A Spatio-temporal Autoregressive Factor Model of the Global Business Cycle*

TOMOHIRO ANDO
Melbourne Business School, University of Melbourne, Carlton, VIC 3053, Australia
T.Ando@mbs.edu

MATTHEW GREENWOOD-NIMMO
Faculty of Business and Economics, University of Melbourne, Carlton, VIC 3053, Australia
Centre for Applied Macroeconomic Analysis, Australian National University, Canberra, ACT 2600, Australia
matthew.greenwood@unimelb.edu.au

YONGCHEOL SHIN
Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, U.K.
yongcheol.shin@york.ac.uk

CHAOWEN ZHENG
Department of Economics, University of Reading, U.K.
chaowen.zheng@reading.ac.uk

Abstract

[Edit??] To study the synchronicity of national business cycles, we propose a new heterogeneous-parameter approach in which the global business cycle is modelled as a spatio-temporal autoregressive process with a common factor error structure. To achieve consistent estimation in the presence of parameter heterogeneity and endogeneity, we develop a modified quasi maximum likelihood estimation approach. We show that the resulting estimators are consistent and asymptotically normally distributed. We employ Monte Carlo simulations to demonstrate that their finite sample performance is satisfactory. Based on the proposed estimator, we further develop network analysis tools at both individual and regional level using diffusion FEVDs and multipliers. These tools are then applied to analyse the business cycle synchronisation covering 79 countries in the world over the period 1970-2019 (50 years).

*Corresponding author: Yongcheol Shin. We are mostly grateful for the insightful comments by Jia Chen, Rui Lin, Massimiliano Marcellino, Camilla Mastromacro, Laura Serlenga, Michael Thornton, and the seminar participants at the JRC at European Commission, and Bocconi University, Italy. Greenwood-Nimmo and Shin are grateful for partial financial support from the Economic and Social Research Council (ESRC) under grant number ES/T01573X/1. The ESRC had no involvement in the research design, the construction, analysis or interpretation of the dataset, the write-up of the paper or the decision to submit to this journal for publication. The usual disclaimer applies.

Key Words: Spatio-temporal Autoregressive Factor (STARF) Model; Parameter Heterogeneity; Network Topology; Global Business Cycle.

JEL Classifications: C23, E32, F41.

1 Introduction

The apparent synchronicity of business cycles across countries is typically modelled in one of two ways. First, it may be treated as the result of strong cross-section dependence (CSD), where a finite number of global and potentially regional factors induce a degree of common behaviour in national business cycles (e.g. Kose et al., 2003, 2008; Hirata et al., 2013). Second, it may be considered as the result of weak CSD, where business cycle comovement arises due to spatial linkages among countries, with spatial weights typically being constructed using either bilateral trade statistics or data on bilateral financial linkages (e.g. Pesaran et al., 2004; Dees et al., 2007; Greenwood-Nimmo et al., 2021). Servén and Abate (2020) were the first to unify these two approaches using a factor-augmented dynamic spatial growth model. A limitation of their approach is the underlying assumption that the slope parameters are homogeneous across countries. This may be a strong assumption in practice, given the marked heterogeneity of national economies. In this paper, we remove the slope homogeneity assumption and develop a new heterogeneous-parameter approach in which the global business cycle is modelled as a spatio-temporal autoregressive process with a common factor error structure.

Servén and Abate approach this problem indirectly via sub-sample estimation but this introduces problems in its own right, notably a loss of efficiency and potentially inaccurate estimation. In this paper, we address the problem directly and propose

considering the stylised fact that output growth is also positively autocorrelated (Nelson and Plosser, 1982).

Moreover, while heterogeneous model include homogeneous model as special case, mistaken heterogeneous as homogeneous could also lead to inconsistent results (Pesaran and Smith (1995), Chen et al. (2021)).

Several attempts have been made to develop a unified characterization of CSD that combines both spatial and factor dependence in recent years. Bai and Li (2014) consider a homogeneous spatial panel data model with common shocks and develop a quasi maximum likelihood (QML) estimation framework. Bailey et al. (2016) develop a multi-step estimation procedure that can distinguish weak CSD that is purely spatial from strong CSD due to the influence of common factors. Mastromarco et al. (2016) propose a technique for modelling stochastic frontier panels by combining the exogenously driven factor-based approach and an endogenous threshold regime selection mechanism. Yang (2021) develops consistent estimators that combine common correlated effects (CCE) and instrumental variable (IV)/generalised method of moments

(GMM) estimation.

All of the studies described above share the assumption of slope homogeneity. In many practical situations, this may prove to be an untenable assumption, as the strength and even the sign of spatial dependence may vary over spatial units. Furthermore, in a data-rich environment, the trade-off between the parsimony that slope homogeneity restrictions offer and the realism that slope heterogeneity offers is likely to favor the latter. Several studies have developed methods for the estimation of heterogeneous parameter models in the spatial literature. Ando et al. (2021d) consider a heterogeneous parameter panel spatial quantile model with interactive effects to accommodate both spatial dependence and factor dependence. For the linear spatial model, Aquaro et al. (2021) and Shin and Thornton (2021) (henceforth ST) develop QML and control function-based estimators (see also LeSage and Chih, 2018). Chen et al. (2021) develop a unifying framework for the analysis of static heterogeneous panel data models that accommodate both spatial dependence and factor dependence.

Following this research trend, our primary objective is to develop a unifying econometric framework for the estimation of dynamic heterogeneous panel data models that can jointly accommodate the spatial and factor dependence. Following ST, we christen our approach the spatio-temporal autoregressive model with unobserved factors (STARF). To tackle the challenging issues for developing consistent estimation in the presence of spatial heterogeneity and endogeneity caused by the spatial lagged term and the correlation between the regressors and unobserved factors, we propose the modified QML methods.

The main aim of this paper is twofold. First, as an input, we derive the consistent individual STARF estimator and establish its asymptotic properties. Then, as an output, we develop the network analysis using the diffusion FEVDs and multipliers. For the practical network-oriented approach we develop, we need only $\sqrt{T}$-consistent estimators of the individual heterogeneous parameters. Our approach contrasts with the pooled or mean-group versions of the estimators, which have been routinely proposed in the main panel data literature. In many applications, there is no economic reason to expect the coefficients of the model to share a common sign (e.g. Masten (2018)). A mean group estimator subject to such netting off has the potential to produce a misleading global picture and failing to examine the relative importance of individual nodes beyond that pre-supposed by the spatial weights matrix, $\boldsymbol{W}$.

Our novel system diffusion FEVDs/multipliers populate a sequence of network connectedness matrices with interpretation as output network matrices capturing spatial/temporal diffusion. Starting with an input network matrix, $\boldsymbol{W}$, the output network matrices resulting from the combined interactions estimated through

our STARF model track the evolving influence of individual nodes on one another. To provide informative network analysis, we also develop summary statistics measuring the connectedness at both individual and aggregated regional level following Diebold and Yılmaz (2014) and Greenwood-Nimmo et al. (2021), Further, two indices quantifying the relative importance of each individual/region are also proposed following Shin and Thornton (2021).

Through conducting extensive Monte Carlo simulations, we find the finite sample performance of our proposed estimator are quite satisfactory: the biases are generally negligible with RMSEs decreasing with both N and T; the sizes are all close to the nominal 5% significance level with powers approaching one quickly as sample increases. We also demonstrate the usefulness of our approach with an application to the business cycle synchronisation covering 79 countries in the world over the period 1970-2019 (50 years). Our main findings are summarised as follows: (i) Both the individual dynamic and contemporary spatial coefficients are quite heterogeneous, but overall positive. These confirm our *a priori* expectation that the output growth is both serially and cross-sectionally positively correlated. (ii) From individual network analysis, we find US, China and Germany to be among the largest transmitters of both country-specific shock and total factor productivity (TFP) growth effects. While some other major economies (e.g. France, Italy and Canada) are also exerting both large country-specific shock and TFP growth effects, their net effects are however negative due to their large spillins. (iii) On the aggregated level, the geographical regions: Northern America and Eastern Asia, and the economic groups: High-income and G7 countries, are identified to be the most influential regions in terms of transmissions of both the regional shock and TFP growth effects.

The remainder of this paper is organised as follows. Section 2 describes the model and the main estimation algorithms. Section 3 develops the asymptotic theory. The network analysis is described in Section 4. Section 5 presents the finite sample performance of the proposed estimator. Section 6 provides an empirical application. Section 7 concludes the paper. Mathematical proofs and additional simulation and empirical results are presented in a separate Online Supplement.

Notational Conventions. C represents a positive constant. For any $N \times N$ real matrix, $\boldsymbol{A} = (a_{ij})$, $\|\boldsymbol{A}\| = \sqrt{tr(\boldsymbol{A}\boldsymbol{A}')}$ denotes the Frobenius norm. The row and column sum norms of $\boldsymbol{A}$ are defined as $\|\boldsymbol{A}\|_\infty = \underset{1\leq i\leq N}{max} \sum_{j=1}^{N} |a_{ij}|$ and $\|\boldsymbol{A}\|_1 = \underset{1\leq j\leq N}{max} \sum_{i=1}^{N} |a_{ij}|$. We denote $\mathrm{vec}(\boldsymbol{A})$ as the vectorisation operator that stacks the columns of $\boldsymbol{A}$ into a column vector, $\lambda_{\max}(\boldsymbol{A})$ as the largest eigenvalue of $\boldsymbol{A}$, and $\mathrm{tr}(\boldsymbol{A})$ as the trace of $\boldsymbol{A}$. We use $\otimes$ as the Kronecker product operator, $a \sim b$ representing that a and b are equivalent

in the order of magnitude, and $(N,T)\to\infty$ implying that N and T tend to infinity jointly.

2 The Model and Estimation

2.1 STARF Model

Consider the following heterogeneous STARF model:

$$y_{it} = \rho_i y_{it-1} + \phi_{0i} y^*_{it} + \phi_{1i} y^*_{it-1} + \boldsymbol{\beta}'_i \boldsymbol{x}_{it} + \boldsymbol{\lambda}'_i \boldsymbol{f}_t + \varepsilon_{it}, \tag{1}$$

where y_{it} is the scalar dependent variable of the ith spatial unit at time t and y_{it-1} is its time lag. y^*_{it} is the spatial/network variable, defined as $y^*_{it} = \sum_{j=1}^{N} w_{ij} y_{jt} = \boldsymbol{w}_i \boldsymbol{y}_t$, with $\underset{N\times 1}{\boldsymbol{y}_t} = (y_{1t}, \ldots, y_{Nt})'$ and with $\boldsymbol{w}_i = (w_{i1}, \ldots, w_{iN})$ being a $1 \times N$ vector of spatial weights determined *a priori*, with $w_{ii} = 0$. Similarly, $y^*_{i,t-1}$ is the diffusion variable, which is defined as $y^*_{i,t-1} = \sum_{j=1}^{N} w_{ij} y_{j,t-1} = \boldsymbol{w}_i \boldsymbol{y}_{t-1}$ with $\underset{N\times 1}{\boldsymbol{y}_{t-1}} = (y_{1,t-1}, \ldots, y_{N,t-1})'$. The coefficients ρ_i, ϕ_{0i} and ϕ_{1i} are the heterogeneous parameters corresponding to y_{it-1}, y^*_{it}, and $y^*_{i,t-1}$, respectively. $\boldsymbol{x}_{it} = (x_{it,1}, \ldots, x_{it,k})'$ is a $k \times 1$ vector of exogenous variables, $\boldsymbol{\beta}_i$ is a $k \times 1$ vector of heterogeneous parameters, $\boldsymbol{f}_t$ is an $r \times 1$ vector of unobserved factors with factor loadings $\boldsymbol{\lambda}_i$, and ε_{it} is an idiosyncratic disturbance.

