Graduate School of Mathematics And Statistics
UNIVERSITY OF GLASGOW

Estimation of Risk Interaction Structure in Financial Networks

Fu Xin

Dissertation submitted to

in fulfillment of the requirements for the of

BSc in Statistics

Supervisor: Prof. Philipp Otto

Submission Date: March 26, 2025

Contents

Chapter 1

Introduction

In the stock market, risk is generally associated with the uncertainty of returns, usually measured by volatility. The greater the volatility, the moreIn the stock market, risk is generally associated with the uncertainty of returns, usually measured by volatility. drastic the fluctuations in stock prices, leading to higher risks for investors.

Early research found that financial markets exhibit volatility clustering, meaning that periods of high volatility are often followed by high volatility, while periods of low volatility tend to persist as well(Mandelbrot, 1963) These findings suggest that volatility is not constant but time-dependent, leading researchers to develop statistical models to describe its variation over time. In order to better model this time-dependent structure, Engle (1982) first introduced the ARCH model to describe the time-varying variance of asset returns. Based on this model, Bollerslev (1986) developed the GARCH (Generalized ARCH) model. These models have become commonly used methods for modeling volatility in financial econometrics and are widely applied in stock markets, foreign exchange markets, and option pricing (Engle and Mustafa, 1992). However, traditional ARCH models focus more on the time series behavior of a single asset, assuming that each asset is independent, and find it difficult to reveal the linkage and risk transmission structure between assets.

As market structures become increasingly complex, the mutual influence between assets is becoming increasingly significant. When the volatility of certain assets changes, its impact tends to spread rapidly in the market, thereby affecting the price volatility of other assets. This horizontal diffusion effect of risk is called "network effects" in recent studies, emphasizing the interdependence between assets and the propagation mechanism of market systemic risks. (Billio et al., 2012). However, traditional ARCH models can only capture volatility in the time dimension and cannot estimate the spatial effects of stock interactions in real markets. To achieve this, we need to develop a volatility modeling framework that captures the interplay between temporal and spatial effects and accurately characterizes the risk Interaction mechanism in the financial network.

Against this background, this paper focuses on estimating the risk interaction structure in financial networks. Unlike networks in natural sciences where the structure is often observable, financial networks are typically unknown, sparse, and evolving over time, especially in response to structural events such as policy shifts or market shocks. This poses substantial challenges for volatility modeling and calls for a framework capable of handling both temporal persistence and dynamic cross-sectional dependencies. Specifically, this paper models the stock market as a multi-node financial network, where each node represents the stock of a company, and its volatility is affected by two driving factors: on the one hand, the impact of the stock's own historical volatility, that is, the temporal autoregressive effect; on the other hand, the impact of the volatility of other nodes on it, that is, the spatial risk spillover effect. These two types of dependency structures jointly determine the trend of stock volatility and also form the basis of the risk interaction structure we are trying to characterize. Different from the traditional time series framework, the modeling approach adopted in this paper not only focuses on the continuity of individual behavior, but also emphasizes the dynamic interaction between the various elements in the system.

This paper models this risk mechanism with both temporal and spatial dependence as a spatiotemporal log-ARCH model, aiming to capture both the asset's own volatility memory effect and the risk spillover effect from other assets. Furthermore, we formalize the risk interaction structure as a spatial weight matrix, in which each element describes the degree of influence of one company's stock on the volatility of another company's stock. The core problem is that the network structure is unknown, sparse, and may mutate over time, which places higher demands on the estimation method of the model.

In the field of volatility modeling, Otto et al. (2018) introduced the spatial ARCH (spARCH) model and later conducted a systematic analysis of its theoretical properties in a follow-up study (Otto et al., 2021) to explore its characteristics and properties.In addition, Sato and Matsuda (2017) explored logarithmic transformations in the volatility equation and proposed the spatial log-ARCH (log-spARCH) model to capture spatial dependence in a more stable manner. These models were extended by Sato and Matsuda (2021a) to the spatial log-GARCH model. Building on this foundation, Otto and Schmid (2023) eveloped a unified framework that extends the spARCH model to accommodate a broader class of spatial GARCH settings. More recently, Mattera and Otto (2024) introduced a spatiotemporal log-ARCH model and applied it to stock volatility forecasting, demonstrating its effectiveness in handling spatiotemporal data.

This study builds on the Network log-ARCH model proposed by Mattera and Otto (2024) which is also the spatiotemporal log-ARCH model referenced in this paper (details in Section 2.2). In their study, Mattera and Otto (2024) argue that incorporating asset network relationships in financial markets can improve volatility forecasting, while the log-ARCH transformation helps reduce the impact of highly volatile assets on model stability.

However, the spatiotemporal log-ARCH model faces many challenges when estimating the risk interaction structure in financial markets. Unlike networks with well-defined physical or geographical locations found in natural sciences (Barabási and Albert, 1999; Barthélemy, 2011). the structure of financial networks is unknown. In addition, the time series of stock fluctuations often have structural mutations, which may be triggered by unexpected events, policy changes or macroeconomic factors, leading to market fluctuations. These mutations can lead to dynamic changes in network structure, posing greater challenges to traditional modeling methods. Furthermore, Mattera and Otto (2024) mention that predefined network structures such as EUCL and COR may fail to detect structural breaks, leading to misestimation. Additionally, the AR.PIC matrix, which relies solely on historical volatility estimates, struggles to adapt to dynamic changes in market conditions.

On the other hand, in the stock market, the modeling process often involves a large number of parameters, far exceeding the available observations, making model construction more challenging. In such high-dimensional settings, the spatiotemporal log-ARCH model is prone to overfitting due to excessive parameters, which affects the accurate estimation of the financial network structure.

To address this issue, we draw on the spatiotemporal model proposed by Otto and Steinert (2023) combined with the Two-Stage LASSO estimation method. This approach allows for the simultaneous detection of structural breaks in a spatiotemporal data framework while accurately estimating the spatial weight matrix of variables, which corresponds to the risk interaction structure between stocks as defined in this paper. Specifically, the method uses the adaptive LASSO method in both stages. By adding a regularization term, the adaptive LASSO can set the weights of unimportant node connections to zero. This reduces the dimensionality of the model and solves the situation where the number of nodes is much larger than the number of observations. Then in the first stage, we focus on detecting structural mutations in the return series of individual stocks and identify important time points when market dynamics change. In the second stage, we combine these breakpoints with spatial and temporal dependencies to further estimate the dynamically adjusted network structure. (See 2.1 for details).

In summary, this project aims to combine the spatiotemporal log-ARCH model and the Two-Stage LASSO method to model volatility in the stock market and accurately estimate the risk interaction structure in financial networks. We focus on the estimation of a dynamic, unknown, and sparse risk interaction structure in financial networks, where both temporal and spatial dependencies are present and structural breaks may occur over time.

Finally, the rest of the paper is organized as follows. Chapter 2 introduces the spatiotemporal Log-ARCH model and the two-stage LASSO estimation framework, explaining how they can be integrated to model volatility and estimate risk interaction structures. Chapter 3 describes the dataset, including stock selection, preprocessing, and spatial dependence analysis using the Monte Carlo semivariogram envelope method. Chapter 4 presents the empirical analysis, applying the two-stage LASSO method to detect structural breaks and estimate the spatial weight matrix, followed by the evaluation of stock network interactions. Chapter 5 summarizes the main findings of the study, discusses the model implications, and proposes potential directions for future research.

Chapter 2

Methodology

This chapter introduces the spatiotemporal Log-ARCH model (Mattera and Otto, 2024) to capture the time-varying volatility clustering effect and the risk interaction between different stocks, that is, the temporal and spatial dependence of stock data. It also combines the two-stage LASSO method (Otto and Steinert, 2023) to estimate the structure of the interaction of the stocks, that is, the spatial weight matrix in the financial network.

First of all, we consider a set of n stocks represented by $V = \{s_1, \ldots, s_n\}$. Over time, we observe a process $\{Y_t(s_i) : t = 1, \ldots, T, s_i \in V\}$, which describes the nodal attributes across time. We assume that its volatility follows a spatiotemporal Log-ARCH model, which incorporates both temporal and spatial components along with the conditional volatility constant term h and an error term ϵ (more details are seen in 2.2.2):

$$Y_t(s_i) = \sqrt{h_t(s_i)}\epsilon_t(s_i), \tag{2.1}$$

By applying a logarithmic transformation to the squared returns, we can convert the model into an additive form, allowing us to introduce the spatial component. This transformation enables the model to extend from a ARCH model to a spatiotemporal framework, and define $Y_t^*(s_i) = (\ln Y_t(s_i))_{t=1,\ldots,T;i=1,\ldots,n}$ as the vector of log-squared returns:

$$Y_t^*(s_i) = \ln h_t(s_i) + \ln \epsilon_t^2(s_i), \tag{2.2}$$

$$\ln h_t(s_i) = Y^*_{\text{temporal},t}(s_i) + Y^*_{\text{spatial},t}(s_i). \tag{2.3}$$

This transformation facilitates estimation by allowing the separation of temporal and spatial effects. The temporal dependence is modeled using a Log-ARCH process, while the spatial dependence is captured through the spatiotemporal model introduced by Otto and Steinert (2023), which estimates the weight matrix W to quantify risk spillovers among different stocks. Unlike models that impose a predefined spatial structure, we employ a Two-Stage LASSO approach to adaptively estimate W, enabling the detection of structural breaks and the precise estimation of the evolving spatial risk interaction structure.

This chapter begins by introducing the univariate Log-ARCH process, which serves as a foundation for the subsequent development of the spatiotemporal Log-ARCH model that integrates both temporal and spatial components. Section 2.1 provides a detailed description of the model specification, outlining how the temporal and spatial dynamics are formally captured within this framework. Section 2.2 then describes the estimation procedure using a Two-Stage LASSO approach, which leverages the model's structure to estimate spatial effects. In particular, we highlight how this method enables the detection of structural breaks in the spatial effect and facilitates accurate estimation of spatial risk interactions within financial networks.

2.1 Spatiotemporal Log-ARCH Models

The spatiotemporal log-ARCH model specified in this section constitutes the central modeling framework upon which the subsequent estimation and analysis are based. To motivate this structure, we begin by introducing the univariate log-ARCH model, which describes how conditional volatility at a single node evolves over time based on its own past information. Establishing the temporal component in isolation allows for a clearer extension to the spatiotemporal setting, in which both intra-node dynamics and inter-node dependencies are jointly modeled. The univariate formulation below provides the core building block for the complete specification developed in the following subsection.Building on the univariate formulation, we then extend the model to a full spatiotemporal log-ARCH specification by incorporating spatial dependence across nodes, resulting in a framework capable of capturing both temporal auto regression and spatial auto regression.

