Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: www.tandfonline.com/journals/ubes20

Multi-Matrix Autoregressive Models with an Application to Multi-Modal Network

Shiqi Ye, Tingguo Zheng, Han Xiao & Yongmiao Hong

To cite this article: Shiqi Ye, Tingguo Zheng, Han Xiao & Yongmiao Hong (10 Aug 2026): Multi-Matrix Autoregressive Models with an Application to Multi-Modal Network, Journal of Business & Economic Statistics, DOI: 10.1080/07350015.2026.2682541

To link to this article: https://doi.org/10.1080/07350015.2026.2682541

JOURNAL OF BUSINESS & ECONOMIC STATISTICS
https://doi.org/10.1080/07350015.2026.2682541

Multi-Matrix Autoregressive Models with an Application to Multi-Modal Network

Shiqi Ye[a,b], Tingguo Zheng[c,d], Han Xiao[e], and Yongmiao Hong[a,b]

[a]State Key Laboratory of Mathematical Sciences and Center for Forecasting Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China; [b]School of Economics and Management, and MOE Social Science Laboratory of Digital Economic Forecasts and Policy Simulation, University of Chinese Academy of Sciences, Beijing, China; [c]Center for Macroeconomic Research, Xiamen University, Xiamen, China; [d]Department of Statistics and Data Science, School of Economics, Xiamen University, Xiamen, China; [e]Department of Statistics, Rutgers University, New Brunswick, NJ

ABSTRACT
Matrix-valued time-series data have become increasingly prevalent across diverse fields, including economics, finance, computer science, engineering, and signal processing. However, mixed matrix-vector structures in real-world data present challenges for existing methods. To fill in this gap, we propose a multi-matrix autoregressive (MMAR) model designed to jointly model matrix time series with varying structures. The key idea is to treat matrices as elements and to model their contemporaneous and intertemporal relationships within a simultaneous equation system. In particular, the matrix-valued autoregressive model and the three-order tensor autoregressive model are special cases of the proposed model. We present three distinct estimation methods for the MMAR model, investigate their statistical properties, and provide numerical simulations to corroborate them. Furthermore, we integrate the MMAR model with connectedness network analysis, proposing a multi-modal connectedness framework to estimate the interconnectedness between the rows and columns of matrices. Finally, we provide an empirical application to concurrently model the macroeconomic matrix time series of China's 31 provinces and a vector time series comprising the economic policy uncertainty, trade policy uncertainty, and geopolitical risk. The findings of the empirical study provide valuable insights for further research and policymaking in the relevant domains.

ARTICLE HISTORY
Received August 2025
Accepted May 2026

KEYWORDS
Autoregressive model; Connectedness network; Matrix time series

1. Introduction

Matrix-valued data have been extensively used in recent years in various fields such as economics, finance, computer science, engineering, and signal processing. In economics, the matrix data under consideration often possesses an additional time dimension, called matrix time series data. Important matrix time series data include time-varying networks or time-varying panel data. For example, international trade data provides monthly matrix data on imports and exports for different countries and regions, while taxi traffic data offers daily rides between various areas. Additionally, to study the dynamic relationships among different national or regional economies, their different economic indicators (such as GDP, money supply, interest rates, inflation rates, etc.) can be arranged in a matrix format. These types of data often contain a large number of variables and exhibit complex temporal dependencies, posing challenges for statistical and econometric modeling.

A traditional and convenient approach to the analysis of matrix time series data is to stack the matrices in vectors and subsequently model and analyze the resulting vector time series (Lütkepohl 2005). However, this method neither accounts for the inherent structure present in the data nor addresses the issue of overparameterization, which often arises due to an excessive number of model parameters. Existing research has made beneficial attempts to fulfill the demand for matrix time series data analysis. Based on the multi-linear regression specification of Hoff (2015), Chen et al. (2021) propose a matrix-valued autoregressive model in bilinear form, effectively capturing the contemporaneous and intertemporal dependencies between matrix-valued variables. Hsu et al. (2021), Xiao et al. (2022), Ren et al. (2022), Zhu et al. (2023), and Chan and Qi (2023) further investigate MAR estimation under spatial-temporal conditions, low-rank situations, network structures, and Bayesian framework.[1] A series of literature analyzes matrix-valued time series based on factor model specifications or detects and decomposes the correlation structures of matrix variables. Some notable examples include Wang et al. (2019), Chen et al. (2020), Chen et al. (2021), He et al. (2023), Han et al. (2024). Others analyze cross-sectional correlations between variables.[2]

Although the methods discussed are adept at leveraging the inherent matrix structure within the data for purposes such as autoregressive relationship estimation, dimensionality reduction, and covariance matrix estimation, they encounter a fundamental limitation when applied to specific economic problems: their focus is on a singular data structure. In other words, these models fail to examine the relationships, both contemporaneous and intertemporal, between mixed forms of data. This includes understanding the dynamics between matrix time

CONTACT Tingguo Zheng zhengtg@gmail.com Center for Macroeconomic Research, Xiamen University, Xiamen, China.
Supplementary materials for this article are available online. Please go to https://www.tandfonline.com/JBES.

[1]Li and Xiao (2021) extend the approach by proposing multilinear tensor autoregressive (TenAR) models for more generalized tensor time series data, Billio et al. (2023) introduce Bayesian Markov switching tensor regression models, with TenAR being a special case.

[2]Chen and Fan (2023), Chen et al. (2022), Barigozzi et al. (2023), Yu et al. (2025), Armillotta et al. (2025) and others further extend the dynamic factor approach to tensor scenarios. Chen et al. (2022) provides an R package to analyze tensor time series.

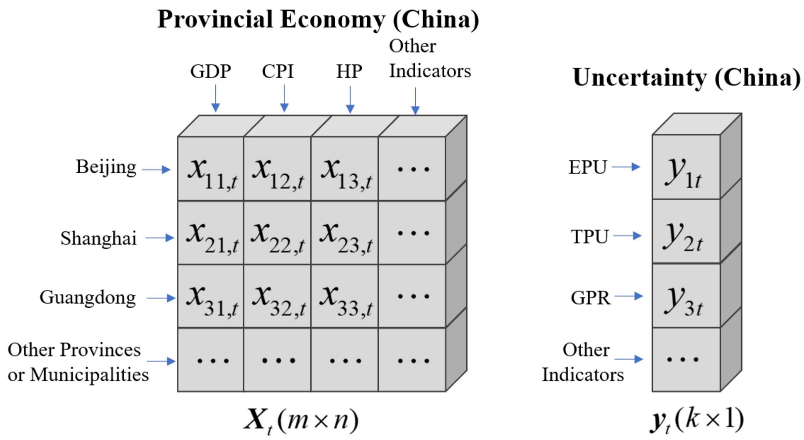


Figure 1. Macroeconomic uncertainties and provincial (state-level) economies.
Notes: The black dots represent the data points. For example, in the provincial economy matrix (X_t), the black dot in the first row and in the first column represents the GDP of Beijing at time t.

series of varying dimensions, as well as between matrix and vector time series. This limitation hinders a comprehensive analysis of the intricate economic interactions represented by these diverse data types. We emphasize that addressing these problems is highly valuable and meaningful and provide three examples to highlight the value of considering these situations in matrix time series data modeling.

Example 1 (Regional economics). Consider a general case using specific economic indicators such as GDP, inflation rate, and other macroeconomic indicators. For demonstration purposes, we take China's provincial economies as a representative example. These economies may also be influenced by various uncertainty indicators such as the economic policy uncertainty index (EPU) by Baker et al. (2016), the trade policy uncertainty index by Baker et al. (2013), and the geopolitical risk index (GPR) by Caldara and Iacoviello (2022). Figure 1 shows the data structure for this case, where the matrix-valued observations $\boldsymbol{X}_t, t = 1, \ldots, T$, collect regional economic indicators (for example, from Chinese provinces), and a corresponding vector $\boldsymbol{y}_t, t = 1, \ldots, T$, gathers uncertainty indicators. At this point, traditional MAR modeling methods are no longer suitable for the above situation, and converting the matrix into a vector by flattening would lose the particular matrix structure present in the regional economy.

Example 2 (International trade). Suppose that we are interested in analyzing global energy trade networks. The available data includes monthly imports and exports of certain energy commodities (such as crude oil, natural gas, and mineral fuels) between countries. Some important economic variables, such as national income, CPI, economic policy, and even some policy uncertainty indicators, can also be considered to study the relationships between economic variables and trade networks. Figure 2 illustrates the matrix structure of the available data, where $\boldsymbol{X}_{1t}, \boldsymbol{X}_{2t}$, and $\boldsymbol{X}_{3t}$ represent the trade networks of crude oil, natural gas, and mineral fuels, respectively. Note that due to the availability of data, the dimensions of $\boldsymbol{X}_{1t}, \boldsymbol{X}_{2t}$, and $\boldsymbol{X}_{3t}$ may not be equal. Moreover, even if we discard some useful information to make the dimensions of the three matrices the same, adding the $k \times 1$ vector $\boldsymbol{y}_t$ makes traditional tensor time series analysis methods difficult to apply.

Example 3 (Stock market connectedness). Let us consider a general model for analyzing relationships between stock markets, using key industry features such as asset returns, market value, liquidity, and trade volume. As a case in point, we examine the stock markets of China and the U.S. This example aids in illustrating how to represent and analyze such issues within a framework of mixed data structures, particularly focusing on the use of matrix time series to capture the nuances and complexities of different stock markets. The data structure is illustrated in Figure 3, with matrix-valued observations representing each market. For example, the matrices $\boldsymbol{X}_{1t}$ and $\boldsymbol{X}_{2t}$ could contain data from the Chinese and US stock markets, respectively. Given the varying industry or sector classifications, the dimensions of these matrices (m_1, n_1 for China and m_2, n_2 for the US) are likely to differ. Furthermore, the vectors $\boldsymbol{y}_{1t}$ and $\boldsymbol{y}_{2t}$ could collect relevant economic indicators from China and the United States, respectively, again with potentially different dimensions (k_1 and k_2). This example serves to illustrate the broader application of the model in contrasting international markets, not limited to China and the US but extendable to other country pairs with distinct economic and stock market characteristics.

The aforementioned examples underscore the potential for mutual correlations between matrix time series of varying dimensions, as well as between matrix and vector time series data. From a statistical perspective, especially in the age of big data, these mixed-data forms are increasingly prevalent, making the characterization of their interrelations critically important. Economically speaking, most economic issues cannot be simplified to modeling problems using a single matrix or tensor time series. Consequently, there is an increasing demand to develop methods capable of modeling the interactions between different matrix time series. This approach is essential for a more nuanced and comprehensive understanding of complex economic phenomena.

A natural and intuitive approach arises when treating matrices as elements within a matrix space. From this perspective, the matrix-valued autoregressive model (MAR) introduced by

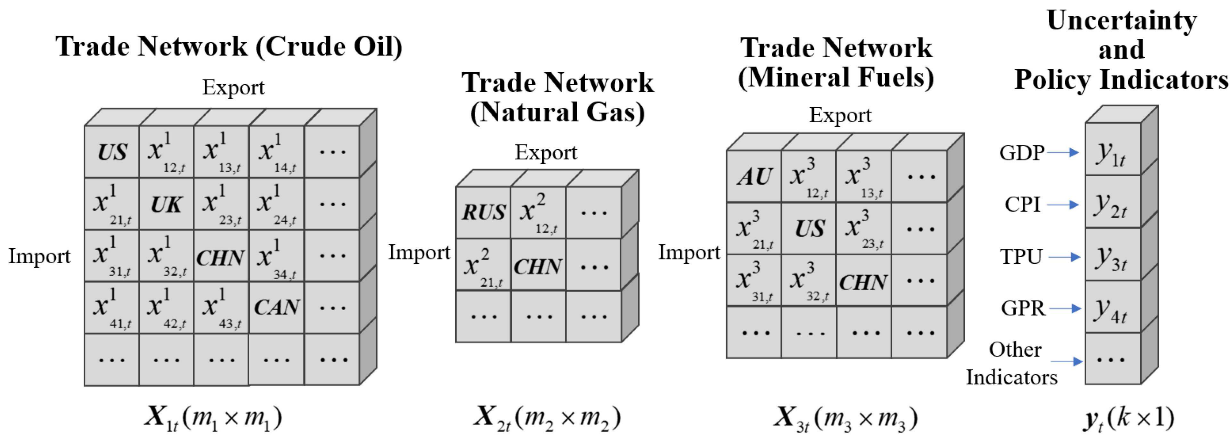


Figure 2. Data structure of international trade network.

