L.A., 2012. The realised–implied volatility relationship: Recent empirical evidence from ftse-100 stocks. Journal of Forecasting 31, 639–660.

Geanakoplos, J., Magill, M., Quinzii, M., Dreze, J., 1990. Generic inefficiency of stock market equilibrium when markets are incomplete. Journal of Mathematical Economics 19, 113–151.

Gers, F.A., Schmidhuber, J., Cummins, F., 2000. Learning to forget: Continual prediction with lstm. Neural computation 12, 2451–2471.

Gettleman, E., Marks, J.M., 2006. Acceleration strategies. SSRN Electronic Journal .

Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., 2016. Deep learning. volume 1. MIT press Cambridge.

Graves, A., Liwicki, M., Fernández, S., Bertolami, R., Bunke, H., Schmidhuber, J., 2008. A novel connectionist system for unconstrained handwriting recognition. IEEE transactions on pattern analysis and machine intelligence 31, 855–868.

Gu, S., Kelly, B., Xiu, D., 2020a. Autoencoder asset pricing models. Journal of Econometrics .

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

Guo, R.J., Lev, B., Shi, C., 2006. Explaining the short-and long-term ipo anomalies in the us by r&d. Journal of Business Finance & Accounting 33, 550–579.

Hafzalla, N., Lundholm, R., Matthew Van Winkle, E., 2011. Percent accruals. The Accounting Review 86, 209–236.

Hansen, L.P., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. Journal of political economy 99, 225–262.

Hansen, P.R., Lunde, A., 2005. A forecast comparison of volatility models: does anything beat a garch (1, 1)? Journal of applied econometrics 20, 873–889.

Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79, 453–497.

Harrison, B., Moore, W., 2012. Forecasting stock market volatility in central and eastern european countries. Journal of forecasting 31, 490–503.

Harvey, D., Leybourne, S., Newbold, P., 1997. Testing the equality of prediction mean squared errors. International Journal of forecasting 13, 281–291.

Herskovic, B., 2018. Networks in production: Asset pricing implications. The Journal of Finance 73, 1785–1818.

Herskovic, B., Kelly, B., Lustig, H., Van Nieuwerburgh, S., 2016. The common factor in idiosyncratic volatility: Quantitative asset pricing implications. Journal of Financial Economics 119, 249–283.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural computation 9, 1735–1780.

Holthausen, R.W., Larcker, D.F., 1992. The prediction of stock returns using financial statement information. Journal of accounting and economics 15, 373–411.

Hong, H., Kacperczyk, M., 2009. The price of sin: The effects of social norms on markets. Journal of financial economics 93, 15–36.

Hosking, J.R.M., 1981. Fractional differencing. Biometrika 68, 165–176. URL: http://www.jstor.org/stable/2335817.

Hou, K., Moskowitz, T.J., 2005. Market frictions, price delay, and the cross-section of expected returns. The Review of Financial Studies 18, 981–1020.

Hou, K., Robinson, D.T., 2006. Industry concentration and average stock returns. The Journal of Finance 61, 1927–1956.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. The Review of Financial Studies 28, 650–705.

Hu, Y., Ni, J., Wen, L., 2020. A hybrid deep learning approach by integrating lstm-ann networks with garch model for copper price volatility prediction. Physica A: Statistical Mechanics and its Applications 557, 124907.

Huang, A.G., 2009. The cross section of cashflow volatility and expected stock returns. Journal of Empirical Finance 16, 409–429.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. The Journal of finance 45, 881–898.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of finance 48, 65–91.

Jiang, G., Lee, C., Zhang, Y., 2005. Information uncertainty and expected returns. Review of Accounting Studies 10, 185–221.

Jin, X., 2017. Time-varying return-volatility relation in international stock markets. International Review of Economics & Finance 51, 157–173.

Justiniano, A., Primiceri, G.E., 2008. The time-varying volatility of macroeconomic fluctuations. American Economic Review 98, 604–41.