2.1.1 Univariate log-ARCH models

Assuming that each time series (such as stocks) is independent and is only affected by its own historical volatility and lagged squared returns, we can use the univariate log-ARCH model to analyze its volatility, which is the basis for studying the risk interaction structure. Specifically, the log-ARCH(p) with the conditional volatility log-$h_t(s_i)$ for the i-th stock can be expressed as a linear combination of lagged squared returns (Sato and Matsuda, 2021b):

$$Y_{\text{temporal},t}(s_i) = \sqrt{h_t(s_i)}\,\varepsilon_t(s_i), \tag{2.4}$$

$$\ln h_t(s_i) = \omega_i + \sum_{p=1}^{P} \gamma_{ip} \ln Y^2_{\text{temporal},t-p}(s_i), \tag{2.5}$$

where $Y_{\text{temporal},t}(s_i)$ is the return of the i-th stock at time t; ω_i and the ARCH parameters γ_{ip}; $\varepsilon_t(s_i)$is a random error term with mean 0 and variance 1 and P is the lag order of the log-ARCH process.

To convert this into linear form we need take the logarithm of the square of $Y_{\text{temporal},t}(s_i)$:

$$\ln Y_{\text{temporal},t}(s_i) = \ln h_t(s_i) + \ln \varepsilon_t^2(s_i), \tag{2.6}$$

By substituting the expression of $\ln h_t(s_i)$, we obtain:

$$\ln Y_{\text{temporal},t}(s_i) = \omega_i + \sum_{p=1}^{P} \gamma_{ip} \ln Y^2_{\text{temporal},t-p}(s_i) + \ln \varepsilon_t^2(s_i). \tag{2.7}$$

This indicates that the logarithm of the squared returns at the current time is jointly determined by the logarithm of the squared returns from lagged periods, the constant term, and the error term.

This transformation allows the spatiotemporal ARCH model to be rewritten as a spatiotemporal autoregressive model, which is compatible with 2-stage LASSO estimation (more details will be discussed in 2.2).

However, existing models still have limitations when studying the risk interaction structure in financial networks. First, these models do not explicitly reflect the intrinsic dependencies formed by connections between nodes (such as stocks) in the network structure, making it challenging to model the complex pathways through which risk propagates within the network. Secondly, as the number of network nodes n increases, the number of model parameters grows quadratically as n^2, leading to high-dimensional challenges and significantly increasing the complexity and computational cost of the model.

To address the above issues, Mattera and Otto (2024) extends the existing model and proposes a dynamic spatiotemporal logarithmic ARCH model to characterize the risk interaction structure between stocks. Given that financial network processes can be represented as multivariate time series, this study combines this feature with spatiotemporal modeling and introduces a novel framework that effectively captures the dependencies among stocks within the network. According to the method proposed by Otto et al. (2024a), by applying a log-squared transformation to the returns in the financial network, the model can effectively capture volatility and its dynamic evolution over time.

Specifically, in this model, the logarithm of the conditional volatility of the i-th stock, $\ln h_t(s_i)$, is influenced by the observations of other stocks in the network, $y_t(s_j)$, where $j = 1, \ldots, n$ and $j \neq i$. Through this modeling approach, when a particular stock exhibits high volatility, such volatility can propagate among other related stocks along the network structure.

2.1.2 Spatiotemporal Log-ARCH Models

This section extends the univariate log-ARCH model to a spatiotemporal log-ARCH framework, following the spatiotemporal ARCH model introduced by Otto et al. (2024b) and combining it with the spatiotemporal model introduced in the Otto and Steinert (2023). This model incorporates temporal and spatial dependencies for volatility modeling and employs a logarithmic representation for the volatility equation.

Similarly to the univariate log-ARCH model, the basic observation process can be expressed as the conditional volatility $h_t(s_i)(i = 1, \ldots, n;\ t = 1, \ldots, T)$ at region s_i and time t is influenced by its own past volatility, lagged squared returns from its neighbors, and instantaneous spatial effects.

The spatiotemporal log-ARCH model is defined as follows:

$$Y^* = \sqrt{\mathbf{h}^*}\epsilon, \tag{2.8}$$

Define $\mathbf{h}^* = (\ln h_t(s_i))_{t=1,\ldots,T;i=1,\ldots,n}$ as the vector of log-squared conditional volatilities $Y^* = \mathrm{vec}(\Upsilon)$ represents the nT-dimensional observation vector obtained by vectorizing the matrix $\Upsilon = (Y_t^*(s_i))_{t=1,\ldots,T;\, i=1,\ldots,n}$, which contains observations across all time points and spatial locations. The network log-ARCH model of order one can be expressed as:

$$\mathbf{h}^* = \underbrace{\omega}_{\text{constant}} + \underbrace{\Gamma Y_{t-1}^*}_{\text{temporal dependence}} + \underbrace{\Phi\tilde{\alpha} + Z\xi}_{\text{spatial dependence}} \tag{2.9}$$

where ω is the constant term, Γ is a diagonal matrix representing the temporal ARCH effects.

Furthermore, substituting the expression for $h_t(s_i)$ into the equation for Y_t^* and $u_t(s_i) = \ln \varepsilon_t^2(s_i) + \epsilon_i$ $u_t = u_t(s_1), u_t(s_2), \ldots, u_t(s_n)^T$ represents the vector of log-square errors. We obtain the following dynamic network log-ARCH model:

$$Y^* = \omega + \Gamma Y_{t-1}^* + \Phi\tilde{\alpha} + Z\xi + u_t, \tag{2.10}$$

However, since the expectation of the log-squared errors $E(\ln \varepsilon_t^2(s_i))$ is typically non-zero, the error term u_t should be adjusted to ensure it has a zero mean. To address this issue, we rewrite the model as:

$$Y^* = \phi_0 + \Phi\tilde{\alpha} + Z\xi + \Gamma Y_{t-1}^* + u_t^*, \tag{2.11}$$

where $\phi_0 = \omega + \mu^*$, with $\mu^*(s_i) = E(\ln \varepsilon_t^2(s_i) + \epsilon_i)$, $\mu^* = (\mu^*(s_1), \mu^*(s_2), \ldots, \mu^*(s_n))^T$. In this formulation, ϕ_0 represents the corrected constant term that incorporates the mean adjustments for the log-squared errors. The error term u_t is now defined as: $u_t^* = (u_t(s_1) - \mu^*(s_1), \ldots, u_t(s_n) - \mu^*(s_n))^T$. The complete specification of the spatiotemporal log-ARCH model introduces both temporal autoregressive components and spatial dependence across nodes. Notably, the spatial component appears in the form of a high-dimensional linear regression, where the conditional log-volatility at each node depends on a potentially large number of neighboring nodes. Since the spatial weight matrix is typically sparse and unknown, direct estimation becomes computationally intensive and prone to overfitting, especially when the number of parameters exceeds the number of observations. To address this challenge, we employ a Two-Stage LASSO framework that exploits the structure of the model to (i) detect structural breaks in the volatility process and (ii) perform sparse estimation of the spatial weight matrix. The following section outlines this estimation procedure in detail.

2.2 Two-Stage LASSO Framework

Then, we focus on the spatial component of the model. In this section, we integrate a spatiotemporal framework with a Two-Stage LASSO approach to construct the spatial component of the spatiotemporal Log-ARCH model and accurately estimate the spatial weight matrix.

One of the key challenges in modeling stock market data lies in the high dimensionality of financial networks, where the number of nodes (stocks) often exceeds the number of observations.In addition, the underlying network structure is often unknown, and time series of the stock market often experience structural breaks due to market shocks or policy changes. Traditional methods may have difficulty accurately estimating the spatial weight matrix.

To address these challenges, this study adopts a two-stage LASSO-based estimation framework. following the approach proposed in Otto and Steinert (2023).

The first stage focuses on identifying structural breaks in individual stock-return series. In this stage, structural breaks are identified mainly by estimating significant changes in the mean level α_t. Since the spatial dependence is not time-varying, no information is lost during this step. In addition, by ensuring that the cumulative number of local change points satisfies $\left|\bigcup_{i=1}^{n} T_i^{(1)}\right| < T$, we effectively reduce the parameter complexity of the complete model, thereby achieving computational feasibility. In the second stage, the set of candidate change points filtered in the first stage is incorporated into the spatiotemporal log-ARCH model to directly estimate the spatial dependence matrix W. In this stage, we apply the adaptive LASSO method to efficiently select relevant connections among stocks, enabling the identification of key interactions within the financial network.

2.2.1 Stage 1 - Candidate Change Points

For each stock s_i, we set linear model to capture all the candidate change points T_i, where:

$$Y^*_{\text{spatial}}(s_i) = K\tilde{\alpha}_{s_i} + \epsilon_{s_i} \tag{2.12}$$

where the matrix K is a $T \times T$ lower triangular matrix, defined as $K = \left(\mathbf{1}_{\{j,\ldots,T\}}(i)\right)_{j=1,\ldots,T}$, and $\mathbf{1}_A(x)$ denotes the indicator function on the set A. This matrix aligns the coefficient vector $\tilde{\alpha}_{s_i}$ with the observed time series data, thereby identifying candidate change points through changes in mean levels. The coefficient vector $\tilde{\alpha}_i$ is defined as $\tilde{\alpha}_{s_i} = (\tilde{\alpha}_{t,s_i})_{t=1,\ldots,T}$, Each element $\tilde{\alpha}_{t,i}$ represents the overall mean level change at time t for stock s_i. These coefficients indicate potential structural changes. When a coefficient significantly differs from zero, it suggests a change point at that time.

However, in this linear model setup, there exists a problem where the number of estimated parameters is exactly the same as the number of observations T. Therefore, we follow the strategy proposed by Zou (2006), which employs the group-LASSO approach to estimate structural breaks in autoregressive time series. Building upon this strategy, we introduce the adaptive LASSO penalty. In their work, Zou (2006) proposed the concept of the adaptive LASSO penalty.