Stacking the individual STARF regressions (1) over the i dimension, we obtain the following system representation:

$$\boldsymbol{y}_t = \boldsymbol{P}\boldsymbol{y}_{t-1} + \boldsymbol{\Phi}_0 \boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{\Phi}_1 \boldsymbol{W}\boldsymbol{y}_{t-1} + \boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{\Lambda}\boldsymbol{f}_t + \boldsymbol{\varepsilon}_t, \tag{2}$$

where we define $\boldsymbol{P} = diag(\rho_1, \ldots, \rho_N)$, $\boldsymbol{\Phi}_0 = diag(\phi_{01}, \ldots, \phi_{0N})$, $\boldsymbol{\Phi}_1 = diag(\phi_{11}, \ldots, \phi_{1N})$, $\boldsymbol{B} = diag(\boldsymbol{\beta}'_1, \ldots, \boldsymbol{\beta}'_N)$, $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_N)'$, $\boldsymbol{x}_t = (\boldsymbol{x}'_{1t}, \ldots, \boldsymbol{x}'_{Nt})'$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$.

To appreciate the generality of the STARF model, it is important to note three points. First, it can accommodate both weak and strong CSD through the spatial lagged terms and common factors. Second, it can accommodate time dependence through the lagged and diffusion terms. Finally, by virtue of its heterogeneous parameters, it allows the data to speak about the relative sensitivity of each spatial unit to external conditions. Consequently, the STARF model encompasses a wide range of existing models, including those of Pesaran (2006), Shi and Lee (2017), Chen et al. (2021).

Note also that the STARF model is related to the burgeoning literature on the network VAR model and its extensions (e.g. Härdle et al., 2016; Zhu et al., 2017, 2019; Greenwood-Nimmo et al., 2019, 2021;

Ando et al., 2021b,c,d), which has contributed greatly to our understanding of macroeconomic and financial networks. Both macroeconomic and financial systems are characterized by high-dimensionality coupled with local and global CSD, which results in networks of unknown and time-varying form. Moreover, in turbulent times, accurate estimation and prediction of systemic and idiosyncratic risks requires econometric techniques that are able to account for the topology of economic and financial networks. In this respect, the STARF model enjoys an important advantage relative to the network VAR model, because it explicitly incorporates the contemporaneous network regressor, y_{it}^*.

Remark 1. *It is straightforward to generalize the STARF model* (1) *by including spatial Durbin terms and additional time dynamics, as follows:*

$$y_{it} = \sum_{s=1}^{p_1} \rho_{si} y_{i,t-s} + \sum_{s=0}^{p_2} \phi_{si} y_{i,t-s}^* + \sum_{s=0}^{q_1} \boldsymbol{\beta}_{si}' \boldsymbol{x}_{i,t-s} + \sum_{s=0}^{q_2} \boldsymbol{\pi}_{si}' \boldsymbol{x}_{i,t-s}^* + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \varepsilon_{it} \tag{3}$$

where $\boldsymbol{x}_{it}^* = \left(x_{it,1}^*, \ldots, x_{it,k}^*\right)' \equiv \left(\sum_{j=1}^N w_{ij} x_{jt,1}, \ldots, \sum_{j=1}^N w_{ij} x_{jt,k}\right)' = \left(\boldsymbol{w}_i' \otimes \boldsymbol{I}_k\right) \boldsymbol{x}_t$. *Although the notation is more complex in this general case, the estimation techniques and asymptotic theory that we develop below can be applied with only straightforward modifications.*

2.2 Estimation

In order to estimate the parameters of the STARF model, it is necessary to handle the endogeneity of y_{it}^*, as well as the time-lagged terms $y_{i,t-1}$ and $y_{i,t-1}^*$ and $\boldsymbol{x}_{it}$ due to the presence of potentially serially correlated factors, $\boldsymbol{f}_t$.[1] We proceed in the same vein as Ando et al. (2021d) and estimate the unknown parameters of (2), which we denote compactly as $\boldsymbol{\theta} \equiv \{\boldsymbol{P}, \boldsymbol{\Phi}_0, \boldsymbol{\Phi}_1, \boldsymbol{B}, \boldsymbol{F}, \boldsymbol{\Lambda}\}$, by minimizing:

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\theta}) = \frac{1}{NT} \sum_{t=1}^{T} \left\| \boldsymbol{y}_t - (\boldsymbol{I} - \boldsymbol{\Phi}_0 \boldsymbol{W})^{-1} \left[\boldsymbol{P} \boldsymbol{y}_{t-1} + \boldsymbol{\Phi}_1 \boldsymbol{W} \boldsymbol{y}_{t-1} + \boldsymbol{B} \boldsymbol{x}_t + \boldsymbol{\Lambda} \boldsymbol{f}_t \right] \right\|^2, \tag{4}$$

subject to the identification conditions that $\boldsymbol{F}'\boldsymbol{F}/T = \boldsymbol{I}_T$ and that $\boldsymbol{\Lambda}'\boldsymbol{\Lambda}/N$ is a diagonal matrix. Instead of minimizing the quantile loss function as in Ando et al. (2021d), we minimize the squared loss function. As the objective function is nonlinear in the unknown parameters, we must update the parameters sequentially.

[1] In the special case where factors are serially uncorrelated, then it is easily seen that $E(y_{i,t-1}, \boldsymbol{f}_t) = 0$ and $E(y_{i,t-1}^*, \boldsymbol{f}_t) = 0$. However, if the factors are serially correlated, then $E(y_{i,t-1}, \boldsymbol{f}_t) \neq 0$ and $E(y_{i,t-1}^*, \boldsymbol{f}_t) \neq 0$ because $E(\boldsymbol{f}_t' \boldsymbol{f}_{t-j}) \neq 0$ for $j \geq 0$.

We begin by re-writing (1) in the following matrix form:

$$\boldsymbol{y}_i = \rho_i \boldsymbol{y}_{i,-1} + \phi_{0i}\boldsymbol{y}_i^* + \phi_{1i}\boldsymbol{y}_{i,-1}^* + \boldsymbol{x}_i\boldsymbol{\beta}_i + \boldsymbol{F}\boldsymbol{\lambda}_i + \boldsymbol{\varepsilon}_i = \boldsymbol{\chi}_i\boldsymbol{\theta}_i + \boldsymbol{u}_i, \tag{5}$$

where $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{iT})'$, $\boldsymbol{y}_{i,-1} = (y_{i0}, \ldots, y_{i,T-1})'$, $\boldsymbol{y}_i^* = (y_{i1}^*, \ldots, y_{iT}^*)'$, $\boldsymbol{y}_{i,-1}^* = (y_{i0}^*, \ldots, y_{i,T-1}^*)'$, $\boldsymbol{x}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{iT})'$, $\boldsymbol{F} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_T)'$, $\boldsymbol{\varepsilon}_i = (\boldsymbol{\varepsilon}_{i1}, \ldots, \boldsymbol{\varepsilon}_{iT})'$, $\boldsymbol{\chi}_i = (\boldsymbol{y}_{i,-1}, \boldsymbol{y}_i^*, \boldsymbol{y}_{i,-1}^*, \boldsymbol{x}_i)$, $\boldsymbol{\theta}_i = (\rho_i, \phi_{0i}, \phi_{1i}, \boldsymbol{\beta}_i')'$ and $\boldsymbol{u}_i = \boldsymbol{F}\boldsymbol{\lambda}_i + \boldsymbol{\varepsilon}_i$.

For now, we assume that the number of unobserved factors, r, is known. In this case, the estimation algorithm for the parameters of (1) proceeds in six steps, as follows:

Step 1 (Initial estimates of $\boldsymbol{\theta}_i$). Abstracting from the unobserved factor structure and the endogeneity issues associated with the model (1), the initial estimator for $\boldsymbol{\theta}_i$, denoted $\hat{\boldsymbol{\theta}}_i^{(0)}$, is obtained by running the following ordinary least squares (OLS) regression for each i:

$$\boldsymbol{y}_i = \boldsymbol{\chi}_i\hat{\boldsymbol{\theta}}_i^{(0)} + \hat{\boldsymbol{u}}_i^{(0)}. \tag{6}$$

All of the OLS residuals are collected into a $T \times N$ matrix, $\hat{\boldsymbol{U}}^{(0)} = \left(\hat{\boldsymbol{u}}_1^{(0)}, \ldots, \hat{\boldsymbol{u}}_N^{(0)}\right)$.

Step 2 (Initial estimator of the factors and the factor loadings). Following Bai (2003), principal components (PC) estimation is applied to:

$$\hat{\boldsymbol{U}}^{(0)} = \boldsymbol{F}\boldsymbol{\Lambda}' + \boldsymbol{\varepsilon}, \tag{7}$$

in order to obtain initial estimators of the factors and factor loadings, denoted $\hat{\boldsymbol{F}}^{(0)}$ and $\hat{\boldsymbol{\Lambda}}^{(0)}$, respectively, where $\hat{\boldsymbol{F}}^{(0)}$ is $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of the $T \times T$ matrix, $\hat{\boldsymbol{U}}^{(0)}\hat{\boldsymbol{U}}^{(0)'}$ and $\hat{\boldsymbol{\Lambda}}^{(0)} = \hat{\boldsymbol{F}}^{(0)'}\hat{\boldsymbol{U}}^{(0)}/T$.

Step 3 (Updating the heterogeneous spatial coefficients). Given the estimators obtained in steps 1 and 2, the spatial coefficients, $\hat{\boldsymbol{\Phi}}_0^{(b)} = diag(\hat{\phi}_{01}^{(b)}, \ldots, \hat{\phi}_{0N}^{(b)})$, are updated by numerical minimization of the following objective function:[2]

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\theta}) = \tfrac{1}{NT}\sum_{t=1}^{T}\left\|\boldsymbol{y}_t - (\boldsymbol{I}_N - \boldsymbol{\Phi}_0\boldsymbol{W})^{-1}\left[\hat{\boldsymbol{P}}^{(b-1)}\boldsymbol{y}_{t-1} + \hat{\boldsymbol{\Phi}}_1^{(b-1)}\boldsymbol{W}\boldsymbol{y}_{t-1} + \hat{\boldsymbol{B}}^{(b-1)}\boldsymbol{x}_t + \hat{\boldsymbol{\Lambda}}^{(b-1)}\hat{\boldsymbol{f}}_t^{(b-1)}\right]\right\|^2, \tag{8}$$

where we define $\hat{\boldsymbol{P}}^{(b-1)} = diag(\hat{\rho}_1^{(b-1)}, \ldots, \hat{\rho}_N^{(b-1)})$, $\hat{\boldsymbol{\Phi}}_1^{(b-1)} = diag(\hat{\phi}_{11}^{(b-1)}, \ldots, \hat{\phi}_{1N}^{(b-1)})$ and $\hat{\boldsymbol{B}}^{(b-1)} =$

[2] We also provide an alternative algorithm in which the spatial coefficients are estimated by QML; see Section S1 in the Online Supplement for further details, including Monte Carlo simulation results.

$diag(\hat{\boldsymbol{\beta}}_1^{\prime(b-1)}, \ldots, \hat{\boldsymbol{\beta}}_N^{\prime(b-1)})$.