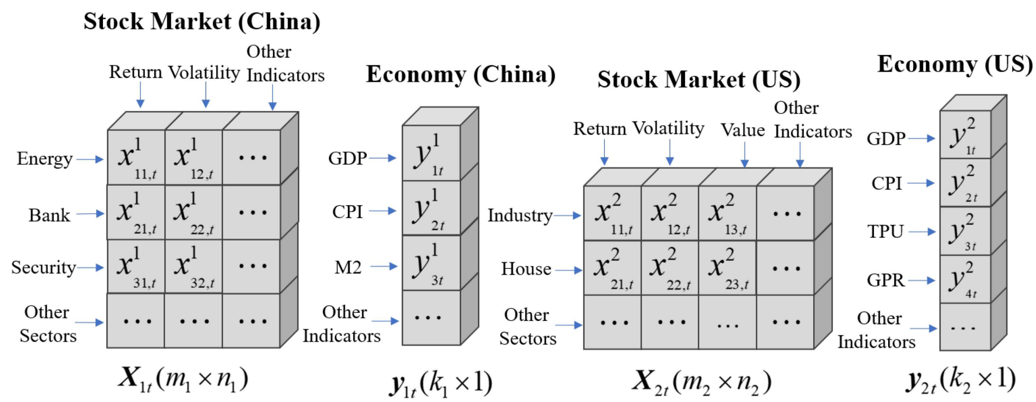


Figure 3. Data structures of Chinese and U.S. stock markets and their macroeconomies.

Chen et al. (2023) can be fundamentally interpreted as an "AR" model for matrices. Notably, Sims (1980) proposed the vector autoregression (VAR) framework, which systematically captures the complex dynamics of multiple time series while offering a user-friendly and interpretable statistical toolkit. Building on these foundations, this paper introduces a novel multi-matrix autoregressive (MMAR) model, extending the capabilities of classical approaches. Specifically, the autoregressive elements or endogenous variables in the model we propose are not restricted to a single matrix. Instead, they can be a composition of multiple matrices or vectors, each potentially differing in dimensions. This flexibility allows our model to effectively capture and characterize the intricate relationships within complex economic systems, as exemplified in the scenarios outlined in Examples 1–3. The model's ability to integrate and analyze these complex structured data sets it apart, offering a more robust and comprehensive framework for understanding the dynamics of multifaceted economic environments.

Compared to the matrix autoregressive (MAR) model of Chen et al. (2021), which corresponds to the special case $K = 1$, the proposed MMAR framework departs in several ways. When $K > 1$, the VAR representation implied by MMAR no longer admits a single Kronecker product structure, but instead takes a block-wise Kronecker form, with each block capturing interactions across different matrices that may have heterogeneous dimensions. This difference raises new theoretical challenges in identification, normalization, and large-sample inference: the parameter space expands to K^2p bilinear blocks, while the limiting covariance becomes block-structured and allows for cross-equation dependence, so that the existing MAR theory and estimation procedures no longer apply directly. To address these challenges, we develop estimation methods tailored to the block-wise Kronecker structure, including block-wise projection estimation, an equation-wise alternating least squares algorithm with reduced computational cost, and an iterated likelihood-based estimator under tractable covariance assumptions, and we establish the asymptotic properties of the resulting estimators. From an applied perspective, the MMAR framework also broadens the scope of empirical analysis by allowing interactions across multiple matrices and matrix-vector systems that cannot be handled within the MAR framework.

Beyond contributions to the model, estimation methods, and theoretical properties, we offer a novel perspective to analyze the connectedness between variables within complex economic systems. Specifically, we extend the Diebold-Yilmaz (DY) connectedness measurement method, as proposed by Diebold and Yılmaz (2014, 2023), to a multi-modal version. The DY connectedness method, based on the VAR framework, illustrates, through generalized forecast error variance decomposition, how much of the future uncertainty associated with the variable i can be attributed to shocks emanating from the variable j. Within the context of the MMAR model, we fully exploit the inherent matrix structure in the data to investigate how much of the future uncertainty associated with the row (column) i can be attributed to the shocks emanating from the row (column) j. This enables researchers not only to focus on the connectedness at the individual variable level but also to delve into multi-modal networks at the row level (such as the provincial level in Example 1, the industry level in Example 2, and the import level in Example 3) and the column level (like the macroeconomic indicators in Example 1, stock indicators in Example 2, and the export level in Example 3). This multi-modal approach opens up new perspectives for analyzing complex systems, enhancing our understanding of their intricate dynamics.

Finally, based on Example 1, we discuss the impacts of economic policy uncertainty, trade policy uncertainty, and geopolitical risk on China's regional economic conditions. Specifically, we construct a 31×7-dimensional matrix time series representing the regional macroeconomic landscape in China and, based on the MMAR model, examine its interrelationships with the aforementioned uncertainty indices. We employ the proposed multi-modal connectedness measure framework to depict the individual-level, provincial-level, and indicator-level connectedness between China's regional macroeconomy and the relationships between the regional macroeconomy and uncertainty indices.

The structure of this paper is as follows. Section 2 introduces the MMAR model. Section 3 of the main paper presents three estimation methods: the projection method, the alternating least squares method, and the maximum likelihood method. The asymptotic theory of these methods is discussed in Section 4. In Section 5, we evaluate the finite-sample performance of our estimations through Monte Carlo simulations. Empirical studies are conducted in Section 6. The final section concludes the paper.

2. Multi-Matrix Autoregressive Model

Let $\{\boldsymbol{X}_{it}\}_{i=1}^{K}$ denote K matrix-valued time series with different structures that can be preliminarily recognized. The $\boldsymbol{X}_{it}$, with dimensions $m_i \times n_i$ $(i = 1, \ldots, K)$, is given by

$$\boldsymbol{X}_{it} = \begin{bmatrix} X_{it,11} & \cdots & X_{it,1n_i} \\ \vdots & \ddots & \vdots \\ X_{it,m_i1} & \cdots & X_{it,m_in_i} \end{bmatrix}. \tag{1}$$

We model the relationships among $\boldsymbol{X}_{it}$s by the following K equations:

$$\begin{aligned} (\boldsymbol{X}_{1t})_{m_1\times n_1} &= \sum_{j=1}^{K} (\boldsymbol{A}_{1j})_{m_1\times m_j}(\boldsymbol{X}_{j,t-1})_{m_j\times n_j}(\boldsymbol{B}_{1j})'_{n_1\times n_j} \\ &\quad + (\boldsymbol{E}_{1t})_{m_1\times n_1}, \\ (\boldsymbol{X}_{2t})_{m_2\times n_2} &= \sum_{j=1}^{K} (\boldsymbol{A}_{2j})_{m_2\times m_j}(\boldsymbol{X}_{j,t-1})_{m_j\times n_j}(\boldsymbol{B}_{2j})'_{n_2\times n_j} \\ &\quad + (\boldsymbol{E}_{2t})_{m_2\times n_2}, \\ &\ \ \vdots \\ (\boldsymbol{X}_{Kt})_{m_K\times n_K} &= \sum_{j=1}^{K} (\boldsymbol{A}_{Kj})_{m_K\times m_j}(\boldsymbol{X}_{j,t-1})_{m_j\times n_j}(\boldsymbol{B}_{Kj})'_{n_K\times n_j} \\ &\quad + (\boldsymbol{E}_{Kt})_{m_K\times n_K}, \end{aligned} \tag{2}$$

where $\boldsymbol{A}_{ij}$ and $\boldsymbol{B}_{ij}$ for $i,j = 1, \ldots, K$, are $m_i \times m_j$ and $n_i \times n_j$ autoregressive coefficient matrices, respectively. In each equation, the lag terms $\boldsymbol{X}_{j,t-1}$ s, $j = 1, \ldots, K$, appear on the right side, and the inclusion of $\boldsymbol{X}_{j,t-1}, j \neq i$, in the ith equation captures the impacts of the variables in $\boldsymbol{X}_{j,t-1}$ on the variables in $\boldsymbol{X}_{i,t}$. As (2) contains only the first-order lag of the variables, we refer to it as the MMAR(1) model.

The random disturbances $\boldsymbol{E}_{it}$ $(i = 1, \ldots, K)$ are assumed to follow a matrix-valued i.i.d. process. Since the covariance matrix of a matrix-valued random variable is a tensor of order 4, which is difficult to express, we focus alternatively on $\boldsymbol{\varepsilon}_t = (\mathrm{vec}(\boldsymbol{E}_{1t})', \ldots, \mathrm{vec}(\boldsymbol{E}_{Kt})')'$, and define $\mathrm{Cov}(\boldsymbol{\varepsilon}_t) = \Sigma$ as the $L \times L$ covariance matrix of $\boldsymbol{\varepsilon}_t$, where $L = \sum_{i=1}^{K} m_i n_i$. For the least squares estimation, we allow Σ to have arbitrary structure and only require that Σ is nonsingular, while for the maximum likelihood estimation, we consider a structured setting of the covariance matrix for $\boldsymbol{E}_{it}$:

$$\mathrm{Cov}(\mathrm{vec}(\boldsymbol{E}_{it})) = \Sigma_{ic} \otimes \Sigma_{ir}, \tag{3}$$

where Σ_{ir} and Σ_{ic} are $m_i \times m_i$ and $n_i \times n_i$ symmetric positive definite matrices. Note that when $\mathrm{vec}(\boldsymbol{E}_{it})$ is normal, we have $\boldsymbol{E}_{it} = \Sigma_{ir}^{1/2}\boldsymbol{Z}_{it}\Sigma_{ic}^{1/2}$, where all elements of $\boldsymbol{Z}_{it}$ follow independent standardized normal distributions. Thus, Σ_{ir} and Σ_{ic} can be interpreted as row-wise and column-wise covariances corresponding to $\boldsymbol{E}_{it}$.

In particular, when $K = 1$, the model reduces to the MAR(1) in Chen et al. (2021):

$$(\boldsymbol{X}_{1t})_{m_1\times n_1} = (\boldsymbol{A}_1)_{m_1\times m_1}(\boldsymbol{X}_{1,t-1})_{m_1\times n_1}(\boldsymbol{B}_1)'_{n_1\times n_1} + (\boldsymbol{E}_{1t})_{m_1\times n_1}.$$

From a more generalized perspective, our model treats matrices of different dimensions as distinct variables. In matrix space, the MAR(1) model becomes an "AR(1)" model for a single matrix variable, and our model can be considered as a vector autoregressive model "VAR(1)" treating multiple matrices as different variables. Furthermore, when the dimensions of the matrices in our variables are the same, our model can be viewed as a more complicated order-3 tensor autoregressive model (TenAR(1)).

Let $\boldsymbol{y}_t = \left(\mathrm{vec}(\boldsymbol{X}_{1t})', \ldots, \mathrm{vec}(\boldsymbol{X}_{Kt})'\right)'$. After vectorization, the MMAR(1) model can be represented as a parsimonious representation of the VAR(1) model:

$$\boldsymbol{y}_t = \Phi\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{4}$$

with the $L \times L$ coefficient matrix Φ takes the form:

$$\Phi = \begin{bmatrix} \boldsymbol{B}_{11}\otimes\boldsymbol{A}_{11} & \cdots & \boldsymbol{B}_{1K}\otimes\boldsymbol{A}_{1K} \\ \vdots & \ddots & \vdots \\ \boldsymbol{B}_{K1}\otimes\boldsymbol{A}_{K1} & \cdots & \boldsymbol{B}_{KK}\otimes\boldsymbol{A}_{KK} \end{bmatrix}, \tag{5}$$

where $\otimes$ denotes the Kronecker product.

From a statistical perspective, the MMAR model extends the MAR framework by allowing interactions across multiple matrix-valued processes. In the MAR case, the VAR representation admits a single Kronecker product structure for the coefficient matrix. By contrast, when multiple matrices are involved, the MMAR model induces a block-wise Kronecker representation of the implied VAR coefficient matrix, in which the (i,j)-th block takes the form $\Phi_{ij} = \boldsymbol{B}_{ij} \otimes \boldsymbol{A}_{ij}$, with block dimensions depending on those of $\boldsymbol{X}_{it}$ and $\boldsymbol{X}_{jt}$. This block-wise Kronecker structure introduces additional complexity, as estimation and inference must account for interactions across multiple heterogeneous blocks rather than a single global Kronecker factorization. As a consequence, existing MAR estimation procedures cannot be applied directly to the MMAR setting, and both estimation methods and asymptotic analysis require a joint block-wise treatment, which motivates the development of new estimation strategies and asymptotic analysis in the subsequent sections.