Kaastra, I., Boyd, M., 1996. Designing a neural network for forecasting financial and economic time series. Neurocomputing 10, 215–236.

Kama, I., 2009. On the market reaction to revenue and earnings surprises. Journal of Business Finance & Accounting 36, 31–50.

Karolyi, G.A., 2001. Why stock return volatility really matters .

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.

Khamsi, M.A., Kirk, W.A., 2011. An introduction to metric spaces and fixed point theory. volume 53. John Wiley & Sons.

Kim, K.H., Kim, T., 2016. Capital asset pricing model: A time-varying volatility approach. Journal of Empirical Finance 37, 268–281.

Kipf, T., Fetaya, E., Wang, K.C., Welling, M., Zemel, R., 2018. Neural relational inference for interacting systems, in: International Conference on Machine Learning, PMLR. pp. 2688–2697.

Kipf, T.N., Welling, M., 2016. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 .

Koop, G., Pesaran, M.H., Potter, S.M., 1996. Impulse response analysis in nonlinear multivariate models. Journal of econometrics 74, 119–147.

Kozak, S., Nagel, S., Santosh, S., 2017. Shrinking the cross section. Technical Report. National Bureau of Economic Research.

Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. The journal of finance 49, 1541–1578.

Lee, S.B., Ohk, K.Y., 1992. Stock index futures listing and structural change in time-varying volatility. The Journal of Futures Markets (1986-1998) 12, 493.

Lerman, A., Livnat, J., Mendenhall, R.R., 2007. Double surprise into higher future returns. Financial Analysts Journal 63, 63–71.

Lettau, M., Pelger, M., 2020a. Estimating latent asset-pricing factors. Journal of Econometrics .

Lettau, M., Pelger, M., 2020b. Factors that fit the time series and cross-section of stock returns. The Review of Financial Studies 33, 2274–2325.

Lev, B., Nissim, D., 2004. Taxable income, future earnings, and equity values. The accounting review 79, 1039–1074.

Lewis, D.J., 2021. Identifying shocks via time-varying volatility. The Review of Economic Studies 88, 3086–3124.

Li, J., 2011. Volatility components, leverage effects, and the return–volatility relations. Journal of Banking & Finance 35, 1530–1540.

Li, Y., Yu, R., Shahabi, C., Liu, Y., 2017. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. arXiv preprint arXiv:1707.01926 .

Liang, C., Wei, Y., Zhang, Y., 2020. Is implied volatility more informative for forecasting realized volatility: An international perspective. Journal of Forecasting 39, 1253–1276.

Lintner, J., 1965. Security prices, risk, and maximal gains from diversification. The journal of finance 20, 587–615.

Lintner, J., 1975. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, in: Stochastic optimization models in finance. Elsevier, pp. 131–155.

Litzenberger, R.H., Ramaswamy, K., 1982. The effects of dividends on common stock prices tax effects or information effects? The Journal of Finance 37, 429–443.

Liu, W., 2006. A liquidity-augmented capital asset pricing model. Journal of financial Economics 82, 631–671.

Liu, Y., 2019. Novel volatility forecasting using deep learning–long short term memory recurrent neural networks. Expert Systems with Applications 132, 99–109.

Malmsten, H., TerÃ, T., et al., 2010. Stylized facts of financial time series and three popular models of volatility. European Journal of pure and applied mathematics 3, 443–477.

Markowitz, H., 1952. The utility of wealth. Journal of political Economy 60, 151–158.

Meddahi, N., 2002. A theoretical comparison between integrated and realized volatility. Journal of Applied Econometrics 17, 479–508.

Meynkhard, A., 2019. Fair market value of bitcoin: Halving effect. Investment Management and Financial Innovations 16, 72–85.

Michaely, R., Thaler, R.H., Womack, K.L., 1995. Price reactions to dividend initiations and omissions: Overreaction or drift? the Journal of Finance 50, 573–608.