Step 4 (Updating $\hat{\rho}_i^{(b)}$, $\hat{\phi}_{1i}^{(b)}$ and $\hat{\boldsymbol{\beta}}_i^{(b)}$). Given $\hat{\boldsymbol{\Phi}}_0^{(b)}$, the following de-factored regressions are estimated for each i by OLS:

$$\boldsymbol{M}_{\hat{\boldsymbol{F}}}^{(b-1)}(\boldsymbol{y}_i - \hat{\phi}_{0i}^{(b)}\boldsymbol{y}_i^*) = \boldsymbol{M}_{\hat{\boldsymbol{F}}}^{(b-1)}(\rho_i \boldsymbol{y}_{i,-1} + \phi_{1i}\boldsymbol{y}_{i,-1}^* + \boldsymbol{\beta}_i'\boldsymbol{x}_i) + \boldsymbol{M}_{\hat{\boldsymbol{F}}}^{(b-1)}\boldsymbol{u}_i, \tag{9}$$

where $\boldsymbol{M}_{\hat{\boldsymbol{F}}}^{(b-1)} = \boldsymbol{I}_T - \hat{\boldsymbol{F}}^{(b-1)}\left(\hat{\boldsymbol{F}}^{\prime(b-1)}\hat{\boldsymbol{F}}^{(b-1)}\right)^{-1}\hat{\boldsymbol{F}}^{\prime(b-1)}$. The residuals are updated as follows:

$$\hat{\boldsymbol{u}}_i^{(b)} = \boldsymbol{y}_i - \hat{\rho}_i^{(b)}\boldsymbol{y}_{i,-1} - \hat{\phi}_{0i}^{(b)}\boldsymbol{y}_i^* - \hat{\phi}_{1i}^{(b)}\boldsymbol{y}_{i,-1}^* - \boldsymbol{\beta}_i^{\prime(b)}\boldsymbol{x}_i, \tag{10}$$

and collected into $\hat{\boldsymbol{U}}^{(b)} = \left[\hat{\boldsymbol{u}}_1^{(b)}, \ldots, \hat{\boldsymbol{u}}_N^{(b)}\right]$.

Step 5 (Updating the factors and factor loadings). Updated estimates of the factors and factor loadings, denoted $\hat{\boldsymbol{F}}^{(b)}$ and $\hat{\boldsymbol{\Lambda}}^{(b)}$, are obtained by applying PC estimation with appropriate normalization to the following:

$$\hat{\boldsymbol{U}}^{(b)} = \boldsymbol{F}\boldsymbol{\Lambda}' + \boldsymbol{\varepsilon}. \tag{11}$$

Step 6. Repeat Steps 3 to 5 B times until convergence. The final converged estimators are denoted $\hat{\boldsymbol{F}}$, $\hat{\boldsymbol{\Lambda}}$, and $\hat{\boldsymbol{\theta}}_i$ for $i = 1, \ldots, N$.

2.3 Determination of the Number of Factors

In practice, the true number of factors is unknown and must be estimated consistently. An important issue is what happens to existing model selection criteria when spatial dependence is controlled for explicitly. We follow the factor literature (e.g. Bai and Ng, 2002; Ando and Bai, 2020; Ando et al., 2021d) and propose to select the number of unobserved common factors by optimization of the following information criterion:

$$\hat{r} = \underset{0 \le m \le r_{max}}{\arg\min}\ IC(m) = \underset{0 \le m \le r_{\max}}{\arg\min}\ ln\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\sigma}_i^2(m)\right) + mq(N,T), \tag{12}$$

where $\hat{\sigma}_i^2(m) = T^{-1}\sum_{t=1}^{T}(\hat{e}_{it}(m))^2$ for $i = 1, \ldots, N$, $\hat{e}_{it}(m)$ denotes the i-th element of $\boldsymbol{y}_t - (\boldsymbol{I}_N - \hat{\boldsymbol{\Phi}}_0\boldsymbol{W})^{-1}\left[\hat{\boldsymbol{P}}\boldsymbol{y}_{t-1} + \hat{\boldsymbol{\Phi}}_1\boldsymbol{W}\boldsymbol{y}_{t-1} + \hat{\boldsymbol{B}}\boldsymbol{x}_t + \hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{f}}_t(m)\right]$, $\hat{\boldsymbol{f}}_t(m)$ is an $m \times 1$ vector of factor estimates and $q(N,T)$ is a penalty function satisfying $q(N,T) \to 0$ and $\delta_{NT}^2 q(N,T) \to \infty$ with $\delta_{NT} = \min(\sqrt{N}, \sqrt{T})$.

3 Asymptotic Theory

3.1 Assumptions

We use the superscript '$*$' to denote the true value of parameters and treat the common factors, $\boldsymbol{f}_t$, and the factor loadings, $\boldsymbol{\lambda}_i$, as parameters.

Assumption A: Common factors

Let $\boldsymbol{\mathcal{F}}$ be a compact subset of $\mathbb{R}^r$. The common factors $\boldsymbol{f}_t^* \in \boldsymbol{\mathcal{F}}$ satisfy $T^{-1}\sum_{t=1}^{T} \boldsymbol{f}_t^* \boldsymbol{f}_t^{*\prime} = \boldsymbol{I}_r$.

Assumption B: Factor loadings and coefficients

(B1) Let $\boldsymbol{\mathcal{P}}$, $\boldsymbol{\mathcal{B}}$ and $\boldsymbol{\mathcal{L}}$ be compact subsets of $\mathbb{R}$, $\mathbb{R}^{p+1}$ and $\mathbb{R}^r$, respectively. The spatial and diffusion parameters, ϕ_{0i}^* and ϕ_{1i}^*, the time lag parameter, ρ_i^*, the slope parameter, $\boldsymbol{\beta}_i^*$, and the factor-loading, $\boldsymbol{\lambda}_i^*$, satisfy that $\phi_{0i}^* \in \boldsymbol{\mathcal{P}}$, $\phi_{1i}^* \in \boldsymbol{\mathcal{P}}$, $\phi_i^* \in \boldsymbol{\mathcal{P}}$, $\boldsymbol{\beta}_i^* \in \boldsymbol{\mathcal{B}}$ and $\boldsymbol{\lambda}_i^* \in \boldsymbol{\mathcal{L}}$ for each $i = 1, \ldots, N$.

(B2) The loading matrix, $\boldsymbol{\Lambda}^* = (\boldsymbol{\lambda}_1^*, \ldots, \boldsymbol{\lambda}_N^*)'$, satisfies $N^{-1}\sum_{i=1}^{N} \boldsymbol{\lambda}_i^* \boldsymbol{\lambda}_i^{*\prime} \xrightarrow{p} \boldsymbol{\Sigma}_\Lambda$, where $\boldsymbol{\Sigma}_\Lambda$ is an $r \times r$ positive definite diagonal matrix with distinct diagonal elements arranged in descending order. Further, the eigenvalues of $\boldsymbol{\Sigma}_\Lambda$ are distinct.

Assumption C: Idiosyncratic error terms

The random error, ε_{it}, is independent and identically distributed (i.i.d) over t and independent over i, with $E[\varepsilon_{it}] = 0$ and $E[\varepsilon_{it}^8]$ being bounded.

Assumption D: Weights matrix

(D1): $\boldsymbol{W}$ is an exogenous spatial weights matrix, the diagonal elements of which are zeros. In addition, $\boldsymbol{W}$ is bounded by some constant C for all N under $\|\cdot\|_1$ and $\|\cdot\|_\infty$.

(D2): Define $P(\boldsymbol{\Phi}_0) = (\boldsymbol{I} - \boldsymbol{\Phi}_0 \boldsymbol{W})^{-1}$ for $\boldsymbol{\Phi}_0 = \mathrm{diag}(\phi_{01}, \phi_{02}, \ldots, \phi_{0N})$ with ϕ_{0i} being an interior point of $\boldsymbol{\mathcal{P}}$ for each i. The matrix $P(\boldsymbol{\Phi}_0)$ is invertible over $\boldsymbol{\mathcal{P}}^N$ and satisfies $\limsup_{N\to\infty} \|P(\boldsymbol{\Phi}_0)\|_1 \vee \|P(\boldsymbol{\Phi}_0)\|_\infty \leq C$ for each $\mathrm{diag}(\boldsymbol{\Phi}_0) \in \boldsymbol{\mathcal{P}}^N$, where C is some positive constant.

Assumption E: Explanatory variables and design matrix

(E1): For a positive constant C_x, the explanatory variables satisfy $\sup_{it} \|\boldsymbol{x}_{it}\| \leq C_x$ almost surely.

(E2): Let $\mathcal{X}(\boldsymbol{P}, \boldsymbol{\Phi}_1, \boldsymbol{B})$ be an $N \times T$ matrix with its (i,t)th entry equal to $\rho_i y_{i,t-1} + \phi_{1i} y^*_{i,t-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta}_i$. Let d^*_{it} be the (i,t)th element of $\boldsymbol{W}(\boldsymbol{I}_N - \boldsymbol{\Phi}^*_0\boldsymbol{W})^{-1}(\mathcal{X}(\boldsymbol{P}^*, \boldsymbol{\Phi}^*_1, \boldsymbol{B}^*) + \boldsymbol{\Lambda}^*\boldsymbol{F}^{*'})$. Define $\boldsymbol{v}_{it} = (d^*_{it}, y_{i,t-1}, y^*_{i,t-1}, \boldsymbol{x}'_{it})'$, $\boldsymbol{V}_i = (\boldsymbol{v}_{i1}, \ldots, \boldsymbol{v}_{iT})'$, $\boldsymbol{A}_i = \frac{1}{T}\boldsymbol{V}'_i\boldsymbol{M}_{\boldsymbol{F}}\boldsymbol{V}_i$, $\boldsymbol{B}_i = (\boldsymbol{\lambda}^*_i\boldsymbol{\lambda}^{*'}_i) \otimes \boldsymbol{I}_T$, $\boldsymbol{C}_i = \frac{1}{\sqrt{T}}[\boldsymbol{\lambda}^*_i \otimes (\boldsymbol{M}_{\boldsymbol{F}}\boldsymbol{V}_i)]'$, $\boldsymbol{\eta} = \frac{1}{\sqrt{T}}\text{vec}(\boldsymbol{M}_{\boldsymbol{F}}\boldsymbol{F}^*)$, and $\boldsymbol{M}_{\boldsymbol{F}} = \boldsymbol{I} - \boldsymbol{F}(\boldsymbol{F}'\boldsymbol{F})^{-1}\boldsymbol{F}'$. Let $\boldsymbol{\mathcal{F}}$ be the collection of $\boldsymbol{F}$ such that $\boldsymbol{\mathcal{F}} = \{\boldsymbol{F} : \boldsymbol{F}'\boldsymbol{F}/T = \boldsymbol{I}_r\}$. We assume that:

$$\inf_{\boldsymbol{F}\in\boldsymbol{\mathcal{F}}} \lambda_{\min}\left[\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{E}_i(\boldsymbol{F})\right] > 0 \text{ with probability approaching one,}$$

where $\lambda_{\min}(\mathbb{M})$ denotes the smallest eigenvalue of the matrix $\mathbb{M}$, and $\boldsymbol{E}_i(\boldsymbol{F}) = \boldsymbol{B}_i - \boldsymbol{C}'_i\boldsymbol{A}^{-1}_i\boldsymbol{C}_i$.

(E3): Let $\mathcal{V}(\Psi)$ be an $N \times T$ matrix, the (i,t)th element of which is equal to $\phi^*_{0i}d^*_{it} + \rho^*_i y_{i,t-1} + \phi^*_{1i}y^*_{i,t-1} + \boldsymbol{x}'_{it}\boldsymbol{\beta}^*_i$, where $\boldsymbol{\Psi} = (\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_N)'$ and $\boldsymbol{\psi}_i = (\rho_i, \phi_{0i}, \phi_{1i}, \boldsymbol{\beta}'_i)'$. For any non-zero $\boldsymbol{\Psi}$, there exists a positive constant $\check{c} > 0$ such that:

$$\frac{1}{NT}\|\boldsymbol{M}_{\boldsymbol{\Lambda}^*}\mathcal{V}(\Psi)\boldsymbol{M}_{\boldsymbol{F}^*}\|^2 \geq \check{c}\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\psi}_i\|^2 \text{ with probability approaching one,}$$

where $\boldsymbol{M}_{\boldsymbol{\Lambda}^*} = \boldsymbol{I}_N - \boldsymbol{\Lambda}^*(\boldsymbol{\Lambda}^{*'}\boldsymbol{\Lambda}^*)^{-1}\boldsymbol{\Lambda}^{*'}$.