Remark 1 (Comparison with VAR, MAR, and TenAR Models). The MMAR model can be viewed as a structured and parsimonious alternative to existing multivariate time series models. In a VAR(1) representation, the coefficient matrix Φ contains $\left(\sum_i m_i n_i\right)^2$ unrestricted parameters, which quickly becomes infeasible in high-dimensional settings. The MMAR(1) model, by contrast, imposes block-wise Kronecker restrictions on Φ, reducing the number of parameters to $\sum_{i=1}^{K}\sum_{j=1}^{K}(m_i m_j + n_i n_j)$ while preserving interpretable interactions across different matrix-valued processes. When matrix-valued series have compatible dimensions, they may alternatively be stacked into a single matrix and modeled using a MAR specification, or further aggregated, possibly after zero-padding, into a higher-order tensor and modeled using a tensor autoregressive (TenAR) framework. Such transformations, however, implicitly impose strong global structural restrictions: TenAR models rely on multilinear parameterizations under which individual autoregressive coefficients no longer correspond directly to block-level interactions in the original data. Consequently, zero-padding or tensorization does not translate into simple zero restrictions on the underlying VAR coefficients and may lead to identification and interpretability issues in economic applications. The MMAR framework occupies an intermediate position between these approaches, accommodating heterogeneous matrix dimensions and exploiting the intrinsic row-column structure through block-wise Kronecker constraints, while retaining coefficient interpretations at the level of individual matrix interactions. This balance between structural parsimony and economic interpretability makes MMAR particularly suitable for applications in which matrix-valued series cannot be naturally stacked into a single matrix or tensor. Additional discussion and illustrative examples are provided in Appendix C of the Online Supplement.

The above MMAR(1) model can be easily extended to a general model of p lag orders with its ith equation given by:

$$\boldsymbol{X}_{it} = \sum_{\ell=1}^{p}\sum_{j=1}^{K} \boldsymbol{A}_{ij}^{(\ell)} \boldsymbol{X}_{j,t-\ell} \boldsymbol{B}_{ij}^{(\ell)\prime} + \boldsymbol{E}_{it}, \tag{6}$$

for $i = 1, \ldots, K$. We refer to the model given by (6) as the MMAR(p) model, where p denotes the lag order. The MMAR(p) model can be also rewritten as the following equivalent VAR(p) representation:

$$\boldsymbol{y}_t = \Phi_1 \boldsymbol{y}_{t-1} + \cdots + \Phi_p \boldsymbol{y}_{t-p} + \boldsymbol{\varepsilon}_t, \tag{7}$$

where $\boldsymbol{y}_t = \left(\mathrm{vec}(\boldsymbol{X}_{1t})', \ldots, \mathrm{vec}(\boldsymbol{X}_{Kt})'\right)'$, $\boldsymbol{\varepsilon}_t = (\mathrm{vec}(\boldsymbol{E}_{1t})', \ldots, \mathrm{vec}(\boldsymbol{E}_{Kt})')'$, and

$$\Phi_\ell = \begin{bmatrix} \boldsymbol{B}_{11}^{(\ell)} \otimes \boldsymbol{A}_{11}^{(\ell)} & \cdots & \boldsymbol{B}_{1K}^{(\ell)} \otimes \boldsymbol{A}_{1K}^{(\ell)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{B}_{K1}^{(\ell)} \otimes \boldsymbol{A}_{K1}^{(\ell)} & \cdots & \boldsymbol{B}_{KK}^{(\ell)} \otimes \boldsymbol{A}_{KK}^{(\ell)} \end{bmatrix}, \quad \ell = 1, \ldots, p. \tag{8}$$

In the MMAR(p) model given by (6), the matrices $\boldsymbol{A}_{ij}^{(l)}$ and $\boldsymbol{B}_{ij}^{(l)}$ for $i, j = 1, \ldots, K$ and $l = 1, \ldots, p$ are not uniquely identifiable. Specifically, the model remains unchanged if $\boldsymbol{A}_{ij}^{(l)}$ is divided by a nonzero constant and $\boldsymbol{B}_{ij}^{(l)}$ is multiplied by the same constant. To address this, we normalize each $\boldsymbol{A}_{ij}^{(l)}$ such that its Frobenius norm equals one, following the approach in the MAR model. The Kronecker product $\boldsymbol{B}_{ij}^{(l)} \otimes \boldsymbol{A}_{ij}^{(l)}$ is then unique up to sign. For complete identification, one can set $\boldsymbol{A}_{11}^{(l)} = 1$.

Remark 2 (Causality). Obtaining the causality condition for the MMAR model is intuitive. For the MMAR(1) model, if $\rho(\Phi) < 1$, then the process $\boldsymbol{X}_{it}, i = 1, \ldots, K$ is causal and stationary, where Φ is defined in (5) and $\rho(\Phi)$ represent its spectral radius. For MMAR(p), if for all $|z| \le 1$, $\det(\boldsymbol{I} - \sum_{\ell=1}^{p} \Phi_\ell z^\ell) \neq 0$, then the process $\boldsymbol{X}_{it}, i = 1, \ldots, K$ is causal and stationary, where Φ_ℓ is defined in (8). For more details, the readers are referred to Chapter 11 in Brockwell et al. (1991) and Chapter 2 in Tsay (2014).

We discuss the model interpretation in Appendix A of the Online Supplement and provide several examples in Appendix B of the Online Supplement to offer a more intuitive understanding of the extensive potential of the MMAR model in empirical applications.

3. Model Estimation

Suppose that the available data set is $\mathcal{F}_T = \{\boldsymbol{y}_T, \boldsymbol{y}_{T-1}, \ldots, \boldsymbol{y}_1, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_{1-p}\}$ and let the initial values $\mathcal{F}_0 = \{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{1-p}\}$, where $\boldsymbol{y}_t = (\mathrm{vec}(\boldsymbol{X}_{1t})', \ldots, \mathrm{vec}(\boldsymbol{X}_{Kt})')'$ and $\boldsymbol{X}_{it}$ are observational matrix data with dimension $m_i \times n_i$.

3.1. Projection Estimation

We begin by introducing a projection estimation method for the MMAR model and first focus on the MMAR(1) specification in (4). The extension to the MMAR(p) case is conceptually straightforward and is discussed in detail later. In contrast to the projection estimator studied in Chen et al. (2021), the presence of multiple matrices induces a block-wise Kronecker structure in the implied VAR representation. As a result, projection estimation in the MMAR setting requires performing block-specific nearest Kronecker product projections and combining them across all (i, j) blocks, which in turn requires a joint asymptotic analysis across these blocks. At the same time, the MAR(1) and MAR(p) models arise naturally as special cases of the proposed framework when $K = 1$.

The basic idea of projection estimation is to first obtain a maximum likelihood estimate or a least squares estimate $\widehat{\Phi}$

under the VAR(1) model in (5), then estimate $\boldsymbol{A}_{ij}$ and $\boldsymbol{B}_{ij}$ by projecting $\widehat{\Phi}_{ij}$ in the space of the Kronecker product under matrix Frobenius norm:

$$\left(\widehat{\boldsymbol{A}}_{ij}, \widehat{\boldsymbol{B}}_{ij}\right) = \underset{\boldsymbol{A}_{ij},\boldsymbol{B}_{ij}}{\arg\min} ||\widehat{\Phi}_{ij} - \boldsymbol{B}_{ij} \otimes \boldsymbol{A}_{ij}||_F^2 \tag{9}$$

for $i, j = 1, \ldots, K$, where $\widehat{\Phi}_{ij}$ is the (i,j)-th block of $\widehat{\Phi}$ with the dimension $m_i n_i \times m_j n_j$.

The above minimization problem is named the nearest Kronecker product (NKP) problem in matrix computation (see Van Loan and Pitsianis (1993) and Van Loan (2000)). The entries in $\boldsymbol{B}_{ij} \otimes \boldsymbol{A}_{ij}$ are the same as $\mathrm{vec}(\boldsymbol{A}_{ij})\mathrm{vec}(\boldsymbol{B}_{ij})'$. Therefore, we can define permutation operators $\mathcal{G}_{ij} : \mathbb{R}^{m_i n_i} \times \mathbb{R}^{m_j n_j} \to \mathbb{R}^{m_i \cdot m_j} \times \mathbb{R}^{n_i \cdot n_j}$, such that $\mathcal{G}_{ij}(\boldsymbol{B}_{ij} \otimes \boldsymbol{A}_{ij}) = \mathrm{vec}(\boldsymbol{A}_{ij})\mathrm{vec}(\boldsymbol{B}_{ij})'$. The NKP problem can be solved through permutation and singular value decomposition (SVD). Under the permutation, we have:

$$||\widehat{\Phi}_{ij} - \boldsymbol{B}_{ij} \otimes \boldsymbol{A}_{ij}||_F^2 = ||\mathcal{G}_{ij}(\widehat{\Phi}_{ij}) - \mathrm{vec}(\boldsymbol{A}_{ij})\mathrm{vec}(\boldsymbol{B}_{ij})'||_F^2.$$

Equation (9) of the main paper is equivalent to find a best approximation of $\widehat{\Phi}_{ij}$, which can be solved by letting:

$$\mathrm{vec}(\boldsymbol{A}_{ij})\mathrm{vec}(\boldsymbol{B}_{ij})' = d_1^{(ij)} \boldsymbol{\mu}_1^{(ij)} \boldsymbol{\nu}_1^{(ij)\prime},$$

where $d_1^{(ij)}$ denotes the largest singular value of $\widehat{\Phi}_{ij}$, $\boldsymbol{\mu}_1^{(ij)}$ and $\boldsymbol{\nu}_1^{(ij)}$ are its corresponding left and right singular vectors, respectively.

Note that the above estimate requires the estimate of the $L \times L$ matrix Φ. Since $L = \sum_{j=1}^K m_j n_j$, which is potentially very large, the estimated $\widehat{\Phi}$ may not be as accurate. We will use the estimates obtained from the projection method as the initial values for the two iterative estimation methods proposed later.

For the MMAR(p) model in (6), consider its equivalent VAR(p) representation: $\boldsymbol{y}_t = \sum_{k=1}^p \Phi_k \boldsymbol{y}_{t-k} + \boldsymbol{\varepsilon}_t$. We can first obtain the ML estimate or the least squares estimate of $\Phi_1, \ldots, \Phi_p$, denoted by $\widehat{\Phi}_1, \ldots, \widehat{\Phi}_p$, and then use the projection estimation procedure to obtain the estimates of $\boldsymbol{A}_{ij}^{(l)}$ and $\boldsymbol{B}_{ij}^{(l)}$ by projecting on each $\widehat{\Phi}_{ij}^{(k)}$, for $i, j = 1, \ldots, K$ and $l = 1, \ldots, p$, where $\widehat{\Phi}_{ij}^{(l)}$ is the (i,j)-th block of $\widehat{\Phi}_l$.

3.2. Alternating Least Squares

Let $\boldsymbol{G}_A$ and $\boldsymbol{G}_B$ denote the collections of $\boldsymbol{A}_{ij}^{(k)}$ and $\boldsymbol{B}_{ij}^{(k)}$ for $i, j = 1, \ldots, K$ and $k = 1, \ldots, p$, respectively, denoted by $\boldsymbol{G}_A = \{\boldsymbol{A}_{ij}^{(k)}, i, j = 1, \ldots, K; k = 1, \ldots, p\}$ and $\boldsymbol{G}_B = \{\boldsymbol{B}_{ij}^{(k)}, i, j = 1, \ldots, K; k = 1, \ldots, p\}$. The least squares estimates of $\boldsymbol{G}_A$ and $\boldsymbol{G}_B$ for the MMAR(p) model are the solution of the following minimization problem:

$$\min_{\boldsymbol{G}_A, \boldsymbol{G}_B} \sum_{t=1}^T \sum_{i=1}^K \left\| \boldsymbol{X}_{it} - \sum_{k=1}^p \sum_{j=1}^K \boldsymbol{A}_{ij}^{(k)} \boldsymbol{X}_{j,t-k} \boldsymbol{B}_{ij}^{(k)\prime} \right\|_F^2. \tag{10}$$

Under the bilinear structure of the MMAR model, the optimization problem in (10) is non-convex, and alternating least squares (ALS) provides a natural solution strategy. A direct extension of existing matrix-wise iterative algorithms would update one coefficient matrix $\boldsymbol{A}_{ij}^{(k)}$ or $\boldsymbol{B}_{ij}^{(k)}$ at a time while keeping all others fixed (Li and Xiao 2021). However, in the MMAR setting this approach would require updating $2K^2p$ coefficient matrices in each iteration, leading to a substantial computational burden as the number of matrices or lags increases. To address this issue, we reformulate the problem at the equation level and derive an equivalent ALS representation in which, for each equation, only two aggregated parameter matrices are updated per iteration. As a result, the proposed ALS procedure achieves reduced computational cost in the multi-matrix setting.