Mohanram, P.S., 2005. Separating winners from losers among lowbook-to-market stocks using financial statement analysis. Review of accounting studies 10, 133–170.

Momtaz, P.P., 2021. The pricing and performance of cryptocurrency. The European Journal of Finance 27, 367–380.

Moratis, G., 2021. Quantifying the spillover effect in the cryptocurrency market. Finance Research Letters 38, 101534.

Morgan, J., 1996. Risk metrics–technical document. new york. JP Morgan/Reuters .

Moskowitz, T.J., Grinblatt, M., 1999. Do industries explain momentum? The Journal of finance 54, 1249–1290.

Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V., Von Weizsäcker, J.E., 1997. Volatilities of different time resolutions—analyzing the dynamics of market components. Journal of Empirical Finance 4, 213–239.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the Econometric Society , 347–370.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of financial economics 108, 1–28.

Ou, J.A., Penman, S.H., 1989. Financial statement analysis and the prediction of stock returns. Journal of accounting and economics 11, 295–329.

Ozsoylev, H.N., Walden, J., 2011. Asset pricing in large information networks. Journal of Economic Theory 146, 2252–2280.

Palazzo, B., 2012. Cash holdings, risk, and expected returns. Journal of Financial Economics 104, 162–185.

Patton, A.J., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. Review of Economics and Statistics 97, 683–697.

Pesaran, H.H., Shin, Y., 1998. Generalized impulse response analysis in linear multivariate models. Economics letters 58, 17–29.

Piotroski, J.D., 2000. Value investing: The use of historical financial statement information to separate winners from losers. Journal of Accounting Research , 1–41.

Poggio, T., Mhaskar, H., Rosasco, L., Miranda, B., Liao, Q., 2017. Why and when can deep-but not shallow-networks avoid the curse of dimensionality: a review. International Journal of Automation and Computing 14, 503–519.

Pontiff, J., Woodgate, A., 2008. Share issuance and cross-sectional returns. The Journal of Finance 63, 921–945.

Richardson, S.A., Sloan, R.G., Soliman, M.T., Tuna, I., 2005. Accrual reliability, earnings persistence and stock prices. Journal of accounting and economics 39, 437–485.

Rosenberg, B., Reid, K., Lanstein, R., 1985. Persuasive evidence of market inefficiency. The Journal of Portfolio Management 11, 9–16.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. nature 323, 533–536.

Sanusi, M.S., Ahmad, F., 2016. Modelling oil and gas stock returns using multi factor asset pricing model including oil price exposure. Finance research letters 18, 89–99.

Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G., 2008. The graph neural network model. IEEE transactions on neural networks 20, 61–80.

Seo, Y., Defferrard, M., Vandergheynst, P., Bresson, X., 2018. Structured sequence modeling with graph convolutional recurrent networks, in: International Conference on Neural Information Processing, Springer. pp. 362–373.

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

Shen, D., Urquhart, A., Wang, P., 2020. A three-factor pricing model for cryptocurrencies. Finance Research Letters 34, 101248.

Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? Accounting review , 289–315.

Soliman, M.T., 2008. The use of dupont analysis by market participants. The accounting review 83, 823–853.

Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of econometrics 53, 165–188.

Stone, C.J., 1977. Consistent nonparametric regression. The annals of statistics , 595–620.

Taylor, S.J., 2008. Modelling financial time series. world scientific.

Thomas, J., Zhang, F.X., 2011. Tax expense momentum. Journal of Accounting Research 49, 791–821.

Thomas, J.K., Zhang, H., 2002. Inventory changes and future returns. Review of Accounting Studies 7, 163–187.

Titman, S., Wei, K.J., Xie, F., 2004. Capital investments and stock returns. Journal of financial and Quantitative Analysis 39, 677–700.