This method not only results in less biased estimates of the true nonzero parameters compared to the classical LASSO but also ensures that the correct parameter entries $\tilde{\alpha}_i$ consistently converge to zero as $T \to \infty$. However, standard LASSO methods do not necessarily exhibit these properties. To address this issue, Zou suggested performing a ridge regression for the pre-estimation of $\tilde{\alpha}_i$ before applying the adaptive LASSO penalty. The coefficients $\hat{\alpha}_i^{(*)}$ obtained from the ridge regression step are then introduced as weighting factors ϖ_i in the adaptive LASSO penalty. So, the adaptive LASSO estimates are given by

$$\hat{\tilde{\alpha}}_i \in \arg\min_{\tilde{\alpha}_i} \left(\|K\tilde{\alpha}_i - Y_i^*\|_2^2 + \lambda_a \|\varpi_i \circ \tilde{\alpha}_i\|_1 \right),$$

where the observed process is $y_{\cdot i}$, the L^p -norm is denoted by $\|\cdot\|_p$, and $\circ$ is the Hadamard product. The elements of the weights vectors $\varpi_i = (\varpi_{t,i})_{t=1,\ldots,T}$ are specified as $\varpi_{t,i} = \frac{1}{|\hat{\alpha}_{t,i}^{(*)}|^\gamma}$, where $\gamma > 0$ is a tuning parameter influencing the threshold function.

2.2.2 Stage 2 - Estimation of the spatial weight matrix

The second stage aims to estimate the sparse spatial weight matrix W. To achieve this, we once again employ a LASSO approach, further refining the network structure by integrating the structural breaks detected during the first stage. And the spatiotemporal log-ARCH model is rewritten to account for both temporal and spatial dependencies(chapter 2.1.2):

$$Y_{\text{spatial}}^* = \Phi\tilde{\alpha} + Z\xi + \epsilon \tag{2.13}$$

In the second-stage model, $Y_{\text{spatial}}^* = \text{vec}(\Upsilon)$ represents the nT-dimensional observation vector obtained by vectorizing the matrix $\Upsilon = \left(Y_{\text{spatial}}^*\right)_{t=1,\ldots,T;\, i=1,\ldots,n}$, which contains observations across all time points and spatial locations. The error vector is denoted by $\epsilon = (\epsilon_1, \ldots, \epsilon_T)$, capturing the random noise in the observations at each time point.

The spatial dependence component is represented by the matrix $Z = I_n \otimes Y_{\text{spatial}}^*$, where I_n is an $n \times n$ identity matrix and $\otimes$ denotes the Kronecker product. This structure ensures that the impact of the spatial weight matrix W on the observations is reflected through the term $Z\xi$, where $\xi = \text{vec}(W)$ is the vectorized form of W.

The changes in mean levels are represented by $\Phi\tilde{a}$, where Φ is a lower triangular block diagonal matrix and $\tilde{a}$ is the vector of mean level change coefficients. To ensure the validity of the estimation, we assume that the coefficients $\tilde{a}_i$ are set to zero at time points not identified as candidate change points in the first stage, thereby reducing model complexity while preserving significant change points.

Since the diagonal elements w_{ii} of the spatial dependence matrix W would introduce self-dependence and increase the number of parameters to estimate, we assume $w_{ii} = 0$ to avoid self-dependence and reduce the complexity of the model, thereby improving the accuracy and stability of the estimation.

In addition to the mean-level component $\Psi\tilde{a}$, we incorporate potential spatial dependence through the term $Z\xi$. Here, the coefficients ξ represent the vectorized spatial weights, defined as $\xi = \mathrm{vec}(W)$. Notably, the diagonal elements of W are assumed to be zero, which eliminates self-dependencies and consequently reduces the number of parameters to be estimated in $Z\xi$.

Finally, the coefficients of the full model are estimated via the following optimization problem:

$$(\hat{a}, \hat{\xi})' \in \arg\min_{(\tilde{a},\xi)} \left\| \Phi\tilde{a} + Z\xi - Y^{*}_{\text{spatial}} \right\|_2^2 + \lambda_b \|\xi\|_1, \quad \text{subject to:}$$

$$\begin{aligned} \tilde{a}_{t,i} &= 0 \quad \text{for all } t \notin T_i^{(1)}, \quad 1 \le i \le n; \\ w_{ij} &\ge 0 \quad \text{and} \quad w_{ii} = 0 \quad \text{for all } 1 \le i, j \le n; \\ \sum_{j=1}^{n} w_{ij} &\le 1 \quad \text{for all } i = 1, \ldots, n. \end{aligned}$$

To estimate the full model, a constrained LASSO approach is applied. The constraint $\sum_{j=1}^{n} w_{ij} \le 1$ ensures the non-singularity of the matrix $(I - W)$. This constraint becomes particularly important when the spatial process is close to nonstationarity, as spatial dependence weakens in such cases.

Additionally, to guarantee sign-consistency of the model under spatial dependence, the weights w_{ij} are required to be nonnegative. This requirement arises because, in practice, negative spatial dependence is extremely rare, and ordinary least squares estimates generally fail to maintain sign-consistency when spatial dependence exists. Therefore, this nonnegativity constraint on the weights enhances the stability and interpretability of the estimation.

In summary, the spatiotemporal log-ARCH model provides a flexible framework to jointly model temporal and spatial dependencies in stock volatility. The autoregressive term ΓY^{*}_{t-1} accounts for past volatility effects, while the spatial dependence term $Z\xi$ captures cross-sectional risk spillovers. This formulation allows us to decompose volatility interactions into time-based and network-based effects, offering a more comprehensive understanding of systemic risk.

The next chapter turns to the empirical application. We describe the data sources, construct volatility measures, and conduct a spatial dependence analysis to justify the use of the spatiotemporal modeling framework in real-world financial networks.

Chapter 3

Data

Before analyzing the structure of risk interactions in financial markets, we need to analyze whether there are spatial effects between different nodes (i.e. the stocks).To achieve this goal, this chapter introduces the Monte Carlo semi-variogram envelope method. By applying this technique to various stock distance measures—including Euclidean distance, correlation-based distance, and log-ARCH parameter-based distance—we can determine whether significant spatial dependence exists in stock volatility. This analysis lays the foundation for constructing the spatial weight matrix used in subsequent modeling.

3.1 Data Preparation

To study the risk interaction structure in financial markets, this research selects 19 representative stocks from the U.S. financial sector, including banks, investment firms, and payment service providers. These companies span different subindustries, including commercial banking (Bank of America, Goldman Sachs, Citigroup, Truist Financial), investment management (Blackstone, Ameriprise, KKR, T. Rowe Price), and payment services (Visa, Mastercard, PayPal). This selection enables a comprehensive analysis of risk transmission within the financial industry while also capturing risk interactions among institutions with different business models. The table 3.1 summarizes the selected financial stocks.

The data for this study come from the public financial data platform Yahoo Finance. The daily closing price data of the selected stocks covers the period 2022-01-01 to 2025-01-01 and are downloaded through the R quantmod package. Subsequently, the raw price data is converted into logarithmic returns and the calculation formula is as follows:

$$\text{Log Return}_t = \log(\text{Price}_t) - \log(\text{Price}_{t-1})$$

.

In addition, the squared logarithmic return is calculated to analyze the interactive characteristics of volatility and risk. The calculation formula is:

$$\text{Log-Squared Log Return}_t = \log\left((\text{Log Return}_t)^2\right)$$

Table 3.1: Companies and Their Sub-Sectors in the Financial Industry

Sub-Sector	Companies
Commercial Banking	Bank of America Corporation, The Goldman Sachs Group, Truist Financial Corporation, Regions Financial Corporation, Huntington Bancshares Incorporated, M&T Bank Corporation, Citizens Financial Group
Investment Management	T. Rowe Price Group, Lazard Ltd, KKR & Co. Inc., Ameriprise Financial, The Blackstone Group Inc.
Insurance	MetLife Inc., Lincoln National Corporation, Cincinnati Financial Corporation
Consumer Finance	Capital One Financial Corporation
Payment Services	Visa Inc., Mastercard Incorporated, PayPal Holdings Inc.

Log returns are used to evaluate daily price changes in stocks and serve as the primary variable for analysis. The log-squared log return, as shown in the equation, represents stock volatility by emphasizing larger fluctuations in returns. This transformation allows for a more effective analysis of risk characteristics in stock data, as it amplifies periods of high volatility while reducing the impact of smaller fluctuations.

3.2 Monte Carlo semi-variogram envelope method

Before constructing the network structure among stocks, a key issue is how to verify whether the similarity between stocks exhibits statistically significant spatial dependence. To address this, Banerjee et al. (2004) introduce the Monte Carlo semi-variogram envelope method, which is used to examine whether spatial dependence exists between different stocks, thereby laying the foundation for the subsequent estimation of risk interaction structures.

3.2.1 Definition of semi-variogram

Spatial dependence measures the degree of similarity between observations that are close in space. A commonly used tool for detecting spatial dependence is the semi-variogram . For given stocks s_i and s_j with volatility series $y(s_i)$ and $y(s_j)$, the semi-variogram is defined as:

$$\gamma(h) = \frac{1}{2}\mathbb{E}\left[(y(s_i) - y(s_j))^2\right] \tag{3.1}$$

where $h = \|s_i - s_j\|$ represents the distance between stocks s_i and s_j. In a financial network, this "distance" can be measured in various ways, such as the Euclidean distance (measuring the difference in stock returns), correlation-based distances, or the volatility dynamic differences based on the log-ARCH model. If the value of $\gamma(h)$ is relatively low for smaller values of h, it indicates that stocks closer in space exhibit higher similarity in terms of returns or volatility, implying the presence of spatial dependence.

3.2.2 Construction of stock distance

Before verifying spatial dependence, it is essential to determine how to measure the similarity between stocks in the feature space, which serves as the basis for constructing the distance h. To construct a reasonable network structure, this study follows three primary similarity measurement approaches proposed by Mattera and Otto (2024):

The most intuitive way to measure similarity is by calculating the Euclidean distance based on the return series of stocks:

$$d_{ij}^{\text{Euclidean}} = \sqrt{\sum_{t=1}^{T} \left(y_t(s_i) - y_t(s_j)\right)^2} \tag{3.2}$$

where $y_t^*(s_i)$ represents the volatility of stock s_i at time t. A smaller Euclidean distance between two stocks indicates a higher similarity in their return series. This metric captures the overall similarity in price dynamics and is particularly suitable for identifying groups of stocks exhibiting similar volatility patterns in the market.