Specifically, we define the following quantities:

$$\begin{aligned}
\boldsymbol{W}_{it}^A &= [\boldsymbol{A}_{i1}^{(1)} \boldsymbol{X}_{1,t-1}, \ldots, \boldsymbol{A}_{iK}^{(1)} \boldsymbol{X}_{K,t-1}, \ldots \ \boldsymbol{A}_{iK}^{(p)} \boldsymbol{X}_{K,t-p}],\\
\boldsymbol{A}_i &= [\boldsymbol{A}_{i1}^{(1)}, \ldots, \boldsymbol{A}_{iK}^{(1)}, \ldots, \boldsymbol{A}_{iK}^{(p)}],\\
\boldsymbol{W}_{it}^B &= [\boldsymbol{B}_{i1}^{(1)} \boldsymbol{X}_{1,t-1}', \ldots, \boldsymbol{B}_{iK}^{(1)} \boldsymbol{X}_{K,t-1}', \ldots, \boldsymbol{B}_{iK}^{(p)} \boldsymbol{X}_{K,t-p}']',\\
\boldsymbol{B}_i &= [\boldsymbol{B}_{i1}^{(1)}, \ldots, \boldsymbol{B}_{iK}^{(1)}, \ldots, \boldsymbol{B}_{iK}^{(p)}].
\end{aligned}$$

Then the ith equation in the MMAR(p) model can be rewritten as:

$$\begin{aligned}
\boldsymbol{X}_{it} &= \sum_{k=1}^p \sum_{j=1}^K \boldsymbol{A}_{ij}^{(k)} \boldsymbol{X}_{j,t-k} \boldsymbol{B}_{ij}^{(k)\prime} + \boldsymbol{E}_{it} = \boldsymbol{A}_i \boldsymbol{W}_{it}^B + \boldsymbol{E}_{it}\\
&= \boldsymbol{W}_{it}^A \boldsymbol{B}_i' + \boldsymbol{E}_{it},
\end{aligned}$$

and the minimization problem is equivalent to

$$\min_{\boldsymbol{G}_A, \boldsymbol{G}_B} \sum_{t=1}^T \sum_{i=1}^K \left\| \boldsymbol{X}_{it} - \boldsymbol{A}_i \boldsymbol{W}_{it}^B \right\|_F^2 \quad \text{or}$$
$$\min_{\boldsymbol{G}_A, \boldsymbol{G}_B} \sum_{t=1}^T \sum_{i=1}^K \left\| \boldsymbol{X}_{it} - \boldsymbol{W}_{it}^A \boldsymbol{B}_i' \right\|_F^2.$$

Using the above equivalent optimization problem, for each equation, we can update the parameter matrix $\boldsymbol{A}_i$ or $\boldsymbol{B}_i$ in a single iteration while keeping the other unchanged and continuing iterating until convergence is reached. Compared to the previous iterative algorithm, our newly proposed algorithm only needs to update two parameter matrices in each iteration, significantly reducing the computation time required for each iteration.

Taking partial derivatives of (10) with respect to $\boldsymbol{A}_i$ s and $\boldsymbol{B}_i$ s, we can obtain the first-order conditions:

$$\begin{aligned}
&\sum_{t=1}^T \boldsymbol{A}_i \boldsymbol{W}_{it}^B \boldsymbol{W}_{it}^{B\prime} - \sum_{t=1}^T \boldsymbol{X}_{it} \boldsymbol{W}_{it}^{B\prime} = \boldsymbol{0},\\
&\sum_{t=1}^T \boldsymbol{B}_i \boldsymbol{W}_{it}^{A\prime} \boldsymbol{W}_{it}^A - \sum_{t=1}^T \boldsymbol{X}_{it}' \boldsymbol{W}_{it}^A = \boldsymbol{0},
\end{aligned} \tag{11}$$

for $i = 1, \ldots, K$. Therefore, for each equation i, given $\boldsymbol{B}_i$, we can update $\boldsymbol{A}_i$ by:

$$\boldsymbol{A}_i \longleftarrow \left(\sum_{t=1}^T \boldsymbol{X}_{it} \boldsymbol{W}_{it}^{B\prime} \right) \left(\sum_{t=1}^T \boldsymbol{W}_{it}^B \boldsymbol{W}_{it}^{B\prime} \right)^{-1}.$$

Then, given $\boldsymbol{A}_i$, we can update $\boldsymbol{B}_i$ by:

$$\boldsymbol{B}_i \longleftarrow \left(\sum_{t=1}^T \boldsymbol{X}_{it}' \boldsymbol{W}_{it}^A \right) \left(\sum_{t=1}^T \boldsymbol{W}_{it}^{A\prime} \boldsymbol{W}_{it}^A \right)^{-1}.$$

Note that the iterations for each equation are independent of each other, allowing us to estimate them equation-wise, thus substantially improving computational efficiency.

The other thing is that $\boldsymbol{A}_{ij}^{(k)}$ s and $\boldsymbol{B}_{ij}^{(k)}$ s have identification problems. To ensure a stable numerical computation, in each iteration, we renormalize $\boldsymbol{A}_{ij}^{(k)}$ and $\boldsymbol{B}_{ij}^{(k)}$ so that $||\boldsymbol{A}_{ij}^{(k)}||_F = 1$ for $j = 1, \dots, K$. In practice, the projection estimators of $\boldsymbol{A}_{ij}^{(k)}$ and $\boldsymbol{B}_{ij}^{(k)}$ can be used as initial values.

Using an equation-wise estimation procedure, the ALS algorithm updates the aggregated parameter matrices $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ for each equation while keeping the other fixed and iterates until convergence. Compared with conventional block-wise iterative procedures that update individual coefficient matrices separately, this approach reduces the number of matrix updates in each full iteration from $2K^2p$ to $2K$. Although each aggregated update may involve a larger matrix operation, the lower update frequency improves computational efficiency in moderate- to large-dimensional MMAR systems.

3.3. MLE When $\mathrm{Cov}(\mathrm{vec}(E_{it}))$ Are Separable

We now present the maximum likelihood estimation of the proposed MMAR(p) model in (6). Let $\boldsymbol{\theta}$ denote the vector collecting all model parameters. As in the least squares case, existing MLE procedures developed for the MAR model do not directly extend to the MMAR setting due to the block-wise Kronecker structure of the implied VAR representation. To obtain a feasible likelihood-based estimator, we impose the assumption that the innovation matrices $\{\boldsymbol{E}_{it}\}$ are contemporaneously uncorrelated and follow the separable structure in (3). Under these assumptions, we develop an iterated MLE procedure for the MMAR model. Given the initial values $\mathcal{F}_0 = \{\boldsymbol{y}_0, \dots, \boldsymbol{y}_{1-p}\}$, the log-likelihood under normality can be written as $\ell(\boldsymbol{\theta}) = \sum_{i=1}^K \ell_i(\boldsymbol{\theta})$ with

$$\ell_i(\boldsymbol{\theta}) = -\frac{Tm_in_i}{2}\log 2\pi - m_i(T-1)\log|\Sigma_{ic}| - n_i(T-1)\log|\Sigma_{ir}| - \sum_{t=1}^T \mathrm{tr}\left(\Sigma_{ir}^{-1}\boldsymbol{E}_{it}\Sigma_{ic}^{-1}\boldsymbol{E}_{it}'\right),$$

where $\boldsymbol{E}_{it} = \boldsymbol{X}_{it} - \sum_{k=1}^p\sum_{j=1}^K \boldsymbol{A}_{ij}^{(k)}\boldsymbol{X}_{j,t-k}\boldsymbol{B}_{ij}^{(k)\prime} = \boldsymbol{X}_{it} - \boldsymbol{A}_i\boldsymbol{W}_{it}^B = \boldsymbol{X}_{it} - \boldsymbol{W}_{it}^A\boldsymbol{B}_i'$, and the quantities $\boldsymbol{W}_{it}^A$, $\boldsymbol{W}_{it}^B$, $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ are defined as before.

Given that the $\boldsymbol{E}_{it}$ s are uncorrelated, we have $\ell_i(\boldsymbol{\theta}) = \ell_i(\boldsymbol{\theta}_i)$, where $\boldsymbol{\theta}_i$ contains all parameters in the ith equation. Consequently, the maximum likelihood estimation of $\boldsymbol{\theta}$ is equivalent to the maximum likelihood estimation of $\boldsymbol{\theta}_i$ equation-wise, that is,

$$\widehat{\boldsymbol{\theta}}_i = \arg\max_{\boldsymbol{\theta}_i} \ell_i(\theta_i), \quad i = 1, \dots, K.$$

For $i = 1, \dots, K$, the first-order conditions of $\boldsymbol{A}_i$, $\boldsymbol{B}_i$, Σ_{ir} and Σ_{ic} are given by:

$$\sum_{t=1}^T \boldsymbol{A}_i\boldsymbol{W}_{it}^B\Sigma_{ic}^{-1}\boldsymbol{W}_{it}^{B\prime} - \sum_{t=1}^T \boldsymbol{X}_{it}\Sigma_{ic}^{-1}\boldsymbol{W}_{it}^{B\prime} = \boldsymbol{0},$$

$$m_i(T-1)\Sigma_{ic} - \sum_{t=1}^T \boldsymbol{E}_{it}'\Sigma_{ir}^{-1}\boldsymbol{E}_{it} = \boldsymbol{0},$$

$$\sum_{t=1}^T \boldsymbol{B}_i\boldsymbol{W}_{it}^{A\prime}\Sigma_{ir}^{-1}\boldsymbol{W}_{it}^A - \sum_{t=1}^T \boldsymbol{X}_{it}\Sigma_{ir}^{-1}\boldsymbol{W}_{it}^A = \boldsymbol{0},$$

$$n_i(T-1)\Sigma_{ir} - \sum_{t=1}^T \boldsymbol{E}_{it}\Sigma_{ic}^{-1}\boldsymbol{E}_{it}' = \boldsymbol{0}.$$

For the estimation of $\boldsymbol{A}_i$, $\boldsymbol{B}_i$, Σ_{ir} and Σ_{ic}, $i = 1, \dots, K$, we can then iteratively update one, while keeping the other parameters fixed, i.e.,

$$\boldsymbol{A}_i \longleftarrow \left(\sum_{t=1}^T \boldsymbol{X}_{it}\Sigma_{ic}^{-1}\boldsymbol{W}_{it}^{B\prime}\right)\left(\sum_{t=1}^T \boldsymbol{W}_{it}^B\Sigma_{ic}^{-1}\boldsymbol{W}_{it}^{B\prime}\right)^{-1},$$

$$\boldsymbol{B}_i \longleftarrow \left(\sum_{t=1}^T \boldsymbol{X}_{it}'\Sigma_{ir}^{-1}\boldsymbol{W}_{it}^{A}\right)\left(\sum_{t=1}^T \boldsymbol{W}_{it}^{A\prime}\Sigma_{ir}^{-1}\boldsymbol{W}_{it}^{A}\right)^{-1},$$

$$\Sigma_{ic} \leftarrow \frac{1}{m_i(T-1)}\sum_{t=1}^T \boldsymbol{E}_{it}'\Sigma_{ir}^{-1}\boldsymbol{E}_{it},$$

$$\Sigma_{ir} \leftarrow \frac{1}{n_i(T-1)}\sum_{t=1}^T \boldsymbol{E}_{it}\Sigma_{ic}^{-1}\boldsymbol{E}_{it}'.$$

To make sure the numerical computation is stable, we address identification problems by renormalizing $\boldsymbol{A}_{ij}^{(k)}$, $\boldsymbol{B}_{ij}^{(k)}$, Σ_{ir} and Σ_{ic} so that $||\boldsymbol{A}_{ij}^{(k)}||_F = 1$ and $||\Sigma_{ir}||_F = 1$, for $i, j = 1, \dots, K$ and $k = 1, \dots, p$.

Remark 3 (Contemporaneous Uncorrelatedness Assumption). The assumption that the innovation matrices $\{\boldsymbol{E}_{it}\}$ are contemporaneously uncorrelated is introduced specifically for the iterated MLE and plays a central role in making likelihood-based estimation of the MMAR model computationally feasible. Under this restriction, the innovation covariance matrix admits a block-diagonal separable structure, so that the likelihood function can be decomposed equation by equation. Although this assumption can in principle be relaxed, doing so leads to a substantially different estimation problem: once contemporaneous cross-equation dependence is allowed, the equation-wise decomposition no longer holds, and the off-diagonal precision blocks must be updated jointly through nonlinear first-order conditions, for which no tractable alternating scheme is currently available. From an economic perspective, after accounting for the lagged cross-matrix effects, contemporaneous uncorrelatedness means that the remaining idiosyncratic shocks do not transmit instantaneously across matrices, while dynamic interdependence is still captured through the lagged terms. This assumption can be appropriate in applications where interactions propagate with delays, such as production networks, but may be less suitable when strong instantaneous spillovers exist. In such cases, estimators that do not rely on this restriction, such as the ALS estimator, may be preferable. A more detailed discussion is provided in Appendix D of the Online Supplement.