Trimborn, S., Härdle, W.K., 2018. Crix an index for cryptocurrencies. Journal of Empirical Finance 49, 107–122.

Tuzel, S., 2010. Corporate real estate holdings and the cross-section of stock returns. The Review of Financial Studies 23, 2268–2302.

Urquhart, A., 2021. Under the hood of the ethereum blockchain. Finance Research Letters , 102628.

Valta, P., 2016. Strategic default, debt structure, and stock returns. Journal of Financial and Quantitative Analysis 51, 197–229.

Vogelsteller, F., Buterin, V., 2015. Eip 20: Erc-20 token standard. Ethereum Improvement Proposals 20.

Wang, Y., Pan, Z., Wu, C., 2017. Time-varying parameter realized volatility models. Journal of Forecasting 36, 566–580.

Wang, Y., Pan, Z., Wu, C., 2018. Volatility spillover from the us to international stock markets: A heterogeneous volatility spillover garch model. Journal of forecasting 37, 385–400.

Wilms, I., Rombouts, J., Croux, C., 2021. Multivariate volatility forecasts for stock market indices. International Journal of Forecasting 37, 484–499.

국문초록

금융 자산은 언제나 리스크에 노출되어 있다. 이 리스크의 크기와, 각 자산이 리스크에 대해 얼마나 보상받는 지를 정확히 측정하는 것은 자산의 특성을 이해하는 데 중요한 문제이다. 자산가격결정모형 (asset pricing model)은 자산의 리스크와 그 보상을 통해서 금융 자산의 수익률을 설명하려 하는 모형이다. 본 연구에서는 여러 자산가격결정모형의 형태 중 팩터 모델에 집중하였다. 팩터 모델은 초과 수익률을 팩터와 베타로 분리해서 설명하는 모델이다. 전통적인 팩터 모델들은 거시 금융 변수나 기업 변수 등을 통하여 팩터와 베타를 추정하는데, 이 때 자산 간의 연결관계를 고려하는 연구는 많이 진행되지 않았다. 금융 자산들은 서로 영향을 주는 관계에 있기 때문에 각각의 수익률 또한 개별적이 아니라 자산 간의 그래프 구조를 고려하며 동시에 평가되어야 한다.

본 논문은 팩터 모델에 자산 간의 연결 구조를 반영하기 위한 인공지능 기반 실증적 자산가격결정모형을 제안한다. 이를 위해 먼저 그래프 인공신경망 (GNN)을 바탕으로 한 멀티 팩터 모델을 개발하였다. 이 때 모델의 구조를 결정하는 것 만큼이나 중요한 것은 자산 간 그래프 구조를 어떻게 정의할 것인가라는 문제이다. GNN은 그 입력 변수로서 잘 정의된 그래프 구조를 요구하지만 자산 간의 연결 구조는 명확하게 정의되지 않았기 때문에, 본 연구에서는 자산 간의 연결성을 피어슨 상관계수를 이용하여 추정하고 이를 특정 임계값을 통해 0과 1로 이진화 시키는 방식을 사용했다. 제안한 모델의 구조는 베타를 추정하는 부분과 팩터를 추정하는 부분으로 나뉘어지는데, 각각 기업 변수와, 수익률을 이용해서 추정한다. 1957년부터 미국에 상장된 주식들을 대상으로 한 실증 실험 결과, 제안한 모델은 설명력과 예측 성능 측면에서 벤치마크 모델들보다 우수한 성능을 보였다. 또한 통계적 성능 이외에도 팩터의 경제적 의미를 측정하는 면에서, 제안한 모델로부터 추정한 팩터가 가장 효율적인 확률적 할인요소 (stochastic

discount factor)를 추정할 수 있다는 점 역시 확인하였다.