Correlation is a crucial metric for measuring stock similarity. The correlation-based distance is defined as:

$$d_{ij}^{\text{Correlation}} = \sqrt{2(1 - \rho_{ij})} \tag{3.3}$$

where ρ_{ij} is the correlation coefficient between the returns of stocks s_i and s_j over the entire observation period. Stock pairs with high correlation tend to exhibit stronger interactions in financial networks, implying more consistent responses to market shocks.

The third approach measures the dynamic differences in volatility based on the log-ARCH model. For each stock s_i, the log-ARCH model is specified as (2.1,2.2):

$$Y_t(s_i) = \sqrt{h_t(s_i)}\,\varepsilon_t(s_i), \quad \ln h_t(s_i) = \omega_i + \sum_{p=1}^{P} \gamma_{ip} \ln Y_{t-p}^2(s_i),$$

where γ_{ip} denotes the autoregressive coefficient of stock s_i at lag p. The model-based distance is defined as:

$$d_{ij}^{\text{AR.PIC}} = \sqrt{\sum_{p=1}^{P} \left(\gamma_{ip} - \gamma_{jp}\right)^2} \tag{3.4}$$

If two stocks exhibit similar autoregressive parameters across all lags, it suggests that they share analogous response mechanisms in terms of volatility dynamics.

3.2.3 Monte Carlo Envelope Method

After defining the similarity measures among stocks, the next step is to verify whether these similarities exhibit statistically significant spatial dependence. To achieve this, we apply the Monte Carlo envelope for the binned empirical semi-variogram method.

Specifically, the Monte Carlo envelope is constructed through the following steps: At each time point t, we randomly reorder the volatilities of n stocks to generate pseudo datasets $y^{(j)}$, ensuring that the dependence structure at each time remains intact while breaking spatial dependencies:

$$y^{(j)}(t) = \left(y^{(j)}(s_1, t), y^{(j)}(s_2, t), \dots, y^{(j)}(s_n, t)\right)$$

where $y^{(j)}(s_i, t)$ denotes the volatility of stock s_i at time t in the j-th permutation. This process is repeated $j = 1, 2, \dots, 1000$ times to ensure the distribution of the data under the null hypothesis of independence.

for each pseudo-data set $y^{(j)}(t)$, the corresponding empirical semi-variogram $\hat{\gamma}^{(j)}(h_m, t)$ is calculated at each time point t as:

$$\hat{\gamma}^{(j)}(h_m, t) = \frac{1}{2N(h_m)} \sum_{\substack{i,j \\ d_{ij} \approx h_m}} \left(y^{(j)}(s_i, t) - y^{(j)}(s_j, t)\right)^2,$$

where $N(h_m)$ represents the number of stock pairs whose distance falls within the interval h_m.

3.2.4 Constructing the Monte Carlo Envelope

For each distance h_m, the 2.5% and 97.5% quantiles of $\hat{\gamma}^{(j)}(h_m)$ across 1000 permutations are calculated to form the 95% confidence interval:

$$\left[\hat{\gamma}_{5\%}(h_m, t), \hat{\gamma}_{95\%}(h_m, t)\right].$$

If the empirical semi-variogram of the actual data $\hat{\gamma}(h_m, t)$ falls entirely within the 95% confidence interval for all h_m values:

$$\hat{\gamma}(h_m) \in \left[\hat{\gamma}_{2.5\%}(h_m), \hat{\gamma}_{97.5\%}(h_m)\right], \quad \forall m,$$

we can conclude that there is no significant spatial dependence among the stocks under the selected similarity measure. Conversely, if $\hat{\gamma}(h_m)$ exceeds the confidence interval, particularly at smaller h_m values, it suggests that the similarity between the stocks is statistically significant in terms of spatial dependence and should be retained when constructing the network structure.

3.3 Testing for Spatial Dependence

To verify the hypothesis that there is spatial autoregressive dependence among stocks, we adopted the Monte Carlo envelope method described in Section 3.2, combined with semi-variogram analysis. Network modeling was carried out according to three distance measures described in formulas (2.15–2.17): Euclidean distance (EUCL), correlation-based distance (COR), and distance based on the parameters of the log-ARCH model (AR.PIC). Figure 3.1 presents the semi-variogram calculation results under different distance measurement methods, where the top row corresponds to the results based on the Euclidean distance, the bottom left panel represents the correlation-based distance and the bottom right panel corresponds to the AR.PIC distance.

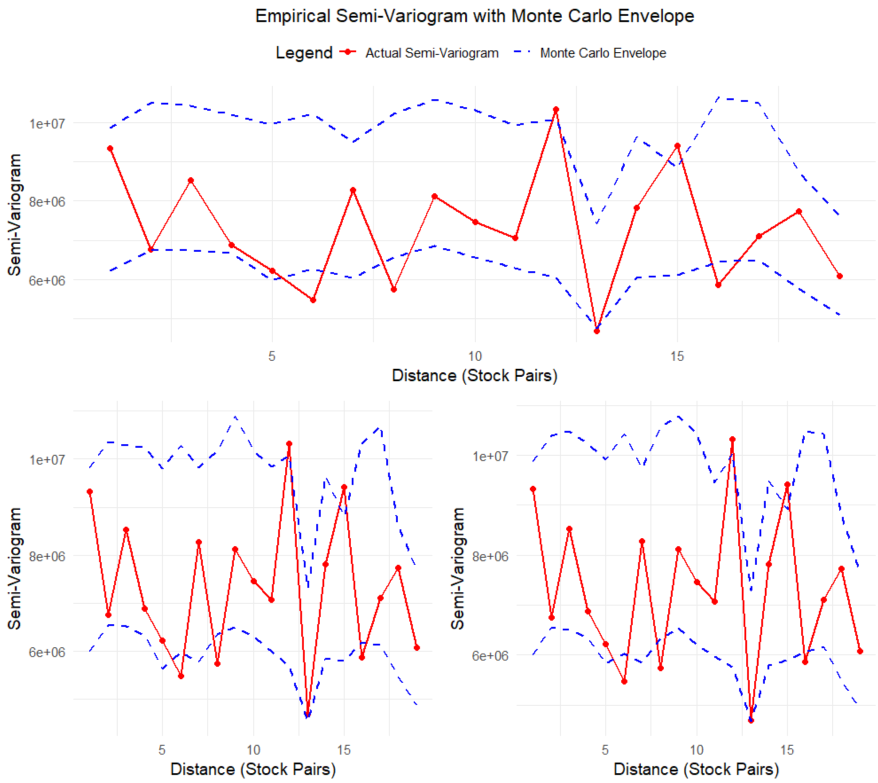

Figure 3.1: Empirical semi-variogram analysis with 95% Monte Carlo envelopes for three distance measures: correlation-based distance (top), Euclidean distance (bottom left), and AR.PIC distance (bottom right). The red solid line shows volatility differences between stocks, while the blue dashed lines represent the confidence interval under the no spatial autocorrelation. Exceedances of the red curve beyond the envelope suggest significant spatial dependence, indicating that volatility in one stock can influence others in the financial network.

The red solid line (actual semi-variogram) in Figure 3.1 represents the volatility differences between different stocks and their changing trend as the distance increases. This curve reflects whether stock volatility shows a spatial autoregressive effect as the distance increases. The blue dashed lines (Monte Carlo Envelope) represent the 95% confidence interval generated through Monte Carlo simulations under the assumption of no spatial autocorrelation, which is used to test the statistical significance of spatial dependence.

The experimental results indicate that, under all three distance measures, the red curve significantly exceeds the boundary of the blue dashed envelope at multiple distance points. This observation suggests the presence of significant spatial dependence in the stock market. Specifically, the volatility of a stock is affected not only by its historical volatility but also by the volatility of other companies' stocks. This spatial dependence means that when certain stocks experience volatility, other stocks that are closer in the financial network are also affected, causing earnings volatility to propagate throughout the market. For example, stocks in the same business often exhibit synchronized movements over short periods of time because they are influenced by similar macroeconomic and market sentiments.

Chapter 4

Analysis

Through the analysis in Chapter 3, we confirmed that there is significant spatial dependence among these 19 stocks. This chapter focuses on applying the spatiotemporal Log-ARCH model combined with the Two-Stage LASSO method to estimate the spatial weight matrix (i.e. risk interaction structure). When applied to real data, the process first eliminates the temporal autoregressive effect, then detects structural mutations, and finally estimates the spatial weight matrix, revealing the intercorrelations of stock market fluctuations. Finally, we verify our findings by analyzing financial events that affect risk transmission within the network.

4.1 Removing Temporal Auto-regressive Effects

Stock returns are typically influenced by the individual market performance of a stock and the risk transmission between different stocks. In industries or market environments with high liquidity and strong information transparency, the stability of stock returns is generally higher than in environments with lower liquidity or insufficient market information.Asparouhova et al. (2010) Additionally, stock return levels often fluctuate over time. These fluctuations may be influenced by external financial events, such as monetary policy, interest rate adjustments, or changes in regulatory policies, which can lead to varying degrees of co-movement in stock returns across multiple companies(Fernández-Amador et al., 2013; Hendershott and Moulton, 2011).

Therefore, in the final section of this paper, we focus on the risk interaction structure among different stocks, specifically the spatial weight matrix of stock data. The data used in this study is sourced from Yahoo Finance, covering stock data of U.S. financial companies from January 1, 2022, to January 1, 2025. The dataset includes a total of T = 752 trading days, with 19 selected stocks analyzed. The ticker symbols of the selected stocks and their main statistical characteristics of volatilities (i.e. Y_t^* in chapter2), including the mean, standard deviation, minimum, and maximum, where Y_t^* represents the log returns, are listed in Table 4.1.