4. Theoretical Properties

4.1. Asymptotics for Projection Estimators

We first show the central limit theorem for the projection estimator of MMAR(p) model: $\widehat{\boldsymbol{A}}_{ij}^{(l)}$*s* and $\widehat{\boldsymbol{B}}_{ij}^{(l)}$*s*. Since the MMAR($p$) has an equivalent VAR($p$) representation in (7) of the main paper, let $\boldsymbol{z}_t = (\boldsymbol{y}_{t-1}', \dots, \boldsymbol{y}_{t-p}')'$ and $\Phi = [\Phi_1, \dots, \Phi_p]$. Note that here we follow the standard theory of VAR models (Lütkepohl 1991; Hamilton 1994), the least square estimator $\hat{\Phi}$ converges to a multivariate normal distribution: $\sqrt{T}\mathrm{vec}(\widehat{\Phi} - \Phi) \to \mathcal{N}(\mathbf{0}, \Gamma^{-1} \otimes \Sigma)$, where $\Sigma = \mathrm{Cov}(\boldsymbol{\varepsilon}_t)$ and $\Gamma = \mathrm{Cov}(\boldsymbol{z}_t, \boldsymbol{z}_t')$. Moreover, define

$$\mathcal{R}(\Phi_l) = \left[\mathcal{G}_{11}(\Phi_{11}^{(l)}), \dots, \mathcal{G}_{K1}(\Phi_{K1}^{(l)}), \dots, \mathcal{G}_{KK}(\Phi_{KK}^{(l)})\right]$$

and that $\mathcal{R}(\Phi) = [\mathcal{R}(\Phi_1), \dots, \mathcal{R}(\Phi_p)]$, then we have:

$$\sqrt{T}\mathrm{vec}\left(\mathcal{R}(\widehat{\Phi}) - \mathcal{R}(\Phi)\right) \to \mathcal{N}(\mathbf{0}, \Xi_1),$$

where the matrix Ξ_1 denotes the rearrangement of $\Gamma^{-1} \otimes \Sigma$ under the operator $\mathcal{R}(\cdot)$.

Theorem 1. Consider the MMAR(p) model. Assume that $||\widehat{\boldsymbol{A}}_{ij}^{(l)}||_F = 1$. *Let* $\boldsymbol{\alpha}_{ij}^{(l)} = \mathrm{vec}(\boldsymbol{A}_{ij}^{(l)})$, $\boldsymbol{\beta}_{ij}^{(l)} = \mathrm{vec}(\boldsymbol{B}_{ij}^{(l)})/||\mathrm{vec}(\boldsymbol{B}_{ij}^{(l)})||$. *Moreover, let*

$$\boldsymbol{V}_{0,ij}^{(l)} = \begin{pmatrix} ||\boldsymbol{B}_{ij}^{(l)}||[\boldsymbol{\beta}_{ij}^{(l)\prime} \otimes (\boldsymbol{I} - \boldsymbol{\alpha}_{ij}^{(l)}\boldsymbol{\alpha}_{ij}^{(l)\prime})] \\ \boldsymbol{I} \otimes \boldsymbol{\alpha}_{ij}^{(l)\prime} \end{pmatrix}$$

$$\boldsymbol{V}_{1,ij}^{(l)} = (\boldsymbol{\beta}_{ij}^{(l)}\boldsymbol{\beta}_{ij}^{(l)\prime}) \otimes \boldsymbol{I} + \boldsymbol{I} \otimes (\boldsymbol{\alpha}_{ij}^{(l)}\boldsymbol{\alpha}_{ij}^{(l)\prime}) - (\boldsymbol{\beta}_{ij}^{(l)}\boldsymbol{\beta}_{ij}^{(l)\prime}) \otimes (\boldsymbol{\alpha}_{ij}^{(l)}\boldsymbol{\alpha}_{ij}^{(l)\prime})$$

for $i = 1, \dots, K$, $j = 1, \dots, K$. *And let the block diagonal matrix* $\boldsymbol{V}_0^{(l)} = \mathrm{diag}\{\boldsymbol{V}_{0,11}^{(l)}, \dots, \boldsymbol{V}_{0,1K}^{(l)}, \dots, \boldsymbol{V}_{0,KK}^{(l)}\}$ *and* $\boldsymbol{V}_1^{(l)} = \mathrm{diag}\{\boldsymbol{V}_{1,11}^{(l)}, \dots, \boldsymbol{V}_{1,1K}^{(l)}, \dots, \boldsymbol{V}_{1,KK}^{(l)}\}$.

Also, define $\boldsymbol{V}_0 = \mathrm{diag}\{\boldsymbol{V}_0^{(1)}, \dots, \boldsymbol{V}_0^{(p)}\}$ *and* $\boldsymbol{V}_1 = \mathrm{diag}\{\boldsymbol{V}_1^{(1)}, \dots, \boldsymbol{V}_1^{(p)}\}$. *Assume that* $\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_T$ *is iid with mean zero and finite second moment. Also, assume that the causality condition holds. Then it holds that:*

$$\sqrt{T}\begin{pmatrix} \mathrm{vec}(\widehat{\boldsymbol{A}}_{11}^{(1)} - \boldsymbol{A}_{11}^{(1)}) \\ \mathrm{vec}(\widehat{\boldsymbol{B}}_{11}^{(1)} - \boldsymbol{B}_{11}^{(1)}) \\ \cdots \\ \mathrm{vec}(\widehat{\boldsymbol{A}}_{1K}^{(1)} - \boldsymbol{A}_{1K}^{(1)}) \\ \mathrm{vec}(\widehat{\boldsymbol{B}}_{1K}^{(1)} - \boldsymbol{B}_{1K}^{(1)}) \\ \cdots \\ \mathrm{vec}(\widehat{\boldsymbol{A}}_{KK}^{(p)} - \boldsymbol{A}_{KK}^{(p)}) \\ \mathrm{vec}(\widehat{\boldsymbol{B}}_{KK}^{(p)} - \boldsymbol{B}_{KK}^{(p)}) \end{pmatrix} \to \mathcal{N}(\mathbf{0}, \boldsymbol{V}_0\Xi_1\boldsymbol{V}_0'). \tag{12}$$

and that

$$\sqrt{T}\begin{pmatrix} \mathrm{vec}(\widehat{\boldsymbol{B}}_{11}^{(1)} \otimes \widehat{\boldsymbol{A}}_{11}^{(1)}) - \mathrm{vec}(\boldsymbol{B}_{11}^{(1)} \otimes \boldsymbol{A}_{11}^{(1)}) \\ \cdots \\ \mathrm{vec}(\widehat{\boldsymbol{B}}_{K1}^{(1)} \otimes \widehat{\boldsymbol{A}}_{K1}^{(1)}) - \mathrm{vec}(\boldsymbol{B}_{K1}^{(1)} \otimes \boldsymbol{A}_{K1}^{(1)}) \\ \cdots \\ \mathrm{vec}(\widehat{\boldsymbol{B}}_{KK}^{(p)} \otimes \widehat{\boldsymbol{A}}_{KK}^{(p)}) - \mathrm{vec}(\boldsymbol{B}_{KK}^{(p)} \otimes \boldsymbol{A}_{KK}^{(p)}) \end{pmatrix} \to \mathcal{N}(\mathbf{0}, \boldsymbol{V}_1\Xi_1\boldsymbol{V}_1'). \tag{13}$$

The proof of the theorem is presented in the Appendix E of the Online Supplement.

4.2. Asymptotics for ALS Estimators

Now, we provide the asymptotics for the least squares estimates $\widehat{\boldsymbol{A}}_i^{(k)}$ s and $\widehat{\boldsymbol{B}}_i^{(k)}$ s. Let $\boldsymbol{\alpha}_i = \mathrm{vec}(\boldsymbol{A}_i)$, $\boldsymbol{\beta}_i = \mathrm{vec}(\boldsymbol{B}_i')$, and let $\boldsymbol{C}_i = [\boldsymbol{c}_{i11}, \dots, \boldsymbol{c}_{iK1}, \dots, \boldsymbol{c}_{i1p}, \dots, \boldsymbol{c}_{iKp}] \in \mathbb{R}^{d_i \times Kp}$, where each column $\boldsymbol{c}_{ijk}$ has $\mathrm{vec}(\boldsymbol{A}_{ij}^{(k)})$ in the corresponding sub-block $\boldsymbol{\alpha}_{ij}^{(k)}$ and zeros elsewhere. Note that the $\boldsymbol{\beta}_i$ considered here is different from the $\boldsymbol{\beta}_{ij}$ defined in Theorem 1. In Theorem 1, $\boldsymbol{\beta}_{ij}$ denotes the normalized vector $\mathrm{vec}(\boldsymbol{B}_{ij})/\|\mathrm{vec}(\boldsymbol{B}_{ij})\|$. In Theorem 2, we establish a central limit theorem for $\mathrm{vec}(\boldsymbol{B}_i')$, denoted by $\boldsymbol{\beta}_i$.

Theorem 2. Define $\boldsymbol{W}_{it}' = [\boldsymbol{W}_{it}^B \otimes \boldsymbol{I} : \boldsymbol{I} \otimes \boldsymbol{W}_{it}^A]$ *where* : *connects the two matrices,* $\boldsymbol{H}_i = E(\boldsymbol{W}_{it}\boldsymbol{W}_{it}') + \boldsymbol{C}_i\boldsymbol{C}_i'$, *and* $\Sigma^{(ij)} = \mathrm{Cov}(\mathrm{vec}(\boldsymbol{E}_{it})\mathrm{vec}(\boldsymbol{E}_{jt})')$. *Let* Ξ_2 *be a block matrix with:*

$$\Xi_2 = \begin{pmatrix} \Xi_{2,(11)} & \cdots & \Xi_{2,(K1)} \\ \vdots & \ddots & \vdots \\ \Xi_{2,(1K)} & \cdots & \Xi_{2,(KK)} \end{pmatrix},$$

where $\Xi_{2,(ij)} = \boldsymbol{H}_i^{-1}E(\boldsymbol{W}_{it}\Sigma^{(ij)}\boldsymbol{W}_{jt}')\boldsymbol{H}_j^{-1}$. *In addition to the conditions of Theorem 1, we assume* $\boldsymbol{E}_{it}$*s are absolutely continuous with respect to Lebesgue measure, then we have:*

$$\sqrt{T}\begin{pmatrix} \mathrm{vec}(\widehat{\boldsymbol{A}}_1 - \boldsymbol{A}_1) \\ \mathrm{vec}(\widehat{\boldsymbol{B}}_1' - \boldsymbol{B}_1') \\ \vdots \\ \mathrm{vec}(\widehat{\boldsymbol{A}}_K - \boldsymbol{A}_K) \\ \mathrm{vec}(\widehat{\boldsymbol{B}}_K' - \boldsymbol{B}_K') \end{pmatrix} \to \mathcal{N}(\mathbf{0}, \Xi_2). \tag{14}$$

Define $\boldsymbol{\phi}_i = ((\mathrm{vec}(\boldsymbol{B}_{i1}^{(1)\prime}) \otimes \mathrm{vec}(\boldsymbol{A}_{i1}^{(1)}))', \dots, (\mathrm{vec}(\boldsymbol{B}_{iK}^{(1)\prime}) \otimes \mathrm{vec}(\boldsymbol{A}_{iK}^{(1)}))', \dots, (\mathrm{vec}(\boldsymbol{B}_{iK}^{(p)\prime}) \otimes \mathrm{vec}(\boldsymbol{A}_{iK}^{(p)}))')'$ for $i = 1, \dots, K$ and $\boldsymbol{\phi} = (\boldsymbol{\phi}_1', \boldsymbol{\phi}_2', \dots, \boldsymbol{\phi}_K')'$. Based on the asymptotic properties of the ALS estimator shown in Theorem 2, we then have the following result.

Corollary 1. Define $\boldsymbol{G}_i^B = \mathrm{diag}\{\mathrm{vec}(\boldsymbol{B}_{ij}^{(1)\prime} \otimes \boldsymbol{I}), \dots, \mathrm{vec}(\boldsymbol{B}_{iK}^{(1)\prime} \otimes \boldsymbol{I}), \dots, \mathrm{vec}(\boldsymbol{B}_{iK}^{(p)\prime} \otimes \boldsymbol{I})\}$ *and* $\boldsymbol{G}_i^A = \mathrm{diag}\{\mathrm{vec}(\boldsymbol{I} \otimes \boldsymbol{A}_{ij}^{(1)}), \dots, \mathrm{vec}(\boldsymbol{I} \otimes \boldsymbol{A}_{iK}^{(1)}), \dots, \mathrm{vec}(\boldsymbol{I} \otimes \boldsymbol{A}_{iK}^{(p)})\}$, *for* $i = 1, \dots, K$. *Let* $\boldsymbol{G}_i = (\boldsymbol{G}_i^B : \boldsymbol{G}_i^A)$ *and* $\widehat{\boldsymbol{\phi}}$ *denote the ALS estimator of* $\boldsymbol{\phi}$, *then we have the following:*

$$\sqrt{T}(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}) \to N\left(\mathbf{0}, \Xi_2^*\right), \tag{15}$$

where $\Xi_2^* = \boldsymbol{G}\Xi_2\boldsymbol{G}'$ *and* $\boldsymbol{G} = \mathrm{diag}(\boldsymbol{G}_1, \dots, \boldsymbol{G}_K)$.

The proof of the theorem is in Appendix F of the Online Supplement and the corollary is in Appendix H of the Online Supplement.