(E4): For each i, there exists a constant $c > 0$ such that:

$$\liminf_{T\to\infty} \lambda_{\min}\Big(\frac{1}{T}\boldsymbol{V}'_i\boldsymbol{M}_{\boldsymbol{F}^*}\boldsymbol{V}_i\Big) \geq c \text{ with probability approaching one.}$$

Assumption F: Stationary condition

The following matrix:

$$\sum_{h=0}^{\infty}\text{abs}\left(\left[(\boldsymbol{I} - \boldsymbol{\Phi}_0\boldsymbol{W})^{-1}(\boldsymbol{P} + \boldsymbol{\Phi}_1\boldsymbol{W})\right]^h(\boldsymbol{I} - \boldsymbol{\Phi}_0\boldsymbol{W})\right),$$

is uniformly bounded in row and column sums in absolute value. Here, $(\text{abs}[A])_{ij} = |a_{ij}|$ for a matrix $A = (a_{ij})$.

CZ(23/02//2023) At the beginning (before conducting the MC simulations), the stationary conditions are given by

One of the following stationary conditions is satisfied for each $i = 1, ..., N$.

$$\begin{aligned}
&\text{Case 1}: \quad \rho_i + (\phi_{0i} + \phi_{1i})\lambda_{\max}(\boldsymbol{W}) < 1 \quad \text{and} \quad \phi_{0i} + \phi_{1i} \geq 0, \\
&\text{Case 2}: \quad \rho_i + (\phi_{0i} + \phi_{1i})\lambda_{\min}(\boldsymbol{W}) < 1 \quad \text{and} \quad \phi_{0i} + \phi_{1i} < 0, \\
&\text{Case 3}: \quad \rho_i - (\phi_{0i} - \phi_{1i})\lambda_{\max}(\boldsymbol{W}) > -1 \quad \text{and} \quad \phi_{0i} - \phi_{1i} \geq 0, \\
&\text{Case 4}: \quad \rho_i - (\phi_{0i} - \phi_{1i})\lambda_{\min}(\boldsymbol{W}) > -1 \quad \text{and} \quad \phi_{0i} - \phi_{1i} < 0,
\end{aligned}$$

where $\lambda_{\max}(\boldsymbol{W})$ and $\lambda_{\min}(\boldsymbol{W})$ denote the largest and the smallest eigenvalues of matrix W.

The stationary conditions are then verified accordingly in both simulation and empirical analysis.

The stationary condition is now however changed to its current form later, and I will think about its verification.

Assumptions A and B on the factors and factor loadings are from Ando et al. (2021d), who study a spatial panel quantile model with a factor structure. Similar to Ando and Bai (2020) and Ando et al. (2021d), the factors and factor loadings are treated as parameters. Assumptions C and D on the idiosyncratic errors and the spatial weighting matrix are standard assumptions in the literature. The i.i.d. assumption imposed on the idiosyncratic errors is not restrictive, because our model explicitly accounts for dynamics and weak/strong CSD, as well as heterogeneity. Assumption E is necessary for deriving the consistency of the proposed estimators; see Ando and Bai (2020) and Ando et al. (2021d) for similar assumptions. Assumption F is a stationary condition (see Yu et al., 2008).

3.2 Asymptotic theory

Define $\boldsymbol{\varphi}_i^* = (\rho_i^*, \phi_{0i}^*, \phi_{1i}^*, \boldsymbol{\beta}_i^{*\prime}, \boldsymbol{\lambda}_i^{*\prime})'$ for $i = 1, \ldots, N$ and $\hat{\boldsymbol{\varphi}}_i$ as its estimator. We provide the following theorems, where all results are obtained under $(N, T) \to \infty$.

Theorem 1. *Suppose that Assumptions A – F hold. Under $\log(N)/T \to 0$, we have:*

$$\frac{1}{N}\sum_{i=1}^{N} \|\hat{\boldsymbol{\varphi}}_i - \boldsymbol{\varphi}_i^*\|^2 = O_p(\delta_{NT}^2),$$

where $\delta_{NT} = \max\{N^{-1/2}, T^{-1/2}\}$.

Theorem 2. *Under the regularity condition in Theorem 1, $N/T \to c$ with $0 < c < \infty$, for each $i =$*

$1, \ldots, N$, *we have:*

$$\sqrt{T}(\hat{\boldsymbol{\varphi}}_i - \boldsymbol{\varphi}_i^*) \xrightarrow{d} N(0, \Upsilon_i^{-1}\Gamma_i\Upsilon_i^{-1}), \tag{13}$$

where

$$\Upsilon_i = \lim_{T\to\infty} \frac{1}{T}\sum_{j=1}^{N}\sum_{t=1}^{T} s_{ji}^{*2}\boldsymbol{\gamma}_{it}^*\boldsymbol{\gamma}_{it}^{*\prime}, \quad \Gamma_i = \lim_{T\to\infty} \frac{1}{T}\sum_{j=1}^{N}\sum_{t=1}^{T} \sigma_j^{*2}s_{ji}^{*2}\boldsymbol{\gamma}_{it}^*\boldsymbol{\gamma}_{it}^{*\prime},$$

with $\boldsymbol{\gamma}_{jt}^* = (d_{jt}^*, y_{j,t-1}, y_{j,t-1}^*, \boldsymbol{x}_{jt}', \boldsymbol{f}_t^{*\prime})'$, d_{jt}^* *is defined in Assumption (E2),* s_{ji}^* *is the* (j,i)*th element of* $(\boldsymbol{I}_N - \boldsymbol{\Phi}_0^*\boldsymbol{W})^{-1}$*, and* σ_j^{*2} *is the variance of the* j*th element of* $\boldsymbol{y}_t - (\boldsymbol{I}_N - \boldsymbol{\Phi}_0^*W)^{-1}[\boldsymbol{P}^*\boldsymbol{y}_{t-1} + \boldsymbol{\Phi}_1^*\boldsymbol{W}\boldsymbol{y}_{t-1} + \boldsymbol{B}^*\boldsymbol{x}_t + \boldsymbol{\Lambda}^*\boldsymbol{f}_t^*]$.

To conduct statistical inference on $\boldsymbol{\varphi}_i$, we plug-in all estimated parameters $\hat{\boldsymbol{\theta}}$ to get Υ_i and Γ_i in Theorem 2. Further, we estimate σ_j^{*2} consistently by $\hat{\sigma}_j^2 = T^{-1}\sum_{t=1}^{T}\hat{e}_{jt}^2$ for $j = 1, \ldots, N$, where $\hat{e}_{jt}$ is the j-th element of $\boldsymbol{y}_t - (\boldsymbol{I}_N - \hat{\boldsymbol{\Phi}}_0W)^{-1}[\hat{\boldsymbol{P}}\boldsymbol{y}_{t-1} + \hat{\boldsymbol{\Phi}}_1\boldsymbol{W}\boldsymbol{y}_{t-1} + \hat{\boldsymbol{B}}\boldsymbol{x}_t + \hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{f}}_t]$.

Theorem 3. *Under the regularity conditions in Theorem 2, we have:*

$$\hat{\sigma}_i^2 \to \sigma_i^{*2} \textit{ for } i = 1, \ldots, N.$$

Assumption G: Identification for the over-fitted model

Let r be the true number of common factors. For each $k > r$, there exists a constant, $\breve{c}_k > 0$, such that, with probability approaching one:

$$\inf_{F^k, \frac{1}{T}F^{k\prime}F^k = I_k} \frac{1}{NT}\|M_{\Lambda_*}\mathcal{V}(\Psi)M_{F^k}\|^2 \geq \breve{c}_k\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\psi}_i\|^2,$$

where $\mathcal{V}(\Psi)$ is defined in Assumption (E3).

Assumption G is adopted from Ando et al. (2021a,d). Ando et al. (2021a) consider the determination of the number of common factors in large-scale panel choice models with interactive effects, while Ando et al. (2021d) use a similar condition to select the dimension of interactive effects in the panel spatial quantile model. As in Ando et al. (2021a,d), this assumption ensures the consistency of the model parameters when the specified dimension of interactive effects is larger than the true dimension.

Theorem 4. *Suppose that Assumptions A – G hold. Under* $N/T \to c$ *with* $0 < c < \infty$, $q(N,T) \to 0$ *and* $\min\{N,T\}q(N,T) \to \infty$*, the information criterion in (12) is consistent, such that* $P(\hat{r} = r) \xrightarrow{p} 1$.

Remark 2. *In practice, for the network-oriented approach that we develop in the next section, we need only $\sqrt{T}$-consistent estimators of the individual heterogeneous parameters. A pooled or mean group estimator will net out heterogeneous signs and, therefore, has the potential to produce a misleading global picture and to fail to examine the relative importance of individual nodes beyond what is assumed ex ante via $\boldsymbol{W}$. In particular, Pesaran and Smith (1995) and Chen et al. (2021) have shown the inconsistency of the pooled estimator in both heterogeneous dynamic and spatial panel data models. For the mean group estimator, although consistency and asymptotic normality could be established under the additional random parameter assumption, in many practical applications, there is no economic reason to expect the coefficients of the model to share a common sign (e.g Masten, 2018). We therefore relegate the development of the mean group estimator to Section S2 in the Online Supplement.*

4 Network Analysis

The STARF model extends the STARDL model advanced by Shin and Thornton (2021) in the sense that each spatial unit is subject to the influence of unobserved common factors. Shin and Thornton develop two measures to summarize the role of each node in the network: (i) individual dynamic multipliers; and (ii) system diffusion multipliers. Of particular interest are the latter, which form a sequence of connectedness matrices that can be interpreted as output network matrices resulting from the input network matrix, $\boldsymbol{W}$, and the heterogeneous spatial coefficients. In the following sections, we first develop two system diffusion multipliers that summarise the dynamic effects among individuals with respect to the influence of idiosyncratic errors and the exogenous regressors, respectively. Based on these calculated system effects, we then propose two further connectedness (network) measures defined at the individual and regional levels following Diebold and Yılmaz (2014) and Greenwood-Nimmo et al. (2021).

To facilitate the ensuing analysis, we first define $\boldsymbol{A}_0 = (\boldsymbol{I}_N - \boldsymbol{\Phi}_0\boldsymbol{W})$, $\boldsymbol{A}_1 = \boldsymbol{P}_1 + \boldsymbol{\Phi}_1\boldsymbol{W}$. We then re-write (2) as follows:

$$\boldsymbol{A}_0\boldsymbol{y}_t = \boldsymbol{A}_1\boldsymbol{y}_{t-1} + \boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{\Lambda}\boldsymbol{f}_t + \boldsymbol{\varepsilon}_t, \tag{14}$$

and derive the associated reduced-form VAR model as follows:

$$\boldsymbol{y}_t = \boldsymbol{\Psi}_1\boldsymbol{y}_{t-1} + \boldsymbol{A}_0^{-1}\boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{A}_0^{-1}\boldsymbol{\Lambda}\boldsymbol{f}_t + \boldsymbol{A}_0^{-1}\boldsymbol{\varepsilon}_t, \tag{15}$$

where $\boldsymbol{\Psi}_1 = \boldsymbol{A}_0^{-1}\boldsymbol{A}_1$.

4.1 Diffusion FEVD and Multipliers

Diffusion Forecast Error Variance Decomposition (FEVD). We first consider the diffusion effects of the idiosyncratic errors based on the FEVD. To this end, following Pesaran and Shin (1998), we write the infinite-order vector moving average (VMA) representation of (15) as:

$$\boldsymbol{y}_t = \sum_{j=0}^{\infty} \boldsymbol{\Theta}_j \boldsymbol{A}_0^{-1} \boldsymbol{B} \boldsymbol{x}_{t-j} + \sum_{j=0}^{\infty} \boldsymbol{\Theta}_j \boldsymbol{A}_0^{-1} \boldsymbol{\Lambda} \boldsymbol{f}_{t-j} + \sum_{j=0}^{\infty} \boldsymbol{\Theta}_j \boldsymbol{A}_0^{-1} \boldsymbol{\varepsilon}_{t-j}, \tag{16}$$

where $\boldsymbol{\Theta}_j = \boldsymbol{\Psi}_1^j$, with $\boldsymbol{\Theta}_0 = \boldsymbol{I}_N$.