To capture both the temporal and spatial dependencies of stock volatility, we

Table 4.1: Summary Statistics of volatilities of Selected Companies

Company	**Symbol**	**Mean**	**St. Dev.**	**Min**	**Max**
Bank of America	BAC	-9.535	1.937	-13.816	-5.028
Goldman Sachs	GS	-9.609	1.933	-13.815	-4.190
Truist Financial	TFC	-9.081	1.936	-13.816	-3.361
Regions Financial	RF	-9.174	1.956	-13.816	-4.048
Huntington Bancshares	HBAN	-9.309	2.005	-13.816	-3.383
M&T Bank	MTB	-9.280	2.018	-13.816	-3.800
Citizens Financial	CFG	-9.036	2.023	-13.816	-4.009
T. Rowe Price	TROW	-9.221	1.963	-13.816	-3.774
MetLife	MET	-9.699	1.899	-13.816	-5.012
Visa	V	-10.005	1.837	-13.816	-4.590
Mastercard	MA	-9.982	1.942	-13.815	-4.876
PayPal	PYPL	-8.711	2.033	-13.816	-2.530
Lincoln National	LNC	-8.573	1.972	-13.816	-1.819
Capital One	COF	-9.061	1.996	-13.816	-3.935
Cincinnati Financial	CINF	-9.594	1.851	-13.816	-3.797
Lazard	LAZ	-8.965	1.929	-13.816	-3.970
KKR	KKR	-8.880	2.002	-13.816	-4.412
Ameriprise	AMP	-9.600	1.995	-13.815	-4.625
Blackstone	BX	-8.745	2.015	-13.816	-3.931

employ the Spatiotemporal Log-ARCH model as the core framework of this study. The full model is formulated as $Y_t^* = \phi_0 + \Phi\tilde{\alpha} + Z\xi + \Gamma Y_{t-1}^* + u_t^*$, (Equation 2.11)

where ϕ_0 is the adjusted intercept term,$\Phi\tilde{\alpha}$ captures structural breaks in volatility, $Z\xi$ represents spatial dependence via the spatial weight matrix W,ΓY_{t-1}^* accounts for temporal autoregressive effects and u_t^* is the adjusted error term.

This equation provides a comprehensive framework that integrates time-series volatility dynamics and cross-sectional risk interactions among stocks. Then, we first separate the temporal autoregressive effect from the data using a Log-ARCH model before estimating the spatial dependence structure.

The core objective of this study is to estimate the risk interaction structure in the finance networks. To achieve this, we first calculated the daily log returns of these stocks and further applied a log-squared transformation to capture their volatility. This transformation enhances the normality of the data and reduces the impact of extreme values on model estimation.

In this study, we consider $n = 19$ stocks (nodes) as individuals in the financial network and analyze their interactions over $T = 752$ trading days. We consider these stocks as nodes in the financial network to study their risk interaction structure. Since stock returns usually exhibit temporal autocorrelation and spatial autocorrelation effects, We first use the Log-ARCH model (Equation 2.5) to remove temporal dependencies. This ensures that the subsequently estimated spatial weight matrix primarily reflects cross-sectional risk interactions rather than temporal autocorrelation. To assess the instantaneous effects of the

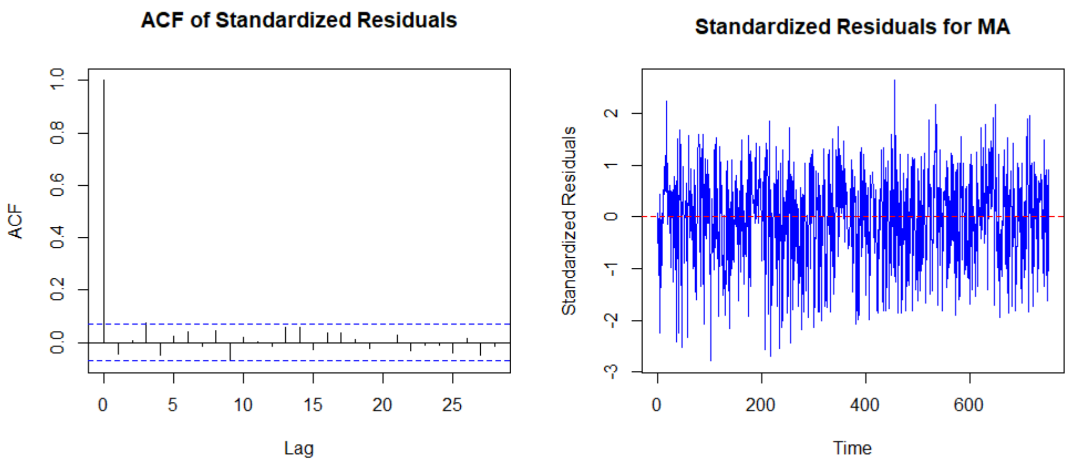

Figure 4.1: Diagnostic plot of the Log-ARCH model for Mastercard Incorporated.The left panel presents the ACF of standardized residuals. The right panel shows the standardized residuals over time.

spatio-temporal ARCH structure, all analyses are based on the same information set, with the time lag order set to $P = 1$. So, the function can be written as $\ln h_t(s_i) = \omega_i + \gamma_{i1} \ln Y^2_{temporal,t-1}(s_i)$ (Equation 2.5), where ω_i is the constant term, and γ_{ip} represents the autoregressive parameters in lag p. This means that only data from the previous trading day are used to calculate the current volatility, thereby capturing the impact of short-term market dynamics. Additionally, we allow each stock to have a different u_t^* (mean parameter) as well as different temporal ARCH parameters. This enables a more precise capture of the information specific to each stock. To evaluate the goodness of fit of the model, we conducted diagnostic analysis on the standardized residuals.

Figure 4.1 shows the ACF plot (left) and standardized residual plot (right) of Mastercard Incorporated, which are used to examine the time dependence and stability of the residuals, respectively. The ACF plot shows that the standardized residuals do not exhibit significant autocorrelation, as most of the values fall within the blue confidence bands. This indicates that the Log-ARCH model effectively eliminates the time autoregressive effect, ensuring that the residuals are approximately white noise. The standardized residual plot on the right further confirms that the residuals remain stable over time, with no obvious pattern or signs of heteroskedasticity. These results indicate that the model has successfully captured and neutralized the volatility clustering effect, making the adjusted return series unaffected by the time dependence at the individual stock level. This adjustment provides a reliable basis for estimating the spatial weight matrix in subsequent analysis, ensuring that the identified financial network structure mainly reflects cross-sectional risk interactions rather than time persistence. In the next section, we will continue to use the two-stage LASSO method to estimate the spatial weight matrix and analyze the risk interaction structure within the financial network.

4.2 LASSO Stage 1: Identifying Change Points

We apply the Spatiotemporal Log-ARCH Model mentioned in Section 2.2 to model the stock data and use the 2-stage LASSO method described in Section 2.3 to estimate the risk interaction structure in two steps, specifically the spatial weight matrix among stocks. In the first stage, we use the adaptive LASSO method to identify potential structural change points (CPs). The model captures change points through a piecewise constant mean process $Y^*_{spatial}(s_i) = K\alpha_{s_i} + \epsilon_{s_i}$, where K is a lower triangular matrix aligning the coefficient vector α_{s_i} with the observed time series. This formulation ensures that structural breaks in the volatility process are detected before estimating the spatial weight matrix.

Figure 4.2 presents the stock volatility (gray lines in the figure) for eight companies and the mean sequences (red lines) estimated in the first stage using the adaptive LASSO method. The gray lines (original data) represent the logarithm of the log-squared returns for each trading day during the sample period. The red lines (mean sequences) represent the piecewise mean estimation results obtained using the adaptive LASSO method. This approach effectively identifies potential structural change points and adjusts the mean levels accordingly, dividing the sample period into several stages, each with a mean level.

In addition, the estimated average level (red curve) will fluctuate greatly due to insufficient data at the beginning and end. Based on this, we introduce an assumption: the mean levels remain unchanged in the initial and final parts of the time series, specifically the first and last 5% of the total sample. In particular, the mean is assumed to be constant during the first 19 days of the sample period (i.e., the first 19 days of 2022) and the last 19 days (i.e., the final 19 days of 2025). This assumption helps reduce estimation bias caused by boundary effects and ensures the accurate identification of potential change points

4.3 LASSO Stage 2: Spatial Matrix Estimation

In the second stage, we estimate the spatial weight matrix W, which corresponds to the parameter $\xi = vec(W)$ in the spatiotemporal log-ARCH model (Equation 2.13): $Y^*_{\text{spatial}} = \Phi\tilde{\alpha} + Z\xi + \epsilon$ where $\xi = vec(W)$ represents the estimated risk spillover structure among stocks. Building upon the candidate structural change points identified in the first stage, we further estimate and incorporate the spatial weight matrix W to capture the risk interaction structure among stocks. The matrix measures the degree of dependence among different stocks, revealing the potential impact on the volatility of other related stocks when the volatility of a particular stock changes.

The blue lines in Figure 4.2 represent the volatility series estimated in the second stage of the model. Compared to the red mean sequences from the first stage, the blue lines retain the piecewise mean level variations in volatility while further incorporating the estimation results of the spatial weight matrix W. Therefore, the blue lines not only reflect the dynamic volatility patterns of individual stocks but also illustrate the risk interactions among different stocks within the financial