4.3. Asymptotics for MLE Estimators

Now we provide the CLT of the MLE estimators $\widetilde{\boldsymbol{A}}_i$ s and $\widetilde{\boldsymbol{B}}_i$ s under the MMAR(p) model. We keep parts of the definitions the same as in the ALS part and let $\boldsymbol{\alpha}_i = \mathrm{vec}(\boldsymbol{A}_i)$, $\boldsymbol{\beta}_i = \mathrm{vec}(\boldsymbol{B}_i')$, then we have the following theorem.

Theorem 3. Consider the MMAR(p) model. Define $\boldsymbol{W}_{it}' = [\boldsymbol{W}_{it}^B \otimes \boldsymbol{I} : \boldsymbol{I} \otimes \boldsymbol{W}_{it}^A]$, $\widetilde{\boldsymbol{H}}_i = E(\boldsymbol{W}_{it}(\Sigma_{ic}^{-1} \otimes \Sigma_{ir}^{-1})\boldsymbol{W}_{it}') + \boldsymbol{C}_i\boldsymbol{C}_i'$, *and* $\Sigma^{(ij)} = \mathrm{Cov}(\mathrm{vec}(\boldsymbol{E}_{it})\mathrm{vec}(\boldsymbol{E}_{jt})')$. *Note that under the structure* *assumptions of* $\mathrm{Cov}(\mathrm{vec}(\boldsymbol{E}_{it}))$ *in (3) and additional assumption that* $\boldsymbol{E}_{it}$ *is uncorrelated with* $\boldsymbol{E}_{jt}$*, we have:*

$$\Sigma^{(ii)} = \Sigma_{ic} \otimes \Sigma_{ir}, \quad \Sigma^{(ij)} = \mathbf{0}, \quad \text{for } i,j = 1,\dots,K \text{ and } i \neq j.$$

Moreover, define a covariance matrix Ξ_3. Ξ_3 *has a structure similar to* Ξ_2*, and its* (ij)*th block takes the form*

$$\Xi_{3,(ij)} = \widetilde{\boldsymbol{H}}_i^{-1} E\left(\boldsymbol{W}_{it}(\Sigma^{(ii)})^{-1}\Sigma^{(ij)}(\Sigma^{(jj)})^{-1}\boldsymbol{W}'_{jt}\right)\widetilde{\boldsymbol{H}}_j^{-1}.$$

Taking the same conditions as in Theorem 2, it holds that:

$$\sqrt{T}\begin{pmatrix} \mathrm{vec}(\widetilde{\boldsymbol{A}}_1 - \boldsymbol{A}_1) \\ \mathrm{vec}(\widetilde{\boldsymbol{B}}'_1 - \boldsymbol{B}'_1) \\ \cdots \\ \mathrm{vec}(\widetilde{\boldsymbol{A}}_K - \boldsymbol{A}_K) \\ \mathrm{vec}(\widetilde{\boldsymbol{B}}'_K - \boldsymbol{B}'_K) \end{pmatrix} \to \mathcal{N}(\mathbf{0}, \Xi_3), \tag{16}$$

Define $\boldsymbol{\phi}_i = ((\mathrm{vec}(\boldsymbol{B}_{i1}^{(1)\prime}) \otimes \mathrm{vec}(\boldsymbol{A}_{i1}^{(1)}))', \dots, (\mathrm{vec}(\boldsymbol{B}_{iK}^{(1)\prime}) \otimes \mathrm{vec}(\boldsymbol{A}_{iK}^{(1)}))', \dots, (\mathrm{vec}(\boldsymbol{B}_{iK}^{(p)\prime}) \otimes \mathrm{vec}(\boldsymbol{A}_{iK}^{(p)}))')'$ for $i = 1,\dots,K$ and $\boldsymbol{\phi} = (\boldsymbol{\phi}'_1, \boldsymbol{\phi}'_2, \dots, \boldsymbol{\phi}'_K)'$. Based on the asymptotic properties of the MLE estimator shown in Theorem 3, we then have the following result.

Corollary 2. Given $\boldsymbol{G}_i$ *defined in Corollary 1, let* $\widetilde{\boldsymbol{\phi}}$ *denotes the MLE estimator of* $\boldsymbol{\phi}$*, then we have the following:*

$$\sqrt{T}(\widetilde{\boldsymbol{\phi}} - \boldsymbol{\phi}) \to N\left(\mathbf{0}, \Xi_3^*\right), \tag{17}$$

where $\Xi_3^* = \boldsymbol{G}\Xi_3\boldsymbol{G}'$ *and* $\boldsymbol{G} = \mathrm{diag}(\boldsymbol{G}_1, \dots, \boldsymbol{G}_K)$.

The proof of the theorem is in Appendix G of the Online Supplement and the corollary is in Appendix H of the Online Supplement.

5. Simulation

In the simulation study, we validate and compare the finite-sample performance of the three estimation methods and examine their asymptotic properties. Specifically, the simulations are designed to (*i*) compare estimation accuracy under different data-generating settings, (*ii*) assess relative asymptotic efficiency as the sample size increases, and (*iii*) evaluate the empirical coverage of confidence intervals implied by the asymptotic theory. Overall, the simulation results lead to several clear conclusions. First, across all data-generating settings, the proposed MMAR-based estimators substantially outperform the benchmark VAR estimator, highlighting the benefit of exploiting the underlying matrix structure. Second, compared with the projection estimator, both ALS and MLE achieve noticeably higher estimation accuracy, particularly as the sample size increases. Third, when the covariance structure is correctly specified, the MLE estimator is the most efficient, while under misspecification it remains competitive with ALS. Finally, the empirical coverage rates are close to the nominal level and improve with the sample size, providing finite-sample support for the asymptotic normality results established in Section 4. For detailed simulation designs and results, please refer to Appendix I of the Online Supplement.

Table 1. Regions in China, names, and abbreviations.

Region	Province & Municipalities(Abbreviation)
North China	Beijing (BJ), Tianjin (TJ), Hebei (HE), Shanxi (SX), Inner Mongolia (IM)
Northeast China	Liaoning (LN), Jilin (JL), Heilongjiang (HL)
East China	Shanghai (SH), Jiangsu (JS), Zhejiang (ZJ), Anhui (AH), Fujian (FJ), Jiangxi (JX), Shandong (SD)
Central China	Henan (HA), Hubei (HB), Hunan (HN)
South China	Guangdong (GD), Guangxi (GX), Hainan (HI)
Southwest China	Chongqing (CQ), Sichuan (SC), Guizhou (GZ), Yunnan (YN), Tibet (XZ)
Northwest China	Shaanxi (SN), Gansu (GS), Qinghai (QH), Ningxia (NX), Xinjiang (XJ)

6. An Empirical Application

In the previous sections, we discuss the setting, estimation, and asymptotic properties of the MMAR model and conduct simulations to verify the effectiveness of the proposed model and estimators. In this section, we will build on the basic ideas of the MMAR model to construct a complex macroeconomic system for various provinces and municipalities in China, examining the high-dimensional and intricate economic connections between regions, capturing economic changes, and exploring the impact of policy uncertainties and geopolitical risk on China's regional macroeconomic.

6.1. Data

The empirical analysis considers 31 provinces and municipalities in mainland China. Table 1 displays the provinces and municipalities under consideration, their English abbreviations, and the administrative regions to which they belong. Specifically, the 31 provinces and municipalities can be divided into seven different administrative regions, which include North China (5), Northeast China (3), East China (7), Central China (3), South China (3), Southwest China (5), and Northwest China (5).

We collect macroeconomic indicators reflecting the macroeconomic levels of each province and municipality, including gross domestic Product (GDP), inflation (CPI), real estate prices (REP), industrial added value (IAV), fixed asset investment (FAI), import (IMP) and export (EXP) values. Moreover, we collect several recently proposed key indicators to measure uncertainty, including China's economic policy uncertainty (EPU), trade policy uncertainty (TPU), and global geopolitical risk (GPR). China's EPU and TPU measurements are based on two mainland Chinese newspapers: the Renmin Daily and the Guangming Daily (Baker et al. 2013). The global geopolitical risk is calculated using 10 US newspapers: the Chicago Tribune, the Daily Telegraph, Financial Times, The Globe and Mail, The Guardian, the Los Angeles Times, The New York Times, USA Today, The Wall Street Journal, and The Washington Post. Caldara and Iacoviello (Caldara and Iacoviello 2022), by determining the proportion of news articles related to negative geopolitical events in each newspaper per month (relative to the total number of news articles). Detailed description of the provincial macroeconomic indicators are in the Appendix J of the Online Supplement. All data starts with January 31, 2001 to December 31, 2022.

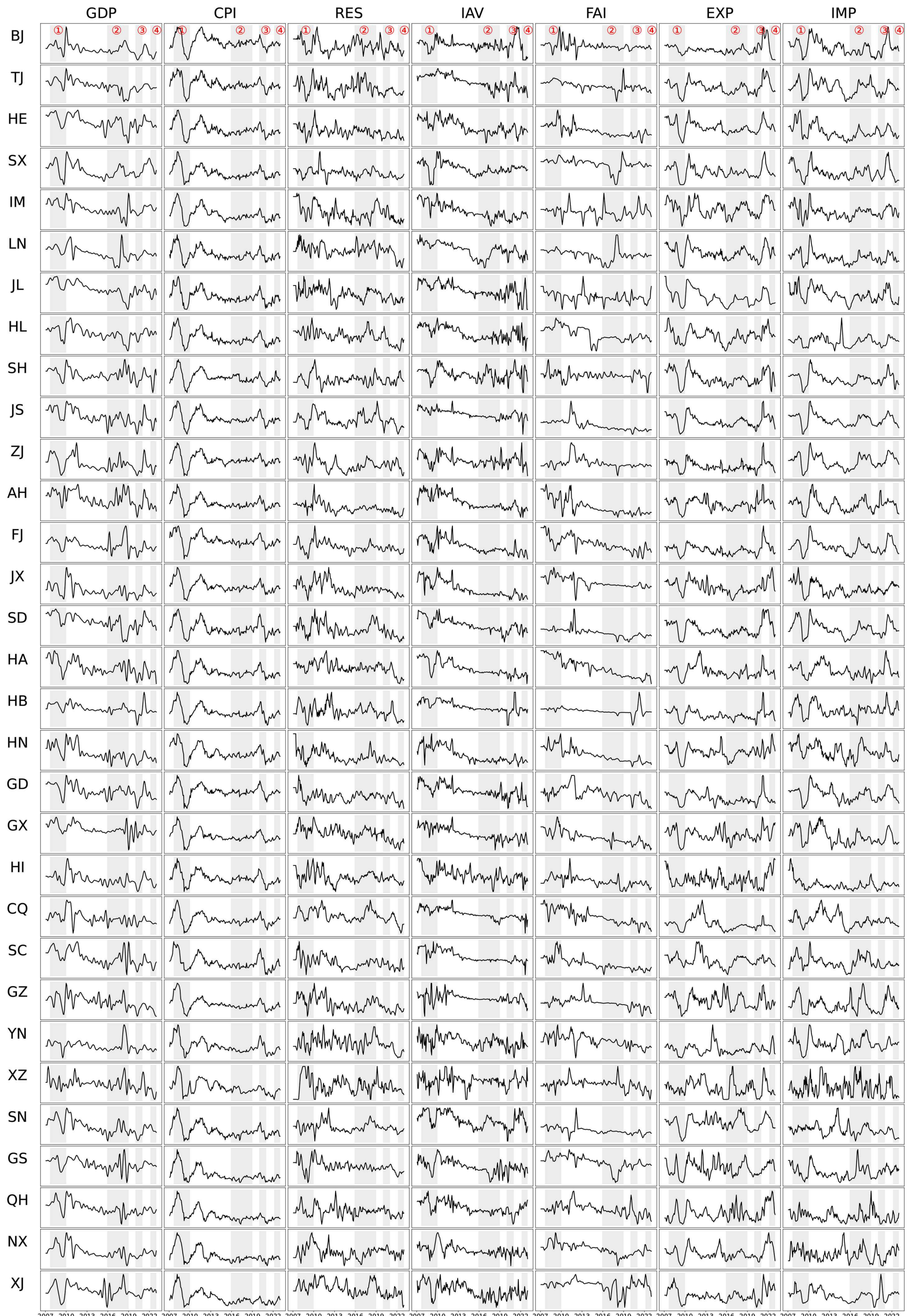


Figure 4. Matrix time series plot of Chinese provincial macroeconomic indicators.
Notes: Rows represent the 31 provinces or municipalities, and columns represent the 7 macroeconomic indicators. We highlight four important periods: (1) Global Financial Crisis: from 2007–2008 to 2009–12. (2) Supply-side reform: from 2015–11 to 2018–12. (3) COVID-19 pandemic: from 2019–12 to 2020–12. (4) Russia-Ukraine conflict: from 2022–02 to 2022–12, the last month of the sample.