Conditional on the information set at time $t-1$ (the regressors and the common factors), we can show that the covariance matrix of forecast errors associated with predicting $\boldsymbol{y}_{t+h}$ is $\sum_{\ell=0}^{h} \boldsymbol{\Theta}_\ell \boldsymbol{A}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \left(\boldsymbol{A}_0^{-1}\right)' \boldsymbol{\Theta}_\ell'$. Further, given the shocks to the ith equation, $(\varepsilon_{it}, \ldots, \varepsilon_{i,t+h})$ and assuming that $\boldsymbol{\varepsilon}_t \sim N(0, \boldsymbol{\Sigma}_\varepsilon)$, this covariance matrix becomes:

$$\sum_{\ell=0}^{h} \boldsymbol{\Theta}_\ell \boldsymbol{A}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \left(\boldsymbol{A}_0^{-1}\right)' \boldsymbol{\Theta}_\ell' - \sigma_{\varepsilon,ii}^{-1} \sum_{\ell=0}^{h} \boldsymbol{\Theta}_\ell \boldsymbol{A}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \mathbf{e}_i \mathbf{e}_i' \boldsymbol{\Sigma}_\varepsilon \left(\boldsymbol{A}_0^{-1}\right)' \boldsymbol{\Theta}_\ell',$$

which implies that, by conditioning further on $(\varepsilon_{it}, \ldots, \varepsilon_{i,t+h})$, the forecast error variance of $\boldsymbol{y}_{t+h}$ would decrease by $\boldsymbol{\Delta}_{ih} = \sigma_{\varepsilon,ii}^{-1} \sum_{\ell=0}^{h} \boldsymbol{\Theta}_\ell \boldsymbol{A}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \mathbf{e}_i \mathbf{e}_i' \boldsymbol{\Sigma}_\varepsilon \left(\boldsymbol{A}_0^{-1}\right)' \boldsymbol{\Theta}_\ell'$. Scaling the jth diagonal element of $\boldsymbol{\Delta}_{ih}$ $\left(\text{i.e. } \mathbf{e}_j' \boldsymbol{\Delta}_{ih} \mathbf{e}_j\right)$ by the h-step ahead forecast error variance of the jth variable in $\boldsymbol{y}_t$, we obtain the generalized FEVD as follows:

$$\text{GFEVD}\left(y_{jt}; \varepsilon_{it}, h\right) = \frac{\sigma_{\varepsilon,ii}^{-1} \sum_{\ell=0}^{h} \left(\mathbf{e}_j' \boldsymbol{\Theta}_\ell \boldsymbol{A}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \mathbf{e}_i\right)^2}{\sum_{\ell=0}^{h} \mathbf{e}_j' \boldsymbol{\Theta}_\ell \boldsymbol{A}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \left(\boldsymbol{A}_0^{-1}\right)' \boldsymbol{\Theta}_\ell' \mathbf{e}_j}. \tag{17}$$

for $\ell = 0, 1, \ldots, h$ and $i, j = 1, \ldots, N$, where $\mathbf{e}_j = (0, \ldots, 1, \ldots 0)'$ is the unit vector selecting the jth predictor and $\mathbf{e}_i$ is the unit vector selecting the ith innovation. This gives the proportion of the h-step ahead forecast error variance of the jth variable explained by the ith innovation (u_{it}) conditional on the information set at time $t-1$.

Diffusion Multipliers Following Shin and Thornton (2021), we consider the cumulative diffusion multipliers with respect to $\boldsymbol{x}_t$, defined as follows:

$$\boldsymbol{DM}_{\boldsymbol{x}}^{H} = \sum_{h=0}^{H} \frac{\partial \boldsymbol{y}_{t+h}}{\partial \boldsymbol{x}_t'} = \sum_{h=0}^{H} \boldsymbol{C}_h, \tag{18}$$

where we derive $\boldsymbol{C}_h = \boldsymbol{\Theta}_h \boldsymbol{A}_0^{-1} \boldsymbol{B}$ from (16). The cumulative diffusion multipliers of x_{jt}^k on $y_{i,t+h}$ are given by the $(i, (j-1)K + k)$th element of the $N \times NK$ matrix, $\boldsymbol{DM}_{\boldsymbol{x}}^{H}$.

4.2 Network Connectedness Measures

Individual Network Connectedness. Following Diebold and Yılmaz (2014), we provide summary statistics measuring the connectedness among individuals based on the diffusion multipliers given above. Suppose that we collect the $N \times N$ connectedness matrix obtained from either (17) or (18) as follows:

$$\underset{(N\times N)}{\mathbf{C}} = \begin{bmatrix} \phi_{1\leftarrow 1} & \cdots & \phi_{1\leftarrow N_1} & \phi_{1\leftarrow N_1+1} & \cdots & \phi_{1\leftarrow N_1+N_2} & \cdots & \phi_{1\leftarrow N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \phi_{N_1\leftarrow 1} & \cdots & \phi_{N_1\leftarrow N_1} & \phi_{N_1\leftarrow N_1+1} & \cdots & \phi_{N_1\leftarrow N_1+N_2} & \cdots & \phi_{N_1\leftarrow N} \\ \phi_{N_1+1\leftarrow 1} & \cdots & \phi_{N_1+1\leftarrow N_1} & \phi_{N_1+1\leftarrow N_1+1} & \cdots & \phi_{N_1+1\leftarrow N_1+N_2} & \cdots & \phi_{N_1+1\leftarrow N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \phi_{N_1+N_2\leftarrow 1} & \cdots & \phi_{N_1+N_2\leftarrow N_1} & \phi_{N_1+N_2\leftarrow N_1+1} & \cdots & \phi_{N_1+N_2\leftarrow N_1+N_2} & \cdots & \phi_{N_1+N_2\leftarrow N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \phi_{N\leftarrow 1} & \cdots & \phi_{N\leftarrow N_1} & \phi_{N\leftarrow N_1+1} & \cdots & \phi_{N\leftarrow N_1+N_2} & \cdots & \phi_{N\leftarrow N} \end{bmatrix}, \tag{19}$$

where each element, $\phi_{i\leftarrow j}$, measures the cumulative effect of individual j on individual i. It is then straightforward to define the following network measures:

(i) Total directional connectedness from others to i: $Spillin_{i\leftarrow\bullet} = \sum_{j=1,j\neq i}^{N} \phi_{i\leftarrow j}$,

(ii) Total directional connectedness to others from i: $Spillout_{\bullet\leftarrow i} = \sum_{j=1,j\neq i}^{N} \phi_{j\leftarrow i}$.

(iii) Total net directional connectedness from i to others: $Net = Spillout_{\bullet\leftarrow i} - Spillin_{i\leftarrow\bullet}$.

Note that the rows of the GFEVD matrix need not to sum unity (see Pesaran and Shin, 1998), so we row-normalize each row by its sum, as in Diebold and Yılmaz (2014).

Group Network Connectedness. When the number of individuals is large, it can be difficult to provide informative network analysis on the basis of disaggregated measures like those given above, because the number of bilateral linkages grows quadratically in N. To address this issue, we adopt the generalized connectedness measures (GCM) approach of Greenwood-Nimmo et al. (2021), which involves introducing

an intermediate level of aggregation in the analysis. Using this technique, we can develop network statistics focused on the $R(R-1)$ bilateral linkages among R groups rather than the $N(N-1)$ bilateral linkages among the N individuals in the system. With $R < N$ by construction, it will often be the case that $R(R-1) \ll N(N-1)$, yielding a significant reduction in the number of linkages under scrutiny.

To implement the GCM approach, we first need to select a grouping criterion that results in the formation of R non-overlapping groups (i.e. such that each individual belongs to one group only). We re-order the individuals so that they are gathered together into desired groups, and then calculate the $N \times N$ connectedness matrix as in (19). Suppose that there are N_j individuals in the jth group, $j = 1, 2 \ldots, R$. The (κ, ℓ)th block of (19), $(\kappa, \ell) = 1, \ldots, R$:

$$\underset{(N_\kappa \times N_\ell)}{\mathbf{B}_{\kappa \leftarrow \ell}} = \begin{bmatrix} \phi_{\tilde{N}_\kappa+1 \leftarrow \tilde{N}_\ell+1} & \cdots & \phi_{\tilde{N}_\kappa+1 \leftarrow \tilde{N}_\ell+N_\ell} \\ \vdots & \ddots & \vdots \\ \phi_{\tilde{N}_\kappa+N_\kappa \leftarrow \tilde{N}_\ell+1} & \cdots & \phi_{\tilde{N}_\kappa+N_\kappa \leftarrow \tilde{N}_\ell+N_\ell} \end{bmatrix}, \tag{20}$$

collects all of the individual effects from region ℓ to region κ, with $\widetilde{N}_\kappa = \sum_{j=1}^{\kappa-1} N_j$ for $\kappa = 2, \ldots, R$, and $\widetilde{N}_1 = 0$. We then sum the elements of $\mathbf{B}_{\kappa \leftarrow \ell}$ and normalize by the average number of countries in the pair as follows:

$$\psi_{\kappa \leftarrow \ell} = \frac{1}{0.5(N_\kappa + N_\ell)} \boldsymbol{\iota}'_{N_\kappa} \mathbf{B}_{\kappa \leftarrow \ell} \boldsymbol{\iota}_{N_\ell}, \tag{21}$$

where $\boldsymbol{\iota}_{N_\kappa}$ is an $N_\kappa \times 1$ column vector of ones. The $R \times R$ connectedness matrix defined at the group level is defined as follows:

$$\underset{(R \times R)}{\mathbf{C}_R} = \begin{bmatrix} \psi_{1 \leftarrow 1} & \psi_{1 \leftarrow 2} & \cdots & \psi_{1 \leftarrow R} \\ \psi_{2 \leftarrow 1} & \psi_{2 \leftarrow 2} & \cdots & \psi_{2 \leftarrow R} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{R \leftarrow 1} & \psi_{R \leftarrow 2} & \cdots & \psi_{R \leftarrow R} \end{bmatrix}. \tag{22}$$

It is now straightforward to derive the group spill-in, spill-out and net effects using (22) in the same way as for the disaggregate case based on (19):

$$RSI_i = \sum_{j=1, j \neq i}^{R} \psi_{i \leftarrow j}; \;\; RSO_i = \sum_{j=1, j \neq i}^{R} \psi_{j \leftarrow i}, \;\; RNE_i = RSO_i - RSI_i.$$

Finally, we define a pair of indices to address two issues of particular interest: (i) 'how dependent is the ith

group on external conditions in other groups'; and (ii) 'to what extent does the ith group influence/is the ith group influenced by the system as a whole?' We follow Shin and Thornton (2021) and construct the External Motivation (EM) and Systemic Influence (SI) indices as follows:

$$\text{EM}_i = \frac{RSI_i}{ATOT_{i\leftarrow\bullet}}; \;\; \text{SI}_i = \frac{RNE_i}{TNP_i}, \tag{23}$$

where $ATOT_{i\leftarrow\bullet} = \sum_{j=1}^{R} |\psi_{i\leftarrow j}|$ is the absolute row-sum for group i and $TNP_i = 0.5\sum_{i=1}^{N} |NE_i|$ is the total absolute net effect. EM_i measures the relative importance and direction of RSI in determining the conditions in the ith group, while SI_i captures the systemic influence of the ith group.[3]

5 Monte Carlo Simulations

In this Section, we investigate the accuracy of the factor selection criterion in (12), as well as the finite-sample performance of our proposed individual estimator.[4]

We begin by specifying the following data generating process (DGP):

$$y_{it} = \rho_i y_{i,t-1} + \phi_{0i}^* y_{it}^* + \phi_{1i}^* y_{i,t-1}^* + \beta_{1i} x_{it1} + \beta_{2i} x_{it2} + \gamma_{1i} f_{1t} + \gamma_{2i} f_{2t} + \sigma_i \varepsilon_{it}, \tag{24}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where both the number of regressors and the number of unobserved factors are set to 2. We generate each factor from the first-order autoregressive process:

$$f_{r,t} = \phi_{f_r} f_{r,t-1} + \xi_{f_{rt}}, \; t = -49, \ldots, T; \; r = 1, 2,$$

with $\phi_{f_r} = 0.5$ and $\xi_{f_{rt}} \sim IIDN(0, 1 - \phi_{f_r}^2)$ for $r = 1, 2$. We discard the first 50 observations as a burn-in sample. We then draw the factor loadings independently from the standard normal distribution.