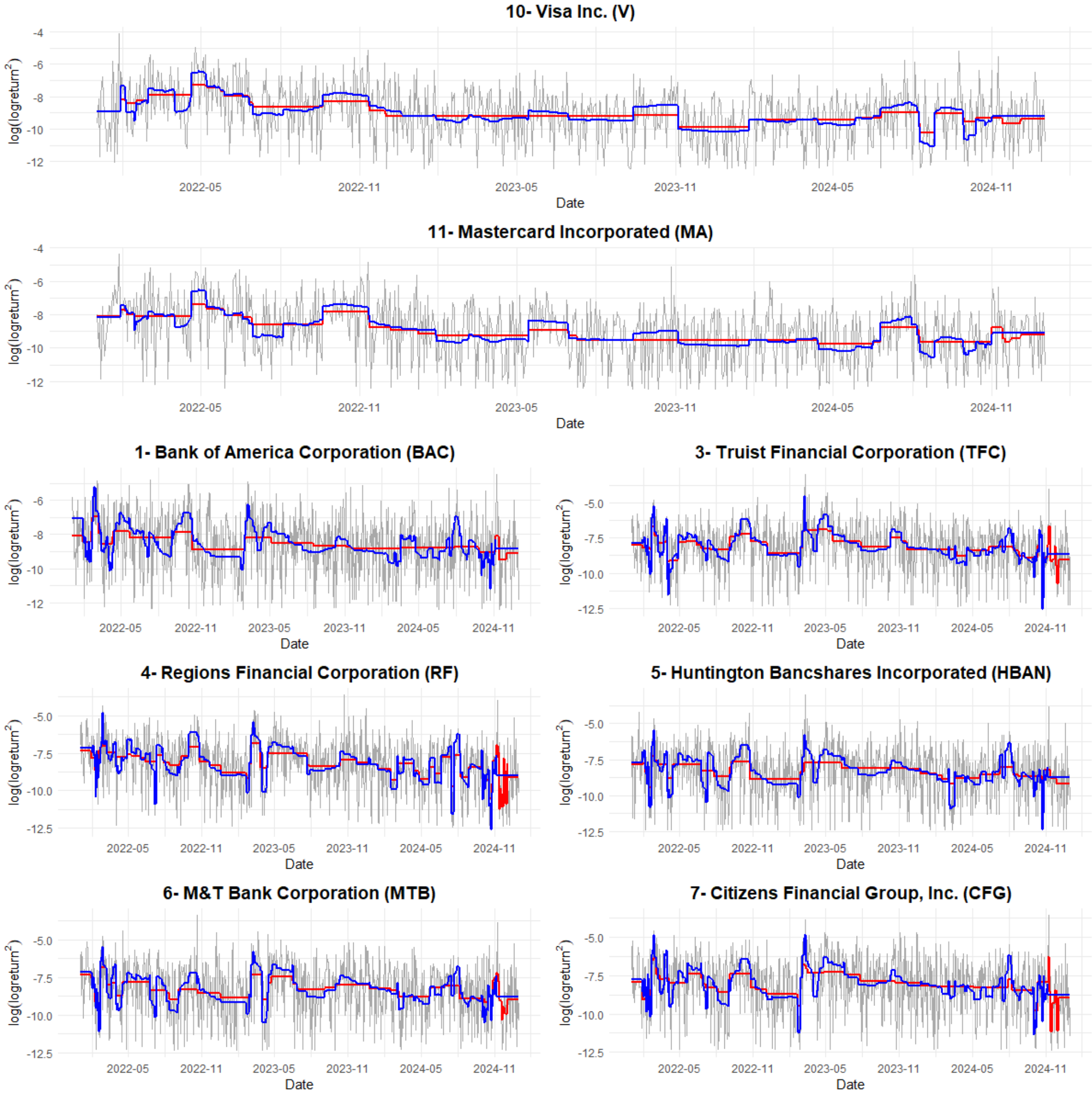

Figure 4.2: Estimation of Volatility for Eight Selected Companies. The figure displays the log-squared return volatility series Y_t^* (gray lines) of eight selected companies, along with the estimated mean sequences $\Phi\tilde{\alpha}$ (red lines) obtained in the first stage using the adaptive LASSO method, and the volatility sequences adjusted by the spatial weight matrix W (blue lines) estimated in the second stage.

network, providing a more comprehensive perspective for understanding the risk interaction structure.

In the second stage, we estimate the spatial weight matrix W, which corresponds to the parameter $\xi = \text{vec}(W)$ in the spatiotemporal log-ARCH model (Equation 2.13):

$$Y^*_{\text{spatial}} = \Phi\tilde{\alpha} + Z\xi + \epsilon$$

where $\xi = \text{vec}(W)$ represents the estimated risk spillover structure among stocks. Building upon the candidate structural change points identified in the first stage, we further estimate and incorporate the spatial weight matrix W to capture the risk interaction structure among stocks. The matrix measures the degree of dependence among different stocks, revealing the potential impact on the volatility of other related stocks when the volatility of a particular stock changes. To determine the optimal regularization parameter λ in the second-stage LASSO estimation, we conduct a cross-validation procedure to evaluate the model's performance under different levels of penalization. Figure 4.3 displays the correlation error across a grid of candidate λ values. The U-shaped curve indicates that both small and large penalties lead to suboptimal model performance. The red point denotes the λ index that minimizes the correlation error, achieving a balance between sparsity and goodness-of-fit.

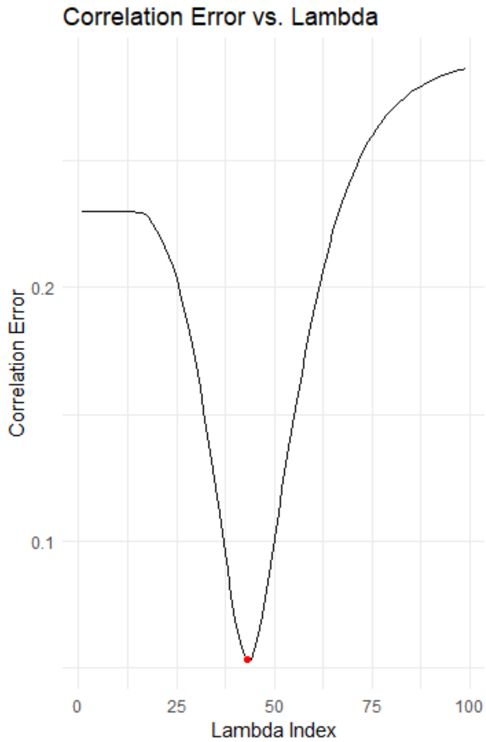

Figure 4.3: Cross-validation curve for LASSO penalty selection in Stage 2. The red dot indicates the lambda index with the lowest correlation error.

Additionally, Figure 4.2 allows for a comparison of the similarity in volatility patterns among different companies. Visa Inc. (V) and Mastercard Incorporated (MA), shown in the first and second rows, exhibit highly consistent phase-based volatility patterns. Over the three-year period, the volatility fluctuations of the two companies are almost synchronized.

The last three rows display the volatility relationships among six financial institutions: Bank of America Corporation (BAC), Truist Financial Corporation (TFC), Regions Financial Corporation (RF), Huntington Bancshares Incorporated (HBAN), M&T Bank Corporation (MTB), and Citizens Financial Group, Inc. (CFG). As shown in Figure 4.2, the volatility of these six banks rises almost

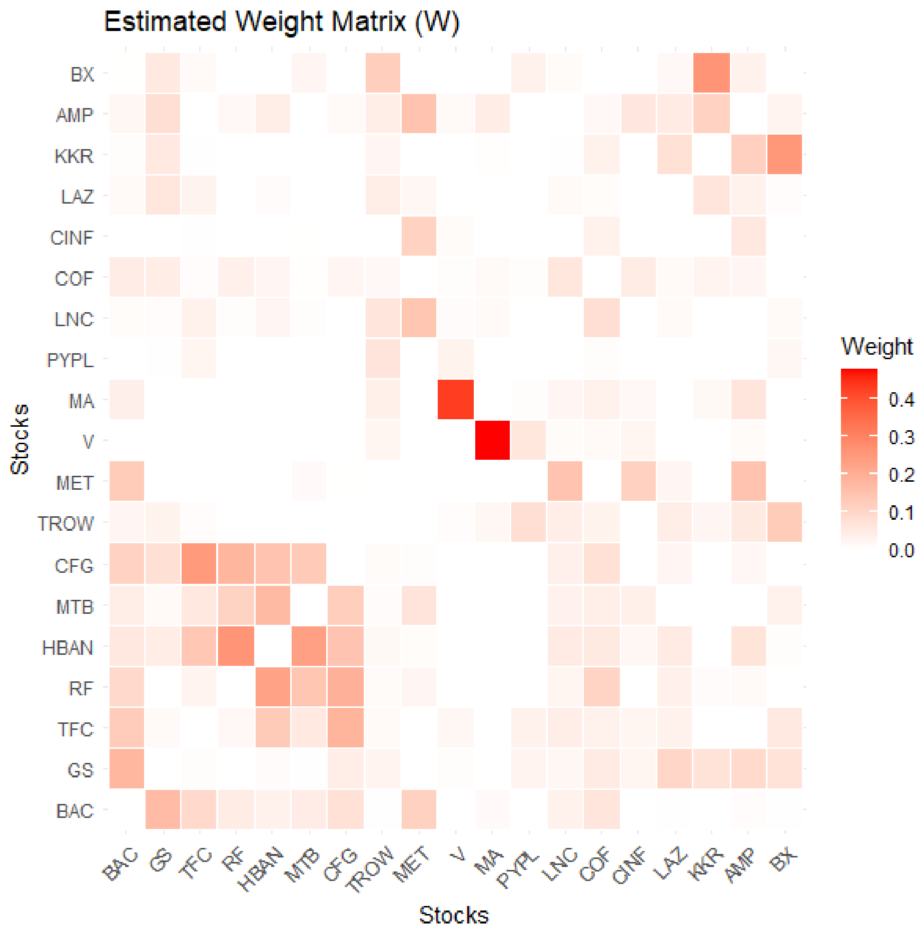

Figure 4.4: Estimated spatial weighting matrix W for the selected financial stocks. The heatmap visualizes the magnitude of risk interaction W_{ij} between each pair of stocks, where W_{ij} represents the estimated spatial dependence in the spatiotemporal log-ARCH model (Equation 2.11). The color gradient from white to red indicates the strength of the interaction, where a deeper red corresponds to a stronger risk interaction.

simultaneously when the market experiences systemic shocks, such as in mid-2023 and early 2024. Additionally, compared to larger nationwide banks such as Bank of America Corporation (BAC) and M&T Bank Corporation (MTB), regional banks like Regions Financial Corporation (RF) and Huntington Bancshares Incorporated (HBAN) exhibit greater sensitivity and larger fluctuations in volatility, as reflected by the blue lines.

Figures 4.4 and 4.5 show the core results of the entire project, the estimated spatial weight matrix (W), which reveals the cross-sectional structure of risk interaction between different stocks. Figure 4.4 is a heat map of W, depicting the strength of risk weights between stocks; Figure 4.5 is a network diagram that visualizes the unknown risk interaction structure and reveals the dynamic risk interaction path between different stocks in the financial market.