The provincial data of interest are obtained from the macroeconomic databases of the Wind financial terminal and the Choice financial terminal.[3] We obtain the monthly year-over-year (YoY) growth rate for the seven province-level indicators. The YoY growth rate g_t is calculated by $g_t = (z_t - z_{t-12})/z_{t-12} \times 100$, where z_t denotes the total value in month t. Specifically, we make a series of transformations to the original data: (1) The housing price index is based on the ratio of the total sales amount of commercial housing and the sales area of commercial housing in each province per month. (2) The original GDP is quarterly data, and we perform a second-order spline interpolation on the seasonally adjusted quarterly YoY GDP growth to obtain the monthly YoY GDP growth. (3) Due to the strong short-term fluctuations and seasonal effects on the housing price index, fixed asset investment, export values, and import values, we calculate their growth rates year-over-year after applying a three-month moving average.

[3]The website for Wind is https://www.wind.com.cn, and the Choice website is https://choice.eastmoney.com

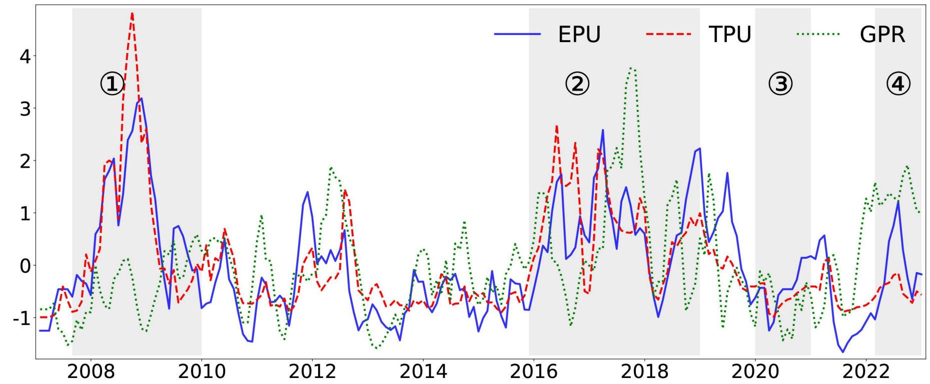


Figure 5. Time series plot of three uncertainty indicators.*Notes:* The blue solid line represents EPU (Economic Policy Uncertainty), the red dashed line represents TPU (Trade Policy Uncertainty), and the green dotted line represents GPR (Geopolitical Risk). We highlight four important periods: (1) Global Financial Crisis: from 2007–08 to 2009–12. (2) Supply-side Reform: from 2015–11 to 2018–12. (3) COVID-19 Pandemic: from 2019–12 to 2020–12. (4) Russia-Ukraine Conflict: from 2022–02 to 2022–12, the last month of the sample.

Figure 4 of the main paper plots the matrix time series of Chinese provincial macroeconomic indicators, where rows represent the 31 provinces or municipalities, and columns represent the seven macroeconomic indicators. The sample period spans from January 2007 to December 2022, with a total of 192 data points. There are clear similarities among the same macroeconomic indicators across different provinces, which point to a column-wise structure. This demonstrates the necessity of constructing matrix time series to capture these differences. Figure 5 of the main paper shows the time series of three uncertainty indices, the EPU, the TPU, and the GPR. The time range is consistent with the provincial-level indicators. Four important periods are marked in the figure. Detailed discussions of the data are in the Appendix J of the Online Supplement.

6.2. Estimating Multi-Modal Connectedness

In this section, we construct multi-modal connectedness networks based on the estimated MMAR model. Our approach builds on the connectedness measures of Diebold and Yilmaz (2009, 2012); Diebold and Yılmaz (2014), which quantify directional spillovers through forecast error variance decompositions in a VAR framework, and adapts them to the MMAR setting by exploiting the row-column structure of matrix-valued data.

The construction proceeds as follows. First, we estimate the MMAR(p) model. Second, we recover its implied restricted VAR(p) representation by assembling the coefficient matrices $\hat{\Phi}_\ell$, where each block satisfies $\hat{\Phi}_{\ell,ij} = \hat{\boldsymbol{B}}_{ij}^{(\ell)} \otimes \hat{\boldsymbol{A}}_{ij}^{(\ell)}$, together with the estimated innovation covariance matrix $\hat{\Sigma}$. Third, based on this estimated VAR(p), we obtain the associated VMA(∞) representation and apply the generalized forecast error variance decomposition (GFEVD) to construct variable-level connectedness measures, which are then aggregated along the row and column dimensions of the MMAR model to form multi-modal connectedness networks.

For notational simplicity, we suppress the hat notation below. Consider the VMA(∞) representation $\boldsymbol{y}_t = \sum_{h=0}^{\infty} \Psi_h \boldsymbol{\varepsilon}_{t-h}$, then the contribution of variable j to the H-step-ahead forecast error variance of variable i is

$$(\mathcal{C}_H)_{i\leftarrow j} = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} \left((\Psi_h \Sigma)_{i,j}\right)^2}{\sum_{h=0}^{H-1} (\Psi_h \Sigma \Psi_h')_{i,i}}, \tag{18}$$

where Σ is the covariance matrix of $\boldsymbol{\varepsilon}_t$. Since $\sum_{j=1}^{N} (\mathcal{C}_H)_{i\leftarrow j}$ does not generally equal one, we normalize $(\mathcal{C}_H)_{i\leftarrow j}$ to obtain the pairwise directional connectedness $(\widetilde{\mathcal{C}}_H)_{i\leftarrow j} = (\mathcal{C}_H)_{i\leftarrow j} / \sum_{j=1}^{N} (\mathcal{C}_H)_{i\leftarrow j}$.

While the variable-level measures provide a complete description of spillovers across all N variables, direct interpretation becomes difficult in high-dimensional systems. To exploit the row-column structure of matrix-valued data, we aggregate these measures to obtain group-wise multi-modal connectedness. Suppose the N variables are partitioned into S groups $\{G_1, \ldots, G_S\}$ with N_s variables in group s. The directional connectedness from group l to group k is defined as

$$(\mathcal{C}_H^g)_{k\leftarrow l} = \frac{1}{N_k N_l} \sum_{i\in G_k} \sum_{j\in G_l} (\mathcal{C}_H)_{i\leftarrow j}, \tag{19}$$

and is normalized analogously to obtain $\widetilde{\mathcal{C}}_H^g$. The corresponding net, "To", and "From" connectedness measures are defined in the usual manner.

By aggregating connectedness along different row or column dimensions of the MMAR model, we obtain multiple layers of group-level connectedness, including indicator-level, provincial-level, and regional measures. Together, these constitute the multi-modal connectedness networks reported below.

6.3. Network Analysis

The empirical analysis is based on a matrix-vector MMAR(1) specification with $K = 2$. Specifically, letting $\boldsymbol{X}_{1,t}$ denote the 31×7 provincial macroeconomic matrix and $\boldsymbol{X}_{2,t} = \boldsymbol{x}_{2,t}$ denote the 3×1 vector of uncertainty indicators, the matrix-vector model is specified as

$$\begin{aligned}\boldsymbol{X}_{1,t} &= \boldsymbol{A}_{11}\boldsymbol{X}_{1,t-1}\boldsymbol{B}_{11}' + \boldsymbol{A}_{12}\boldsymbol{x}_{2,t-1}\boldsymbol{b}_{12}' + \boldsymbol{E}_{1,t}, \\ \boldsymbol{x}_{2,t} &= \boldsymbol{A}_{21}\boldsymbol{X}_{1,t-1}\boldsymbol{b}_{21} + \boldsymbol{A}_{22}\boldsymbol{x}_{2,t-1} + \boldsymbol{e}_{2,t}.\end{aligned} \tag{20}$$

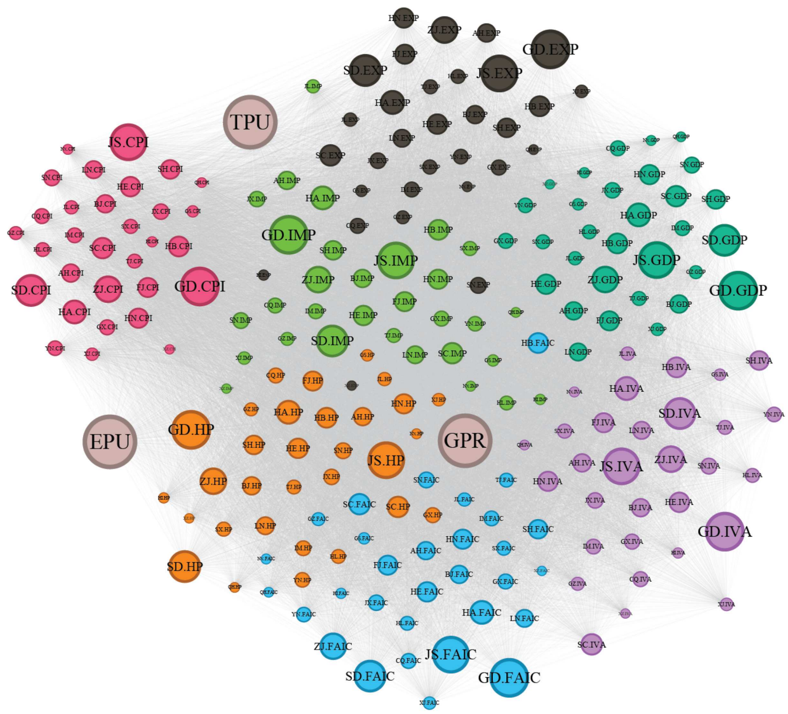


Figure 6. Individual-level connectedness network.
Notes: Each node in the graph represents the economic indicators of different provinces and the national-level economic uncertainty indices. The different colors of the nodes represent different types of indicators, while the size of the nodes reflects the average GDP size of different provinces during the sample period. The closer the nodes are, the stronger their risk association.

The model is estimated using the ALS procedure described in Section 3.2. Before turning to network-based summaries, we first examine the estimated MMAR(1) coefficient and covariance matrices. These results reveal a economically interpretable block-wise structure that are absent in an unrestricted VAR model of the same dimension. A detailed discussion of the coefficient estimates and their economic implications is provided in Appendix K of the Online Supplement.

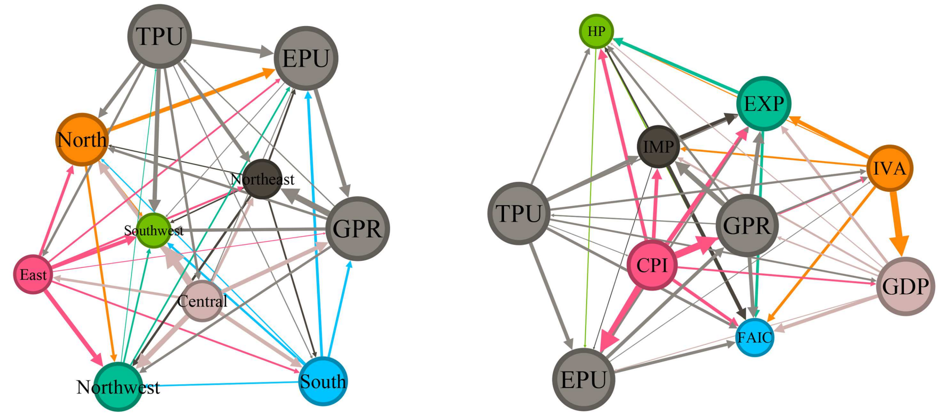
The total of $N = 220$ variables makes the traditional analysis of impulse response functions or variance decompositions considerably challenging due to the substantial dimensionality and intricate interrelationships. In particular, approximately 48,000 pairwise directional connectedness among the 220 variables render a direct examination of individual relationships impractical. Given the scale and complexity of the system, it is therefore necessary to introduce network-based methods to further summarize and interpret the high-dimensional dependence structure. We construct high-dimensional connectedness networks to capture the key relationships among regional economic variables, and the detailed construction of the multi-modal network is presented in Appendix L of the Online Supplement. In the subsequent analysis, we build on the estimated coefficients of MMAR(1) and the covariance matrix and set the forecast horizon of the generalized variance decomposition at $H = 12$.

6.3.1. *Individual-Level Connectedness Network*

We first present the connectedness network graph at the individual level, which includes 220 nodes, 24,090 connections, and 48,180 arrows. As shown in Figure 6, we summarize several important findings. First, the individual-level connectedness network exhibits a strong clustering effect. The nodes representing the same macroeconomic indicators (the same color) are generally clustered in the network, indicating a strong connectedness among the same economic indicators in different provinces. Second, by analyzing the distances between the clusters of various indicators, we find that GDP (dark green) is adjacent to imports (light green), exports (black), and industrial added value (purple), suggesting that China's economic growth is closely related to its domestic industrial conditions and imports and exports. This might also explain why the "dual circulation" strategy has been repeatedly mentioned recently. CPI (red) is connected to imports, exports, and real estate (orange), indicating that inflation has a stronger impact on these economic indicators in China. Finally, examining the relative positions of TPU, EPU, and GPR in the network reveals some interesting findings: (1) TPU is closer to CPI, imports, and exports, suggesting that changes in China's trade policy are closely related to the CPI and import-export situation; (2) EPU is close to CPI and real estate prices, reflecting that changes in China's economic policies have a significant connection with real estate and inflation; (3) Compared to EPU and TPU, GPR is located at the center of the network. This central position suggests that China's geopolitical risk plays a crucial role within the network and is closely associated with the general economic conditions of the country.