Next, for each i, we generate the heterogeneous parameters as follows:

$$\rho_i = 0.3 + \xi_i, \;\; \phi_{0i}^* = 0.3 + \xi_i, \;\; \phi_{1i}^* = 0.3 + \xi_i, \;\; \beta_{i1} = 1 + \xi_i, \; \beta_{i2} = 1 + \xi_i,$$

[3] Both EM_i and SI_i are bounded in $[-1, 1]$. If $\text{EM}_i \to 1(-1)$, then conditions in group i are dominated by positive (negative) RSIs, as opposed to direct effects. If group i receives contradictory spill-ins and/or if the magnitude of RSI is small in comparison to the direct effects, then $\text{EM}_i \to 0$. If $0 \leq \text{SI}_i \leq 1$ ($-1 \leq \text{SI}_i \leq 0$), then group i is a net shock transmitter (receiver). If SI_i is close to zero, then the RSOs and RSIs of group i are approximately equal, and net out.

[4] Additional simulation results for the mean group estimator may be found in the Online Supplement.

with $\xi_i \sim IIDU(-0.2, 0.2)$.[5] If the stationary condition in Assumption F is violated, each of ρ_i, ϕ^*_{0i} and ϕ^*_{1i} is normalized by the absolute value of their sum.[6] We consider heteroskedastic errors by generating $\sigma_i \sim IIDU(0.5, 1.5)$ and $\varepsilon_{it} \sim IID(\chi^2 - 2)/2$.[7] Following Shi and Lee (2017), we generate the two exogenous regressors as follows:

$$x_{it,1} = 0.25(\gamma_{1i} f_{1t} + \gamma_{2i} f_{2t} + (\gamma_{1i} f_{1t} + \gamma_{2i} f_{2t})^2 + \gamma_{1i} + \gamma_{1i} + f_{1t} + f_{2t}) + v_{it,1}, \tag{25}$$

$$x_{it,2} = v_{it,2}, \tag{26}$$

where $v_{it,l} \sim N(0, 1)$ for $l = 1, 2$. Note that the first regressor is correlated with both factors and factor loadings, while the second regressor is independent of both.

The spatial weights matrix, $\boldsymbol{W}$, is constructed using the standard h-ahead and h-behind neighbor specification, whereby all elements of the weights matrix are set to zero apart from those off-diagonal elements that are within h rows on either side of the principal diagonal, which are set to $1/2h$. We then apply a row-sum normalisation. Here, we report simulation results using $h = 4$.[8] Each experiment is replicated 1,000 times for $(N, T) = \{50, 100, 200\}$.

5.1 Determining the Number of Factors

We first examine the finite-sample performance of the information criterion, $IC(m)$, which can take the following two forms:

$$IC_1(m) = \ln\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\sigma}^2_{i,m}\right) + m\frac{N+T}{NT}\ln(\min[N, T]), \tag{27}$$

$$IC_2(m) = \ln\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\sigma}^2_{i,m}\right) + m\frac{N+T}{NT}\ln\left(\frac{N+T}{NT}\right), \tag{28}$$

[5] Additional simulation results obtained with the means of $(\rho_i, \phi^*_{0i}, \phi^*_{1i})$ set to $(0.4, 0.4, 0)$, $(0.4, 0.4, -0.4)$, $(0.8, 0, 0)$, $(0, 0.8, 0)$ and $(0.4, 0.6, -0.3)$ are reported in the Online Supplement.

[6] CZ(23/2/2023): Notice that the previous stationary condition is verifiable. After generating the heterogeneous coefficients, I then checked if the stationary condition was satisfied. If not, for example, it is violated due to Case 1 (See Assumption F), I will then divide ρ_i,ϕ_{0i}, and ϕ_{0i}, by $\rho_i + (\phi_{0i} + \phi_{1i})\lambda_{\max}(\boldsymbol{W})$. Since the stationary condition has been changed, I will think about it more.

[7] Additional simulation results obtained with $\varepsilon_{it} \sim IIDN(0, 1)$ are reported in the Online supplement.

[8] Additional simulation results obtained using $h = 2$ and $h = 6$ are provided in the Online Supplement; they are qualitatively similar to the case with $h = 4$.

where $\hat{\sigma}^2_{i,m}$ is defined in (12). Setting $r_{\max} = 6$, we report the ratios of underestimation ($\hat{r} < r$), correct estimation ($\hat{r} = r$), and overestimation ($\hat{r} > r$) in Table 1. The performance of both IC_1 and IC_2 is satisfactory, with the probability of selecting the true number of factors rapidly converging to 1 with T or N for all values of h. In small samples, IC_2 slightly outperforms IC_1.

– Insert Table 1 Here –

5.2 Finite Sample Performance of the Proposed Estimators

Given the satisfactory performance of the information criterion in selecting the true number of factors, we henceforth assume that the number of factors is known. We now examine the finite sample performance of the proposed estimators as follows, using the notation θ_i to denote a generic parameter:

- **Bias** and **RMSE**. For each $i = 1, \ldots, N$, the bias and RMSE of $\hat{\theta}_i$, are given by $\sum_{m=1}^{M}(\hat{\theta}_i^m - \theta_i^*)/M$ and $\sqrt{\sum_{m=1}^{M}(\hat{\theta}_i^m - \theta_i^*)^2/M}$, where M is the number of replications, θ_i^* is the true value and $\hat{\theta}_i^m$ is the estimate for the m-th replication. We report the average bias and RMSE using $M = 1,000$ replications.

- **Size** and **Power**. The size of the t-test for $\hat{\theta}_i$ at the 5% significance level is evaluated as $\text{Size}_{\hat{\theta}_i} = \frac{1}{M}\sum_{m=1}^{M} \mathbb{1}\left(\frac{|\hat{\theta}_i^m - \theta_{i0}|}{\hat{\sigma}_{\hat{\theta}_i^m}} > 1.96\right)$, where $\mathbb{1}(.)$ is the indicator function, and $\hat{\sigma}_{\hat{\theta}_i^m}$ is the variance estimated by (13). To evaluate the power of the test, we set the alternative hypothesis as $\rho_i^a = \phi_{i0} + 0.2$. We again report averages using $M = 1,000$ replications.

Simulation results for the individual estimator are reported in Table 2. The biases of the individual estimators for all parameters are relatively small, even with sample sizes of just $(N, T) = (50, 50)$. Note that the RMSEs are decreasing with T only, with little change over N. This is in line with Theorem 2, which establishes the $\sqrt{T}$-consistency of the individual estimator. It is also interesting to observe that the RMSEs rise slightly with h. For small sample sizes (particularly small T), we observe some mis-sizing of the t-test, which is rectified as both N and T increase. As T increases, the power of the t-test quickly approaches 1. In sum, our simulations reveal that the finite sample performance of the individual estimator is satisfactory.[9]

– Insert Table 2 Here –

[9] In Tables S2-S3 of the Online Supplement, we report additional simulation results using the alternative QML estimation algorithm described in footnote S1. The simulation results are quite similar to those presented here, although the RMSEs and size performance are slightly better in small samples.

6 An Empirical Analysis of Business Cycle Synchronization

To investigate the synchronization of international business cycles, we specify and estimate the following heterogeneous STARF model:

$$g_{it} = \rho_i g_{i,t-1} + \phi_{0i} g^*_{it} + \phi_{1i} g^*_{i,t-1} + \beta_i \text{TFP}_{it} + \boldsymbol{\lambda}'_i \boldsymbol{f}_t + \varepsilon_{it}, \tag{29}$$

where g_{it} is the real output growth in country $i = 1, \ldots, N$ at time $t = 1, \ldots, T$, obtained by differencing the logarithm of real GDP, $g^*_{it} = \sum_{j=1}^{N} w_{ij} y_{jt}$ is the spatial-lagged variable and $g_{i,t-1}$ and $g^*_{i,t-1}$ are their respective time lags. The specification of (29) is more flexible than most existing studies such as those of Ho et al. (2013) and Ertur and Koch (2007), as it allows for heterogeneous spatial dependence, unobserved common factors and growth persistence. In addition, our specification captures the diffusion impact of total factor productivity (TFP) on the growth rate of real GDP. Mastromarco et al. (2016) identify TFP as the primary contributor to GDP growth. It is also believed to play a central role in driving rapid growth in less developed countries (e.g. Romer, 1993; Felipe, 1999). Data on both annual real GDP and TFP are sourced from version 10.0 of the Penn World Table (PWT), yielding a balanced panel covering 79 countries over the 50 years from 1970 to 2019, inclusive. See Table S55 for a complete list of the countries in our sample.

We obtain the spatial weights matrix using the normalized bilateral trade weight for the (i, j)th country pair, which we construct as follows:

$$w_{ij} = \frac{\text{EXP}_{ij} + \text{IMP}_{ij}}{\sum_{j=1}^{N} \text{EXP}_{ij} + \sum_{j=1}^{N} \text{IMP}_{ij}}, \tag{30}$$

where EXP_{ij} denotes the exports from country i to country j and IMP_{ij} denotes the imports of country i from country j. We denote the resulting weights matrix $\boldsymbol{W}_{tra}$.[10] To avoid the potential endogeneity of the spatial weights arising via the correlation of trade with GDP growth, we use the 5-year average of exports and imports for each country pair over the period 2014-2019 (see Greenwood-Nimmo et al., 2012, and the literature cited therein for a similar approach).[11]

[10] In the Online Supplement, we also report estimation results obtained using inverse distance-based weighting schemes. These results are available on request.

[11] As a robustness check, we also construct the matrix of trade weights using a 10-year average over the period 2009-2019. The results are similar and are available from the authors upon request.

6.1 Estimation Results

We begin by analyzing the individual estimation results for each country. Both of the information criteria defined in (28) indicate that the optimal number of common factors is 4. Table S55 indicates that the stationary condition in Assumption F is satisfied in the large majority of cases and that none of $\hat{\rho}_i$ or $\hat{\phi}_{0i}$ fall outside $[-1, 1]$. We exclude the unstable cases in the subsequent analysis.

In Figures 1a–1d, we provide a concise visualization of the heterogeneous parameter estimates from (29) in the form of a choropleth.[12] First, we analyse the dynamic coefficients plotted in Figure 1a. The estimated parameters exhibit notable heterogeneity. Most are positive (72, 91.1%) and statistically significant (57, 72.2%); in fact, none of the negative estimates are significant. This is consistent with the stylized fact that output growth is positively autocorrelated (e.g. Nelson and Plosser, 1982; Campbell and Mankiw, 1987). Furthermore, the figure reveals that estimated autoregressive coefficient is typically larger among advanced economies. Of the ten countries with the largest estimates of ρ_i, nine of them are advanced countries, including the U.S., Spain and Australia. Recall that the speed of adjustment to the steady state is captured by $1 - \rho_i$ (β-convergence). Consequently, our results imply that developed economies converge to equilibrium more slowly than their less developed counterparts, which reflects the finding of Lee et al. (1997) that the adjustment speed for OECD member countries is slower than for the intermediate group of countries.