In Figure 4.4, the horizontal and vertical axes list names of stocks, with each

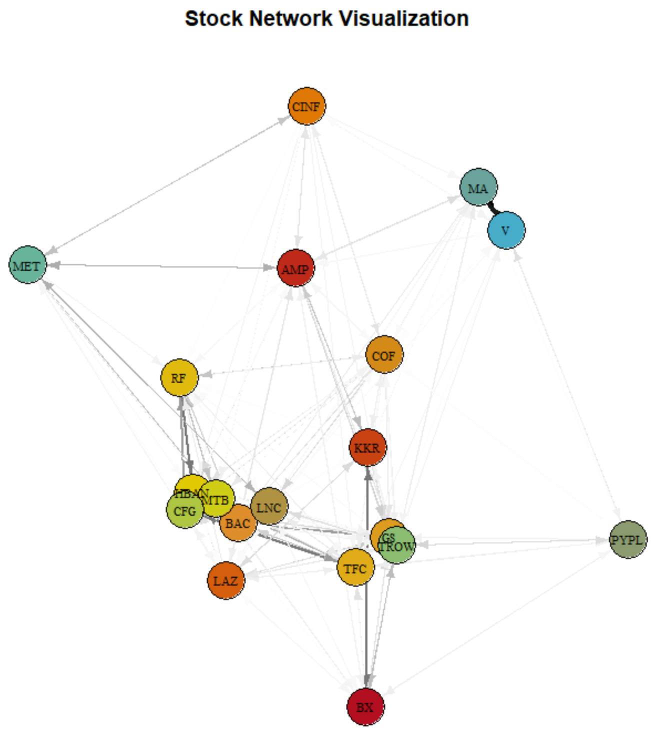

Figure 4.5: Estimated Financial Network of Risk Interactions. The figure visualizes the estimated risk interaction structure among the 19 financial stocks, where each node represents a company, and the edges illustrate how volatility propagates between them. The intensity of the edges reflects the strength of the risk interaction, with darker lines indicating stronger connections. Closer nodes signify higher interdependencies, while more distant nodes indicate weaker relationships.

cell representing the corresponding weight value in the spatial weight matrix obtained through the 2-stage LASSO method. The color gradient from light to dark (white to red) maps the magnitude of the weight values, where a deeper red indicates stronger risk interactions between the corresponding pair of stocks.

In Figure 4.4, the weight cell between Visa (V) and Mastercard (MA) displays the deepest red color, indicating an extremely strong risk interaction between the two companies. This suggests that when one company faces a market shock, the risk can quickly propagate to the other, reflecting the high internal connectivity and lower risk-buffering capacity within the payment industry. This result is consistent with the highly synchronized volatility sequences of the two companies shown in Figure 1, further confirming the synchronous response capability of major participants in the payment industry to systemic risks.

The lower-left corner of Figure 4.4 reveals the complex risk interaction relationships among six financial institutions: Bank of America Corporation (BAC), Truist Financial Corporation (TFC), Regions Financial Corporation (RF), Huntington Bancshares Incorporated (HBAN), M&T Bank Corporation (MTB), and Citizens Financial Group, Inc. (CFG). For the nationwide large bank group (such as BAC and MTB), the weight distribution is relatively dispersed, indicating that these large banks possess stronger risk-buffering capacity when facing systemic risks and are relatively less affected by market shocks. In contrast, for the regional bank group (such as RF, HBAN, and CFG), the weight values are relatively high, suggesting that these banks have closer risk transmission relationships. When regional economic or policy changes trigger risks, the mutual impacts among these banks become more significant.

Figure 4.5 depicts the complete risk interaction structure of these 19 financial companies. Each node represents a different financial company, while the edges illustrate how volatility propagates between them, with the color intensity indicating the strength of the volatility impact. In addition, Figure 4.5 also shows the aggregation effect of risk, where stocks with similar market characteristics are located closer, while stocks with weaker similarities are located farther away.

From an overall perspective, the payment sector (Visa and Mastercard), the banking sector (BAC, TFC, RF, HBAN, MTB, CFG), and investment firms (BX, AMP, KKR) play distinct roles in the structure of risk transmission. Among them, investment companies are the core of this network. They not only connect the banking and payment industries, but also maintain close ties with other financial institutions, which makes them the key disseminators and those affected by market fluctuations. They are like "bridges" in the financial market risk network. Any volatility in any company will affect the investment company and be amplified to the whole financial network.

In addition, the payment sector (Visa and Mastercard) is located in the upper right corner of Figure 4.5. From the figure, we can see that the distance between the two nodes is very close, and the color of the connected edge is very dark, but the connection with the stocks of companies in other industries is not very significant. This indicates that the volatility of these two companies is highly influenced by each other. When Visa experiences a market shock, its risk is quickly transmitted to Mastercard, and vice versa, while their exposure to fluctuations from other industry sectors remains limited. This finding is consistent with the conclusions drawn from 4.4.

Furthermore, Figure 4.5 reveals the risk interaction patterns within the banking sector, particularly among BAC, TFC, RF, HBAN, MTB, and CFG (located in the lower-left corner of the figure). The clustering of these nodes indicates a strong risk transmission mechanism among banks, suggesting that fluctuations within one institution are more likely to propagate throughout the sector. To further illustrate the internal risk interactions within the banking sector, Figure 4.6 (Banking Sector Network) extracts a subnetwork specifically consisting of these bank companies. In this figure, national banks BAC and MTB are positioned at the two ends of the banking network structure, connecting multiple regional banks (CFG, RF, HBAN). Additionally, the edges linking them are light in color, indicating that these national banks are less affected by the volatility of other firms in the financial market. Compared to regional banks, they demonstrate a greater capacity to withstand market shocks, reinforcing their role as more stable entities within the banking network. Additionally, the edge between HBAN and RF is the darkest, indicating a strong mutual dependence between the two banks.

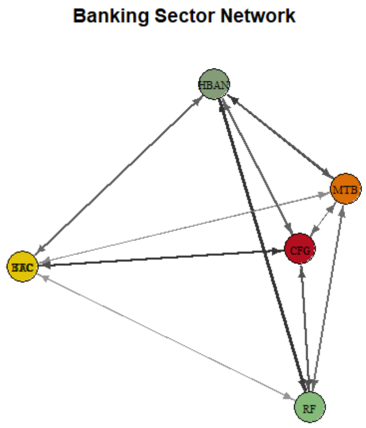

Figure 4.6: Subnetwork of Risk Interactions within the Banking Sector. The network consists of the following banks: Bank of America (BAC), Truist Financial (TFC), Regions Financial (RF), Huntington Bancshares (HBAN), M&T Bank (MTB), and Citizens Financial Group (CFG). The edges represent estimated risk connections between these institutions, illustrating their interdependencies within the sector.

This suggests that when there is a change in the volatility of HBAN, it is likely to affect the volatility of RF. This finding emphasizes that regional banks are more likely to react collectively during periods of market turmoil because they tend to be more interconnected and more sensitive to local economic conditions.

Overall, the weight matrix figure 4.4 and the network graph figure 4.5 complement each other, with the weight matrix quantitatively revealing the intensity of risk interactions between stocks, while the network graph structurally illustrates dynamic risk linkages. Together, they validate the transmission mechanism of systemic risk in the market and highlight the distinct roles that different financial institutions play within this network.

4.4 Event Impact Analysis

Figure 4.7 shows the volatility trends of Visa Inc. and Mastercard Incorporated during two periods: from February to June 2022 and from August to December 2024. The two charts on the left show that from February to June 2022, the stock volatility of Visa Inc. and Mastercard exhibited a continuous upward trend with a high degree of similarity. This phenomenon is primarily attributed to the global financial market uncertainty triggered by the Russia-Ukraine conflict and the decline in cross-border transaction volumes and revenue caused by the withdrawal of both companies from the Russian market. In February 2022, Russia launched a full-scale military operation against Ukraine, causing significant turbulence in global financial markets. Shortly thereafter, Visa and Mastercard announced the suspension of their business operations in Russia. Given that Russia is a key region in the cross-border payments market, the exit from this market led to a sharp decline in cross-border transaction volumes, which directly impacted the revenue

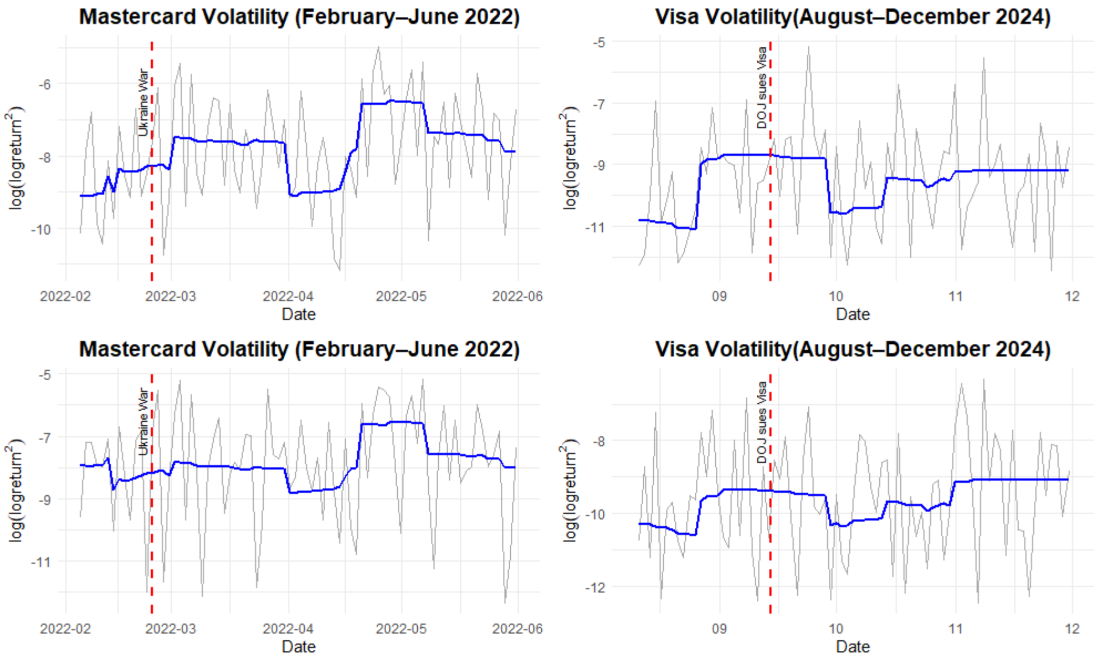

Figure 4.7: Volatility trends of Visa Inc. and Mastercard during two periods: February–June 2022 (left) and August–December 2024 (right).The vertical dashed red lines indicate events: the onset of the Russia-Ukraine War (February 2022) on the left and the U.S. Department of Justice lawsuit against Visa (September 2024) on the right.

streams of both companies. As a result, the combination of increased market volatility and reduced revenue caused the stock volatility of both companies to rise continuously. This also indicates that when facing global market fluctuations, the market's risk assessment of Visa and Mastercard is similar, leading to their stock prices exhibiting volatility almost at the same time.