6.3.2. *Regional-Level and Indicator-Level Net Connectedness Network*

We further delve into the net connectedness networks at regional and indicator levels to analyze the roles of net transmitter and net receiver that each region or indicator assumes within the connectedness network. Figure 7a exhibits the regional-level



(a) Regional level (b) Indicator level

Figure 7. Multi-modal net connectedness network. (a) Regional level; (b) indicator level.
Notes: Each node in the graph represents different regions (indicators) and the national-level economic uncertainty indices. The different colors of the nodes represent different regions and uncertainty indices, and the size of the nodes reflects the average GDP of different regions during the sample period (since the uncertainty indices are national-level indicators, the EPU, TPU, and GPR nodes are the largest). The closer the nodes are, the stronger the pairwise connectedness from the uncertainty index to the province.

connectedness network, and we sequentially present several key observations: (1) EPU serves as net receivers for all the regions. This indicates that, compared to the directional connectedness from uncertainty indices to regional economies, the directional connectedness from China's economic policy to regional economies is more pronounced. This highlights that China's macroeconomic regulation follows a problem-solving approach after identifying pertinent issues. Furthermore, the TPUs serve as net transmitters to all regions, indicating that trade uncertainty fluctuations have a significant impact on regional economic conditions in China. (2) TPU, EPU, and GPR form an internal triangular circuit, with TPU pointing to EPU, EPU pointing to GPR, and eventually GPR pointing back to TPU. This pattern implies that trade policy changes influence economic policy changes, subsequently impacting geopolitical risk fluctuations. Ultimately, geopolitical risk changes are transmitted back to trade policy. (3) By comparing the net connectedness across various regions, it becomes evident that the eastern and central region serves as a net transmitter for all other regions, signifying its strong influence on China's regional economy. While the northwest and northeast regions function primarily as net receivers, they still maintain a net transmission of connectedness to the EPU, indicating their continued substantial impact on economic policy uncertainty.

Upon examining the net connectedness network at the indicator level (Figure 7b), we would like to emphasize several key findings: First, EPU and TPU consistently act as a net receiver and a net transmitter of connectedness to other macroeconomic indicators. The CPI (red) is notable for exerting a significant net transmission effect on both the EPU and the GPR, while GDP shows a slightly net transmission on the EPU and GPR. These results highlight the important impact of inflation on China's economic and political stability. Secondly, when analyzing various macroeconomic indicators, it becomes evident that the industrial added value (orange) is a net transmitter of connectedness for all indicators, excluding the CPI. Following IAV, GDP and CPI are also found to be strong net transmitters. This evidence underscores the continued prominence of the industrial sector within China's macroeconomic framework.

7. Conclusion

In this study, we introduce an innovative multi-matrix autoregressive (MMAR) model. The proposed model exploits the inherent matrix structure within the data and effectively captures contemporaneous and intertemporal relationships among matrices of varying dimensions, offering a wide range of empirical applications. We illustrate the application landscape of the proposed MMAR model and discuss potential extensions. We propose three different estimation methodologies, examining the statistical properties of the associated estimators. We present simulation studies that demonstrate the versatility and applicability of the proposed model in various cases to substantiate these properties. In the empirical study, we examine the interrelationships between China's macroeconomic conditions with some uncertainty indicators of China.

Our work extends the matrix AR model to a matrix "VAR" framework, thus opening up a vast potential for further expansion and broad applicability. Regarding model expansion, researchers could consider multi-matrix factor models or multi-matrix network models or consider imposing sparsity assumptions on the coefficient matrices of multi-matrix models. The ideas can also be extended more broadly to the field of tensors. Empirically, the research framework proposed in this paper includes group-level connectedness measures and multi-modal network construction methods, which can be applied to various empirical research scenarios, including studying international trade networks of multiple commodities and interactions among various countries' stock markets.

Acknowledgments

We would like to thank Brendan K. Beare, Peng Bin, Jamie Cross, Rong Chen, Jiti Gao, Whitney Newey, Peter C. B. Phillips, Thomas J. Sargent, Zhengjun Zhang, and participants in the following conferences for useful discussions and comments: 31st Australia New Zealand Econometric Study Group Meeting in Adelaide; Time Series and Forecasting Symposium 2023 in Sydney; 33rd Annual Conference of Chinese Economics Society Australia in Sydney; International Symposium on Advance in Panel Data and Time Series Econometrics in Honor of Cheng Hsiao's 80th Anniversary in Xiamen; 7th IAER Econometrics Workshop in Dalian; and 2023 Symposium on Recent Developments in Time Series Econometrics and Applied Macroeconometrics in Xiamen.

Disclosure Statement

The authors report there are no competing interests to declare.

Funding

S. Ye is supported by the National Science Foundation of China (NSFC) [Grant No. 23B2020], the China Postdoctoral Science Foundation [Grant Nos. 2024M763465, 2025T181016, GZC20252267], and the National Social Science Fund of China under the Social Science Academic Societies Program [Grant No. 25SGC066]. T. Zheng is supported by the National Social Science Fund of China Major Project (Grant No. 23&ZD074). Y. Hong and S. Ye are supported by the MOE Social Sciences Innovative Group on Complex Systems Modeling in Economic Management in the Era of Digital Intelligence, University of Chinese Academy of Sciences [Grant No. E5820801] and NSFC [Grant No. 71988101].

ORCID

Shiqi Ye http://orcid.org/0009-0008-2208-5742
Yongmiao Hong http://orcid.org/0000-0002-9711-9061

References

Armillotta, M., Gorgi, P., and Lucas, A. (2025), Copula Tensor Count Autoregressions for Modeling Multidimensional Integer-Valued Time Series, Technical Report, Tinbergen Institute Discussion Paper. [1]

Baker, S. R., Bloom, N., and Davis, S. J. (2016), "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics*, 131, 1593–1636. [2]

Baker, S., Bloom, N., Davis, S. J., and Wang, X. (2013), *A Measure of Economic Policy Uncertainty for China*, Chicago: Work-in-Progress, University of Chicago. [2,9]

Barigozzi, M., He, Y., Li, L., and Trapani, L. (2023), "Robust Estimation of Large Factor Models for Tensor-Valued Time Series," arXiv Preprint arXiv:2303.18163. [1]

Billio, M., Casarin, R., Iacopini, M., and Kaufmann, S. (2023), "Bayesian Dynamic Tensor Regression," *Journal of Business & Economic Statistics*, 41, 429–439. DOI: 10.1080/07350015.2022.2032721. [1]

Brockwell, P. J., Davis, R. A., Brockwell, P. J., and Davis, R. A. (1991), "Stationary Time Series," in *Time Series: Theory and Methods*, pp. 1–41, New York: Springer. DOI: 10.1007/978-1-4419-0320-4$_1$. [5]

Caldara, D., and Iacoviello, M. (2022), "Measuring Geopolitical Risk," *American Economic Review*, 112, 1194–1225. DOI: 10.1257/aer.20191823. [2,9]

Chan, J. C., and Qi, Y. (2023), Large Bayesian Matrix Autoregressions. [1]

Chen, E. Y., and Fan, J. (2023), "Statistical Inference for High-Dimensional Matrix-Variate Factor Models," *Journal of the American Statistical Association*, 118, 1038–1055. DOI: 10.1080/01621459.2021.1970569. [1]

Chen, E. Y., Tsay, R. S., and Chen, R. (2020), "Constrained Factor Models for High-Dimensional Matrix-Variate Time Series," *Journal of the American Statistical Association*, 115, 775–793. DOI: 10.1080/01621459.2019.1584899. [1]

Chen, R., Han, Y., Li, Z., Xiao, H., Yang, D., and Yu, R. (2022), "Analysis of Tensor Time Series: Tensorts," *Journal of Statistical Software*, 323, 1–41. [1]

Chen, R., Xiao, H., and Yang, D. (2021), "Autoregressive Models for Matrix-Valued Time Series," *Journal of Econometrics*, 222, 539–560. DOI: 10.1016/j.jeconom.2020.07.015. [1,3,4,5]

Chen, R., Yang, D., and Zhang, C.-H. (2022), "Factor Models for High-Dimensional Tensor Time Series," *Journal of the American Statistical Association*, 117, 94–116. DOI: 10.1080/01621459.2021.1912757.

Chen, X., Yang, D., Xu, Y., Xia, Y., Wang, D., and Shen, H. (2023), "Testing and Support Recovery of Correlation Structures for Matrix-Valued Observations with an Application to Stock Market Data," *Journal of Econometrics*, 232, 544–564. [3]

Diebold, F. X., and Yilmaz, K. (2009), "Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets," *Economic Journal*, 119, 158–171. DOI: 10.1111/j.1468-0297.2008.02208.x. [11]

Diebold, F. X., and Yilmaz, K.. (2012), "Better to Give than to Receive: Predictive Directional Measurement of Volatility Spillovers," *International Journal of Forecasting*, 28, 57–66. DOI: 10.1016/j.ijforecast.2011.02.006. [11]

Diebold, F. X., and Yılmaz, K. (2014), "On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms," *Journal of Econometrics*, 182, 119–134. DOI: 10.1016/j.jeconom.2014.04.012. [3,11]

Diebold, F. X., and Yılmaz, K.. (2023), "Reprint Of: On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms," *Journal of Econometrics*, 234, 70–90. DOI: 10.1016/j.jeconom.2023.03.003. [3]

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton, NJ: Princeton University Press. [8]

Han, Y., Chen, R., Zhang, C.-H., and Yao, Q. (2024), "Simultaneous Decorrelation of Matrix Time Series," *Journal of the American Statistical Association*, 119, 957–969. DOI: 10.1080/01621459.2022.2151448. [1]

He, Y., Kong, X., Trapani, L., and Yu, L. (2023), "One-Way or Two-Way Factor Model for Matrix Sequences?," *Journal of Econometrics*, 235, 1981–2004. DOI: 10.1016/j.jeconom.2023.02.008. [1]

Hoff, P. D. (2015), "Multilinear Tensor Regression for Longitudinal Relational Data," *Annals of Applied Statistics*, 9, 1169–1193. DOI: 10.1214/15-AOAS839. [1]

Hsu, N.-J., Huang, H.-C., and Tsay, R. S. (2021), "Matrix Autoregressive Spatio-Temporal Models," *Journal of Computational and Graphical Statistics*, 30, 1143–1155. DOI: 10.1080/10618600.2021.1938587. [1]

Li, Z., and Xiao, H. (2021), "Multi-Linear Tensor Autoregressive Models," arXiv Preprint arXiv:2110.00928. [1,6]

Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Berlin, Heidelberg: Springer Science & Business Media. [8]

——— (2005), *New Introduction to Multiple Time Series Analysis*, Berlin, Heidelberg: Springer Science & Business Media. [1]

Ren, Y., Zhu, X., and Ma, Y. (2022), "Matrix-Valued Network Autoregression Model with Latent Group Structure," arXiv Preprint arXiv:2212.02107. [1]

Sims, C. A. (1980), "Macroeconomics and Reality," *Econometrica*, 48, 1–48. DOI: 10.2307/1912017. [3]

Tsay, R. S. (2014), *An Introduction to Analysis of Financial Data with R*, Hoboken, NJ: John Wiley & Sons. [5]

Van Loan, C. F. (2000), "The Ubiquitous Kronecker Product," *Journal of Computational and Applied Mathematics*, 123, 85–100. DOI: 10.1016/S0377-0427(00)00393-9. [6]

Van Loan, C. F., and Pitsianis, N. (1993), "Approximation with Kronecker Products," in *Linear Algebra for Large Scale and Real-Time Applications*, Dordrecht: Springer, pp. 293–314. [6]

Wang, D., Liu, X., and Chen, R. (2019), "Factor Models for Matrix-Valued High-Dimensional Time Series," *Journal of Econometrics*, 208, 231–248. DOI: 10.1016/j.jeconom.2018.09.013. [1]

Xiao, H., Han, Y., Chen, R., and Liu, C. (2022), "Reduced Rank Autoregressive Models for Matrix Time Series," *Journal of Business & Economic Statistics*. [1]

Yu, C., Li, D., Jiang, F., and Zhu, K. (2025), "Matrix Garch Model: Inference and Application," *Journal of the American Statistical Association*, 120, 1747–1762. DOI: 10.1080/01621459.2024.2415719. [1]

Zhu, X., Wang, F., Li, Z., and Ma, Y. (2023), "Network Autoregression for Incomplete Matrix-Valued Time Series," arXiv Preprint arXiv:2302.02768. [1]