Turning now to the contemporaneous spatial coefficient, ϕ_{i0}, plotted in Figure 1b, we again observe considerable heterogeneity. Of the 79 countries, the parameter estimate is positive for 63 (79.8%) and statistically significant for 44 (55.7%). This finding reflects the positive international spillover of output growth due to technology transfer and trade connections (Ho et al., 2013). While the spatial coefficients for most of the world's largest economies by GDP are positive and significant, it is interesting to see that two developing countries, China and Mexico, display negative but insignificant estimates of -0.048 and -0.089, respectively (see Table S55). Re-write: Our conjecture is as follows: China trades extensively with Western countries such that their growths are expected to be positively correlated. But, China has also a very close connection with many Asian countries geographically, culturally and economically. These countries, such as India, Indonesia, Malaysia, Philippines, Thailand and Vietnam, share similar industrial and export structures with China, featuring in cheap, hardworking labors and significant positive trade balance with Western countries. These economies are more likely to be competitors than collaborators for China, which

[12] Detailed estimation results could be found in Table S55 in the Online Supplement.

may lead to their negative growth relations.[13] Similarly, Mexico also competes with countries like Brazil and Colombia. As the largest immigration country for US, Mexicans constitute a non-negligible part of the U.S. labour market, which may reduce the human capital in Mexico, thus a detrimental factor to its own economic development.

Another interesting finding is that European countries tend to display high spatial dependence. Among the twenty countries with largest estimates of ϕ_{0i}, twelve are European. This likely reflects economic cooperation and convergence within the European Union and wider European Economic Area.

The estimates of the diffusion parameter, ϕ_{1i}, also exhibit notable heterogeneity, as seen in Figure 1c. Contrary to the contemporaneous spatial coefficients, ϕ_{0i}, the estimates of which are mostly positive, more than half of the estimated diffusion parameters are negative (46, 58.2%) and most are insignificant (53, 70.9%). This is consistent with the findings of Servén and Abate (2020).

Finally, our estimates of β_i, displayed in Figure 1d, are again quite heterogeneous. Inklaar and Timmer (2013) note that the TFP data reported in the PWT is defined as a Solow (1957) residual that controls for individual and time-specific labour shares only, under the assumption of cross-section independence. This offers an intuitive explanation for our finding that the estimates of β_i are all significantly positive and take values around 1, as we add time and space variation into the model without controlling for capital and labour inputs. Still, it is interesting to observe that the US and Japan are the two countries with largest estimated coefficient. This is suggestive of the importance of technical progress for the leading developed economies, considering that the US and Japan are the largest developed economies over our sample period. In a similar vein, Kim and Lau (1994) also document the importance of TFP growth on industrialised countries.

Mean group estimation results are presented in Table S56. In addition to the overall average, we also report mean group estimates based on four different grouping criteria: the *Continent* and *Subregion* criteria, where countries are grouped according to their geographical location, and the *Income Level* and *Development* criteria, where countries are grouped according to their economic status. A summary of the membership of each group and a detailed view of each country's group identity are provided in Tables S53 and S54. We use the same grouping criteria in our subsequent GCM analysis. Across all of these mean group specifications, only the coefficient on TFP is consistently statistically significant. This suggests that heterogeneity and volatility of the individual estimates for the other parameters leaves the mean group estimator largely uninformative (see Remark 2). For further discussion of the mean group estimates, see

[13] See Marukawa (2021) for more analysis on the competition relationship between Asian countries and China.

Section S4.3 in the Online Supplement.

6.2 Network Analysis

6.2.1 Individual Network Analysis

In this subsection, we analyze the individual network estimation results using the tools introduced in Section 4.2, for both the diffusion FEVDs defined in (17) and diffusion effects of TFP calculated according to (18).

FEVD Analysis

In Table S57, we report the network of diffusion FEVDs for each country over horizons $h = 0, 2, 5$.[14] The estimation results are relatively similar across horizons, which implies rapid convergence to the long-run values as h rises. This is a reflection of the stationarity of the system. The US, China and Germany are the three countries with the largest net spillover effects in terms of a unit (one standard error) country-specific shock. Unlike the US and Germany, the spill-in effect that China receives is very small and essentially negligible when compared to its large spill-out effect. This is not a surprising finding considering that China was the first major economy in the world to emerge from the 2008 financial crisis and maintain relatively steady growth over the past few decades regardless of the global economic climate.

Several developed countries, including three G7 members, namely Canada, France and Italy, exhibit negative spillover effects, driven primarily by their relatively large spill-in effects from the US and Germany. For example, around one third of the total shocks that France and Italy receive are from Germany, and more than eighty percent of the total shocks Canada receives are from the US. Small economies are typically net shock recipients, as one may expect. Luxembourg, the country with the highest GDP per capita in our sample, provides a good example, being the largest net receiver of shocks in our analysis. This reflects its openness, its dependence on its partners, its small size and its location in the heart of Western Europe.[15] Most developing economies are also net receivers. The central American countries in our sample provide a good illustration: Honduras, Costa Rica and Guatemala are all among the biggest net shock receivers and all have the common feature of significant exposure to US shocks. India and Brazil, as some of the largest emerging economies in the world, exhibit modest positive net spill-out effects.

[14] Results using the distance-based weighting matrix $\boldsymbol{W}_{dist}$ are reported in Table S59 of the Online Supplement.

[15] According to the World Bank, Luxembourg is very open to foreign trade, which represented 390.3% of its GDP in 2020, the highest level in the world.

TFP analysis

In Table S58, we report the diffusion effects for TFP calculated according to (18) over horizons $h = 0, 2, 5$.[16] Similar to the FEVD analysis, the estimated results converge rapidly with h. The five largest economies in the world by GDP—the US, China, Japan, Germany and UK—are the largest net transmitters of TFP spillovers. This finding is suggestive of a dominant role of TFP spillovers in driving global growth.

TFP growth in France and Italy also exerts large positive effects overseas, although the net effects for both countries are negative, due to the spill-in effects that they experience from both the US and Germany. The spill-out effect from Canada is weaker than those of other major economies. Furthermore, the spill-in effect from the US to Canada is so strong that Canada records one of the lowest net effects in our sample. This finding is not surprising given the well-documented dependence of the Canadian economy on the US.

Interestingly, we find that Mexican growth is negatively correlated with TFP growth in US. We conjecture that, while Mexico undoubtedly benefits from technology spillovers from the US, this positive effect may be offset by technological developments that reduce the incentives for US firms to operate offshore and innovations that create a 'brain-drain' from Mexico, reduce its human capital. Nevertheless, the spill-out effects from Mexico are large due to its regional impacts on central and south American economies. Other examples of regional leadership can also be seen in our results; for example India and Korea are major technology transmitters with respect to the fast-developing economies in South-Eastern Asia, such as the Philippines and Thailand.

6.2.2 Regional/Group Network Analysis

In this subsection, we develop an enriched understanding of regional/group connectedness using GCM analysis subject to the four grouping criteria mentioned in Section 6.1.

FEVD analysis

First, the results by continent are presented in Table 3. Most of the off-diagonal elements are very small, implying that shocks occurring in one region tend to primarily affect that region. This finding may partly reflect our use of a common factor structure that accounts for much of the comovement among regions. Nonetheless, some interesting results can be seen. For all continents, the largest spill-in effect is from Asia. Consequently, Asia has the largest spill-out effect of any continent. North America exerts the second largest

[16] Results using $\boldsymbol{W}_{dist}$ are provided in Tables S60 of the Online Supplement.

spill-out effect, which increases over the forecast horizon. Europe exhibits the third largest spill-out effect, but this is dominated by spill-ins from both north America and Asia, resulting in a net negative spillover effect. Africa, Oceania and South America are all net recipients and exhibit negligible spill-out effects. In Figures 2a-2b, we plot the coordinate pair $(\mathrm{EM}_i, \mathrm{SI}_i)$, which provides a vivid representation of the role of each continent in the global economy. The figure reveals a tendency for regions to cluster along a line from north-west to south-east, as positive (negative) spill-ins contribute negatively (positively) to a region's net effect. The figure highlights the unique position of Asia and North America as the most influential regions in the World, being the only continents with positive EM values. Europe is the most externally dependent continent, while South America appears to be relatively isolated, with small values of both the EM and SI indices. It is noteworthy that all of the EM indices take relatively small values, which implies that continent-specific shocks are not transmitted strongly between continents. It is also interesting to note that a comparison of Figures 2a-2b reveals little sensitivity across forecast horizons.

We should plot all panels on the same scale to make the figure easier to read.

CZ:In Table 4, we present results by sub-region. Among the Asian regions, it is Eastern Asia that has the largest net spill-out effect. While Northern America has the largest spillouts, it also receives relatively large spillin effect from both Eastern Asia and Central America. The spillout effect from Western European Region are also generally larger than the rest regions. Its net effects is however negative due to the spillins from Northern America and Eastern Asia. For all the other regions, the spillouts are relatively small with negative net effects.

For the EM and SI indices plotted in Figure 2c-2d, we find that the dependence of Western European Region on external shocks are almost twice stronger than that of other European Regions, and it is only second to Central America that is largely affected by Northern America. Northern America and Eastern Asia are the two influential regions in terms of transmitting their regional-specific shock to the global economy, and the former also has large external dependence on the latter. For most of the other subregions, they are having both small external dependence and system influence. Moreover, the role that each region plays in the global economy is generally stable across time.

Combining all of the analysis above, we find that the three geographical regions, North America (including both Northern and Central America), Eastern Asia and most of Europe, are having deeper mutual linkage than the rest regions, in terms of either receiving/ transmitting regional specific shocks. The rest regions are weakly connected both mutually and with the above three areas, except that the linkage between Pacific

Islands and Eastern Asia and Northern America are getting stronger with time.

Second, the results by economic grouping are presented in Tables 5.

In table 5, we report the results by grouping countries according to their income level. The high income and Upper-middle income countries are mutually dependent in the sense that they both receive/transmit relatively large group-specific shock effects to each other. Moreover, they are the only two groups with positive net shock effects due to their spillout effects and almost negligible spillins from the to Lower-middle countries. Low income countries are the only group receive positive spillins but with no spillouts. All the findings are robust to time implying the stationarity of the system. Further, from the EM and SI indices plotted in Figure 2e-2f, we find that except for low income countries, all the other three groups are having relatively large mutual external dependence. Upper-middle income countries are the most influential transmitters of shock effects whereas Lower-middle income countries are the biggest shock effects receivers. As time goes by, high income countries also become more and more influential shock effects transmitters.

TFP analysis

We also report the regional/group TFP network results, using all the given grouping criteria and by the weighting scheme, $\boldsymbol{W}_{tra}$.

We first analyze the results by geographical grouping in Tables 6-7.

From Table 6, we can see that the TFP growths in both North America and Asia have large spillout effects on all others continents, in particular on the European, while the reverse effects are relatively smaller. As a result, they are the only two continents with positive net effects. Although the TFP growth in Europe has the third largest spillout effects, it has the smallest net effects due to large spillins from North America and Asia. The effects of TFP growth in the rest three regions are similar in the sense that, while they could benefit from the TFP growths from the other three regions, the improvement of TFP in these regions has relatively small effects on other regions, making them effects receivers in the world. It is interesting to see that as time goes by, the TFP growth in other regions, in particular those in North America and Europe, will have negative effects on Africa. Considering that the major export for Africa is primary commodities and raw materials (e.g. petroleum and copper), TFP growth in other regions would reduce their demand for these commodities in the future, which in turn would be detrimental to Africa's economy. Meanwhile, from Figures 3a-3b, we find that Europe benefits the most from the TFP growth in other regions, while South America receives the least boost in economics from the TFP growth in other regions. This might be sensible considering that, on

the one hand, the economy in South America could benefit from adopting technology developed in other regions, its export (same as Africa, mainly primary commodities and raw materials, e.g. plant products and fuels) would also suffer from the development of new technology that reduces demand. These two effects cancel out with each other. Owing to large spillout effects, Asia and North America are the most influential continents in terms of transmitting TFP growth effects, and all the other regions are net receivers benefiting from the increase of TFP in other regions.