The two charts on the right of Figure 4.7 show that in September 2024, the stock volatility of Visa Inc. increased significantly, primarily due to an antitrust lawsuit filed by the U.S. Department of Justice. The lawsuit alleged that Visa engaged in monopolistic practices in the U.S. debit card market, harming consumer interests. If the lawsuit is upheld, Visa would not only face substantial fines but might also be forced to restructure its business operations, potentially affecting future revenue growth. This lawsuit introduced uncertainty regarding the market's expectations of Visa's future earnings, which in turn impacted the stock price, causing the volatility to remain elevated during this period.At the same time, the volatility figure of Mastercard in the lower right corner of Figure 4.7 shows that, although Mastercard was not directly affected by the antitrust lawsuit, its stock also exhibited a similar upward volatility trend during the same period. This indicates that the risk faced by Visa not only impacted its own stock but also propagated to Mastercard, affecting its stock volatility. This phenomenon demonstrates the presence of spatial correlation effects between the two companies.

Chapter 5

Conclusion

To estimate the risk interaction structure within the financial network, this paper employs the spatiotemporal log-ARCH model combined with the two-stage LASSO method to model stock data and capture its autoregressive effects in both temporal and spatial dimensions.

Stock volatility exhibits not only temporal autocorrelation but also spatial autocorrelation within the financial network, which cannot be overlooked. However, traditional ARCH/GARCH models primarily focus on the time-series volatility of individual assets and fail to account for the dynamic interactions between different stocks in the financial market. To address this limitation, we introduce the spatiotemporal log-ARCH model to capture the spatiotemporal autoregressive effects of stock volatility.

We consider the stocks of different companies as nodes in the financial network, while the volatility interactions between stocks represent the risk interaction structure. The strength of spatial autocorrelation reflects the degree of risk interaction. Unlike physical or geographical networks, financial networks are generally unknown, as stock connections are not determined by fixed spatial locations. Although tools such as spatial variograms can help identify the presence of spatial dependence, irregular dependencies between different stock volatilities often remain undetected. Furthermore, the challenge of high-dimensional data (i.e., the number of parameters may exceed the number of observations) complicates the estimation process. To address this, many studies use predefined weight matrices to approximate the risk interaction structure (e.g., AR.PIC, EUCL, and COR matrices). However, the stock market is subject to numerous external factors that introduce uncertainty into volatility, such as market crashes or policy adjustments. These external shocks lead to structural changes in the unknown spatial network, affecting both the mean level of volatility at specific locations and across the entire network. Since predefined matrices are often unable to detect structural breakpoints, they may result in estimation bias.

Therefore, we introduced the 2-stage LASSO method. In the first stage, we identified the structural change points and replaced the data in the adjacent change points with the mean level to achieve the effect of reducing the dimension. In the second stage, we estimated the complete model including the spatial weight matrix. With the addition of cross-validation and other procedures, we can ensure the model's fitting performance on real data while sparsely constraining the parameters.

Finally, we illustrate the spatiotemporal log-ARCH model and the 2-stage LASSO method through an empirical example of the volatility of US financial company stocks. The results show that the estimated spatial weight matrix finds that risk interactions are not only affected by industry attributes, but also driven by broader political events, regulatory actions, and investor sentiment. For example, Visa and Mastercard are both in the payment industry, and their volatility trends are very similar; and events such as the Russia-Ukraine conflict and antitrust investigations have also affected the volatility of these two stocks to varying degrees.These findings demonstrate the model's ability to uncover meaningful and time-varying risk transmission structures that would be difficult to identify using static or predefined network assumptions.

In addition, companies of different sizes show different characteristics when dealing with risks. Companies with a wider business coverage tend to be less affected by risk transmission, while smaller companies are more vulnerable. For example, national banks usually have stronger buffering capacity when dealing with systemic risks and are relatively less affected by stock fluctuations of other companies. In contrast, regional banks are more closely connected and the mutual transmission of risks is more significant.

These empirical findings highlight the capacity of the proposed model to capture not only industry-specific risk transmission but also broader systemic patterns driven by macro events and institutional characteristics. The heterogeneity in risk responses across firms of different sizes further illustrates the importance of modeling flexible, data-driven network structures rather than relying on static or assumed relationships.

Against this backdrop, this study advances the literature by proposing a unified spatiotemporal modeling framework that integrates time-varying volatility dynamics with high-dimensional estimation techniques. Through the use of a two-stage adaptive LASSO procedure, the model simultaneously identifies structural breaks in individual return series and recovers a sparse, interpretable spatial weight matrix—thereby uncovering evolving risk interaction structures that are otherwise unobservable. Compared to traditional approaches based on predefined network assumptions, this framework allows for a more realistic, adaptive, and interpretable representation of financial interaction.

Although this study provides a new modeling method and analytical framework, there are still some limitations. First, although our model can capture the spatial dependence in the market, financial networks are inherently complex and are affected by unobservable factors such as investor behavior and market sentiment. Future research can explore more flexible modeling methods, such as multivariate methods and deep learning, to improve the accuracy of risk estimation.

In summary, this study addresses the challenge of modeling dynamic and unknown risk interaction structures in financial networks by integrating the spatiotemporal log-ARCH model with a two-stage LASSO procedure. This approach allows for simultaneous detection of structural changes and sparse estimation of evolving network structures, offering a more accurate and flexible framework for systemic risk analysis. Compared with traditional approaches that rely on fixed or pre-defined network structures, our method adaptively identifies time-varying risk transmission patterns without prior assumptions, providing a data-driven mechanism to capture evolving systemic risks. By identifying market structural change points and accurately estimating dynamic risk interactions, this study provides a new approach to systemic risk assessment in financial markets. In the future, we can further optimize the financial network modeling method and improve our understanding of the risk propagation mechanism in global financial markets.

Bibliography

Asparouhova, E., Bessembinder, H., and Kalcheva, I. (2010). Liquidity biases in asset pricing tests. *Journal of Financial Economics*, 96(2):215–237.

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2004). *Hierarchical modeling and analysis for spatial data*, volume 101;101.;. Chapman & Hall/CRC, London;Boca Raton, FL;.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science (American Association for the Advancement of Science)*, 286(5439):509–512.

Barthélemy, M. (2011). Spatial networks. *Physics reports*, 499(1-3):1–101.

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of financial economics*, 104(3):535–559.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. and Mustafa, C. (1992). Implied ARCH models from options prices. *Journal of econometrics*, 52(1):289–311.

Fernández-Amador, O., Gächter, M., Larch, M., and Peter, G. (2013). Does monetary policy determine stock market liquidity? new evidence from the euro zone. *Journal of Empirical Finance*, 21:54–68.

Hendershott, T. and Moulton, P. C. (2011). Automation, speed, and stock market quality: The nyse's hybrid. *Journal of Financial Markets*, 14(4):568–604.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of business (Chicago, Ill.)*, 36(4):394–419.

Mattera, R. and Otto, P. (2024). Network log-ARCH models for forecasting stock market volatility. *International Journal of Forecasting*, 40:1539–1555.

Otto, P., Doğan, O., and Taşpınar, S. (2024a). Dynamic spatiotemporal ARCH models. *Spatial Economic Analysis*, 19(2):250–271.

Otto, P., Doğan, O., Taşpınar, S., Schmid, W., and Bera, A. K. (2024b). Spatial and spatiotemporal volatility models: A review. *Journal of Economic Surveys*, 38(1):1–55.

Otto, P. and Schmid, W. (2023). A general framework for spatial GARCH models. *Statistical papers (Berlin, Germany)*, 64(5):1721–1747.

Otto, P., Schmid, W., and Garthoff, R. (2018). Generalised spatial and spatiotemporal autoregressive conditional heteroscedasticity. *Spatial Statistics*, 26:125–145.

Otto, P., Schmid, W., and Garthoff, R. (2021). Stochastic properties of spatial and spatiotemporal ARCH models. *Statistical Papers*, 62:623–638.

Otto, P. and Steinert, R. (2023). Estimation of the spatial weighting matrix for spatiotemporal data under the presence of structural breaks. *Journal of Computational and Graphical Statistics*, 32(2):696–711.

Sato, T. and Matsuda, Y. (2017). Spatial autoregressive conditional heteroskedasticity models. *J. Japan Statist. Soc.*, 47(2):221–236.

Sato, T. and Matsuda, Y. (2021a). Spatial extension of generalized autoregressive conditional heteroskedasticity models. *Spatial Economic Analysis*, 16(2):148–160.

Sato, T. and Matsuda, Y. (2021b). Spatial extension of generalized autoregressive conditional heteroskedasticity models. *Spatial Economic Analysis*, 16(2):148–160.

Zou, H. (2006). The adaptive Lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476):1418–1429.

Peer Review Reflection

I appreciate the detailed and helpful feedback I received during the peer review process. These comments were very useful and helped me reflect on the content and structure of my report. I made some specific improvements based on this feedback, making the article more accurate, clear and standardized.

Regarding the structure of the article, one reviewer pointed out that it would be more appropriate to have the data preparation and model sections as independent chapters rather than subsections of the introduction. I completely agree with this observation, as these sections represent different stages of the analysis. Therefore, I modified the structure of the report to reflect this, making each section clearly independent.

Regarding the figures, the reviewer pointed out that there was no reference to Table 1.1 in the text and suggested adding a sentence such as "can be seen in Table 1.1" at the end of the relevant paragraph. This was a very useful reminder, and I have now ensured that all tables and figures are clearly referenced and explained at the appropriate place in the main text. Another issue was that "Table 1.1 is slightly clear showing companies and their sub-sector however I would have companies in the first column to the left and then the sub-sector column with a row for each company as that would make it clear what companies you are looking at and then a column describing what sub-sector they are in. The table isbclearly labelled with an informative caption." After some reflection, I did not take this suggestion. Given the relatively large number of companies involved in the dataset, listing each company individually would make the table too long and difficult to read. Instead, grouping companies under shared sub-industries in the same row would provide a more concise and thematically clear view, better suited to the objectives of this summary table.

The reviewers also highlighted some minor formatting and grammatical issues, such as missing spaces after commas or periods. I carefully proofread the manuscript to address these inconsistencies and improve the professional appearance of the report. In addition, I updated several in-text references to follow the correct reference placement (i.e. before punctuation) and removed redundant reference formatting using both the manual and hyperlinked versions. Information on the selection of US companies and the covariates you used was also added to the data section accordingly.

Overall, the peer review process helped me refine the content and structure of my project, and I appreciate the helpful and specific feedback provided.