자산가격결정모형의 가장 중요한 목적은 수익률이지만, 변동성 또한 금융 자산의 움직임을 설명하는 데 중요한 성질이다. 많은 사전 연구에서 밝혀졌듯 수익률과 변동성 사이에는 상관관계가 존재하기 때문에 변동성은 수익률을 설명하는 요인이 될 수 있다. 자산가격결정모형에서와 마찬가지로 자산들 간의 연결 구조를 고려하는 것은 변동성 예측에서도 성능 향상에 큰 영향을 미칠 수 있다. 변동성 분석에서는 여러 자산의 변동성이 서로 영향을 미치는 것을 스필오버 (spillover)라 부른다. 본 논문에서는 스필오버 효과를 직접적으로 반영하는 변동성 예측 모델을 개발하였다. 제안한 모델은 변동성의 측면에서 자산 간 연결 구조를 변동성 스필오버 지수로 구성한 인접행렬로 정의하며, 모델의 구조로는 시공간적 그래프 인공신경망 (spatial-temporal GNN)를 사용하였다. 글로벌 시장 지수들에 대한 실증 실험을 통해서 제안한 모델은 단기와 중기 변동성 예측에서 벤치마크 모델에 비해 가장 좋은 예측 성능을 보이고, 다른 시장에 큰 영향을 주는 시장을 이용하여 다른 시장에 대한 예측 성능을 크게 높일 수 있음을 보였다.

자산가격결정모형에 변동성을 직접적으로 반영하기 위해서는 모형 내에서 변동성이 어떻게 정의되는가를 먼저 살펴보아야 한다. 변동성은 자산가격결정모형 내에서 잔차의 표준편차로 해석할 수 있다. 그러나 시계열 기반 방법론을 사용하여 추정하는 기존의 자산가격결정모형은 시간에 따라 불변하는 변동성을 가정한다. 본 논문에서는 시간에 따라 변하는 변동성을 예측 모델을 이용하여 추정하고, 이를 팩터 모델의 손실함수에 정규화로 사용함으로써 시간에 따라 변화하는 변동성의 특성을 반영하는 팩터 모델을 제안하였다. 미국 상장 주식에 대한 실증 실험 결과 제안한 모델은 시간 불변 변동성 조건을 완화하지 않은 모델에 비해 변동서이 낮은 시기에서 통계적 성능이 큰 폭으로 상승함을 확인하였다.

현재 무시할 수 없는 규모로 성장한 가상화폐 시장에는 구조적으로 확실하게 연결된 자산이 존재한다. 같은 블록체인 상에 존재하는 토큰들은 해당 블록체인 위에서

발행되고 거래되므로 네트워크 구조 상으로 연결성을 지닌다. 본 연구에서는 앞서 진행된 연구에 대한 응용으로, 명확히 구조적으로 연결된 자산들이 초과 수익률을 설명할 수 있는 측정 가능한 공통된 팩터를 가짐을 보이고자 했다. 연구의 대상을 이더리움 블록체인 상의 토큰들로 제한하여 실증 실험을 진행한 결과, EIP-1559 적용 이후에 이더리움 가스 수익률이 시장 수익률과 함께 토큰의 수익률을 설명할 수 있는 팩터로서 작용함을 보였다. 또한, 이더리움 가스 수익률은 토큰의 변동성에 영향을 주는 요소로, 토큰 변동성 예측에도 도움을 줄 수 있는 요소임을 스필오버 기반 변동성 예측 모델을 통해 확인하였다.

본 논문은 자산 간의 연결성을 고려한 자산가격결정모형을 구성하였으며, 이를 통해서 금융 자산들이 갖는 그래프 구조가 실질적으로 수익률에 영향을 미침을 확인할 수 있다. 이 연구결과는 향후 새로운 금융 시장에 대해서도 적용 가능한 확장성 있는 모델이며, 금융 자산의 평가에 있어 여러 자산을 동시에 상관관계를 고려하며 평가해야 한다는 함의점을 제공하고 있다.

주요어: 자산가격결정모형, 변동성 예측, 그래프 인공신경망

학번: 2017-27701