CZ: When looking at the effects of TFP growth on a smaller scales at sub-regions (Table 7), we find that the large spill out TFP effects on GDP growth of Asia and North America are mainly attributed to Northern America and Eastern Asia countries. Specifically, Northern America countries exert large spillover effects of TFP growth to countries in Western and Central America areas, while those of Eastern Asia countries have large positive spillout effects on countries in South-Eastern Asia, Pacific islands, and Western Europe. It is also interesting to see both the spillin TFP effects in this two subregions are relatively small among all the regions, implies weak dependence of their GDP growth on TFP growth in other regions. This finding may contribute to the debate that the growth miracle in Eastern Asia is mainly "input-driven" instead of acquisition and mastery of foreign technology, see e.g. Young (1995),Collins et al. (1996). Western Europe also exerts large positive spill over effects of TFP growth on outputs growth in other European areas. It is however a net effects receiver due to the large spillins from both Northern America and Eastern Asia, and similarly for Other Europe countries. Central America and South-Eastern Asia are receiving very large spills effects of TFP growth: while Central America receives the largest spillins from Northern America, South-Eastern Asia receives the largest spillins from Eastern Asia. This might be explained by their geographical closeness that could boost trade and other economic interactions. Africa and South America tend to be "closed" regions in the sense of both receiving and transmitting small effects from TFP growth on GDP growth. This is in line with existing findings that the average GDP growths in African/South American countries are driven primarily by factor accumulation with little or no role for TFP (see e.g. Bosworth and Collins (2003),Tahari et al. (2004).). The plots of SI and EM indices in Figures 3c-3d further confirm the above findings, that Eastern Asia and Northern America are the most influential TFP effects transmitters whereas the former has smaller external dependence. GDP growth in Western European countries benefit a lot form technology developed in Northern America and Eastern Asia, which is also case for Central America and Pacific Islands. Asian countries other than Eastern Asia are all relatively neutral to the TFP growth in other regions, and their roles in the global economy growth remians stable across time.

We now investigate how the effects of TFP growth transmit across countries with different economic development.

We first look the transmission of TFP effects among countries grouped by their income level in Table 8. As could be seen, High and Upper Middle income countries are having a relatively strong mutual dependence, that both of them receives/exerts large TFP effects on the outputs growth in the other region. These two regions also both have non-negligible influence on the GDP growth in Lower-middle countries, although the effects from Lower-middle countries to them are generally of much smaller magnitude. Meanwhile, it is interesting to see that Low income countries receive largest spillins from Lower-middle countries and least from High income group. This may reflect that economic contentedness or cooperation is more likely to be made between countries/regions that does not differ much (see Lucas (1990)). As h changes, although Upper-middle income countries still have large spillout effects, the increase is however smaller than the increase of boosting effects it receives from High income countries, making it now a net effect-receiver group. Another notable finding is that both the spill over effects from High and Upper Middle income countries on Lower-middle countries are getting smaller, which may be explained by that TFP increase in higher income countries will lead to reduction of demand for export from lower incomes countries. From the EM and SI indices in Figures 3e-3f, we also find countries with higher income tend to lead the world economy in the sense of having smaller external dependence (EM) but larger spillout effects (SI). However, for countries with lowest incomes which means they might still at the early stage of economic development, the transmission of capital and human capital from richer countries might be more important than technology transmission, see e.g. Cohen et al. (2004). In line with the changing patters of regional spill in/out effects, Lower-middle income countries are becoming more influential owing to its net positive effects to the whole world economy.

7 Concluding Remarks

In this paper, we introduce and study a heterogeneous spatio-temporal autoregressive process with a common factor error structure. This process is appealing by virtue of its generality, as its offers a simple and analytically tractable unified characterization of CSD that combines both spatial and factor dependence. Following Ando et al. (2021d), we develop a modified QML method to overcome the problem of endogeneity and achieve consistent estimation. Under mild regularity conditions, we show that the individual estimator is

$\sqrt{T}$ consistent and asymptotically normally distributed. Monte Carlo simulations indicate that the estimator achieves satisfactory finite-sample performance. Based on the resulting estimators, together with the input network matrix, we further developed system diffusion FEVDs/multipliers and both individual and regional network topology measuring tools, following Diebold and Yılmaz (2014) and Shin and Thornton (2021).

Applying these tools to analyse the business cycle synchronisation covering 79 countries in the world over the period 1970-2019, we found that (i) Both the individual dynamic and contemporary spatial coefficients are quite heterogeneous, but overall positive. These confirm our *a priori* expectation that the output growth is both serially and cross-sectionally positively correlated. (ii) From individual network analysis, we found US, China and Germany to be among the largest transmitters of both country-specific shock and TFP growth effects. While some other major economies (e.g. France, Italy and Canada) are also exerting both large country-specific shock and TFP growth effects, their net effects are however negative due to their large spill-ins. (iii) On the aggregated level, the geographical regions: Northern America and Eastern Asia, and the economic groups: High-income and G7 countries, are identified to be the most influential regions in terms of transmissions of both the regional shock effects and TFP effects.

We conclude by noting a few avenues for future research. A natural development is to extend the current analysis to the nonlinear/quantile case. Such extensions are interesting considering that nonlinear relationships are abound in real economics (Granger and Terasvirta (1993)) and researches/policy makers are often interested in distributional effects beyond mean analysis (see e.g. Chernozhukov et al. (2020)). Another interesting research direction might be the exploration of alternative estimation strategy. While we use the QML approach, we may also develop the alternative IV/GMM estimation as in Cui et al. (2020) though they consider a homogeneous model.

Appendix A Proof of the Main Theorems

Ando et al. (2021d) obtain the same results presented in our Theorems 1, 2 and 4 for the panel spatial quantile model with interactive effects and heterogeneous slope coefficients. Unlike the quantile loss function studied by Ando et al., in contrast, our objective function is smooth. We therefore adapt Ando et al.'s technical analysis to prove Theorems 1, 2 and 4 by replacing the quantile loss function with the squared loss function. To avoid unnecessary repetition of material from Ando et al., we limit our discussion to the key points of the technical proof.

A.1 Proof of Theorem 1

Define $U(\boldsymbol{\theta}) \equiv \boldsymbol{\mathcal{L}}(\boldsymbol{\theta}) - \boldsymbol{\mathcal{L}}(\boldsymbol{\theta}^*)$. We first borrow the argument used for the proof of Proposition 1 in Bai (2009). Given the definition of $\hat{\boldsymbol{\theta}}$, we have the following inequality (A): $U(\hat{\boldsymbol{\theta}}) \leq U(\boldsymbol{\theta}^*) = 0$. Putting $\boldsymbol{y}_t = (\boldsymbol{I}_N - \boldsymbol{\Phi}_0^* \boldsymbol{W})^{-1}(\boldsymbol{P}^* \boldsymbol{y}_{t-1} + \boldsymbol{\Phi}_1^* \boldsymbol{W} \boldsymbol{y}_{t-1} + \boldsymbol{B}^* \boldsymbol{x}_t + \boldsymbol{\Lambda}^* \boldsymbol{f}_t^* + \boldsymbol{e}_t)$ into $U(\boldsymbol{\theta})$, we obtain the following equality (B): $U(\boldsymbol{\theta}) = \frac{1}{NT} \sum_{t=1}^T \|(\boldsymbol{I}_N - \boldsymbol{\Phi}_0^* \boldsymbol{W})^{-1}(\boldsymbol{P}^* \boldsymbol{y}_{t-1} + \boldsymbol{\Phi}_1^* \boldsymbol{W} \boldsymbol{y}_{t-1} + \boldsymbol{B}^* \boldsymbol{x}_t + \boldsymbol{\Lambda}^* \boldsymbol{f}_t^*) - (\boldsymbol{I}_N - \boldsymbol{\Phi}_0 \boldsymbol{W})^{-1}(\boldsymbol{P} \boldsymbol{y}_{t-1} + \boldsymbol{\Phi}_1 \boldsymbol{W} \boldsymbol{y}_{t-1} + \boldsymbol{B} \boldsymbol{x}_t + \boldsymbol{\Lambda} \boldsymbol{f}_t)\|^2 + o_p(1)$. Here, we have used the independence property of the idiosyncratic error term and the assumptions imposed on the weights matrix. Combining conditions (A) and (B) together leads to: $\frac{1}{NT} \sum_{t=1}^T \|(\boldsymbol{I}_N - \boldsymbol{\Phi}_0^* \boldsymbol{W})^{-1}(\boldsymbol{P}^* \boldsymbol{y}_{t-1} + \boldsymbol{\Phi}_1^* \boldsymbol{W} \boldsymbol{y}_{t-1} + \boldsymbol{B}^* \boldsymbol{x}_t + \boldsymbol{\Lambda}^* \boldsymbol{f}_t^*) - (\boldsymbol{I}_N - \hat{\boldsymbol{\Phi}}_0 \boldsymbol{W})^{-1}(\hat{\boldsymbol{P}} \boldsymbol{y}_{t-1} + \hat{\boldsymbol{\Phi}}_1 \boldsymbol{W} \boldsymbol{y}_{t-1} + \hat{\boldsymbol{B}} \boldsymbol{x}_t + \hat{\boldsymbol{\Lambda}} \hat{\boldsymbol{f}}_t)\|^2 = o_p(1)$. In this step, we have applied the same argument used for the proof of Proposition 1 in Bai (2009). Next, we apply the technical proof of Theorem 1 in Ando et al. (2021d). We then immediately have $\frac{1}{N} \sum_{i=1}^N \|\hat{\boldsymbol{\varphi}}_i - \boldsymbol{\varphi}_i^*\|^2 = o_p(1)$ and $\frac{1}{T} \sum_{t=1}^T \|\hat{\boldsymbol{f}}_t - \boldsymbol{f}_t^*\|^2 = o_p(1)$. Finally, we apply the proof of Theorem 2 in Ando et al. (2021d). The claim immediately follows.

A.2 Proof of Theorem 2

Following the proof of Theorem 3 in Ando et al. (2021d), we replace the quantile loss function with the squared loss function. The claim immediately follows.

A.3 Proof of Theorem 3

Let $\hat{\eta}_{it}$ be the i-th element of $(\boldsymbol{I}_N - \hat{\boldsymbol{\Phi}}_0 \boldsymbol{W})^{-1}(\hat{\boldsymbol{P}} \boldsymbol{y}_{t-1} + \hat{\boldsymbol{\Phi}}_1 \boldsymbol{W} \boldsymbol{y}_{t-1} + \hat{\boldsymbol{B}} \boldsymbol{x}_t + \hat{\boldsymbol{\Lambda}} \hat{\boldsymbol{f}}_t)$. From the result of Theorem 2, we have: $T^{-1} \sum_{t=1}^T (y_{it} - \hat{\eta}_{it})^2 = T^{-1} \sum_{t=1}^T e_{it}^2 + o_p(1)$. Thus, the claim holds.

A.4 Proof of Theorem 4

The proof follows from the technical proof of Theorem 4 in Ando et al. (2021d) by replacing the quantile loss function with the squared loss function.

References

Ando, T. and Bai, J. (2020). Quantile co-movement in financial markets: A panel quantile model with unobserved heterogeneity. *Journal of the American Statistical Association*, 115:266–279.

Ando, T., Bai, J., and Li, K. (2021a). Bayesian and maximum likelihood analysis of large-scale panel choice models with unobserved heterogeneity. *Journal of Econometrics*, forthcoming.

Ando, T., Bai, J., Lu, L., and Vojtech, C. M. (2021b). Quantile connectedness of the U.S. interbank liquidity risk network: a bayesian nuclear estimation approach. *mimeo.,* Federal Reserve Bank, Boston.

Ando, T., Greenwood-Nimmo, M., and Shin, Y. (2021c). Quantile connectedness: Modeling tail behavior in the topology of financial networks. *Management Science*, forthcoming.

Ando, T., Li, K., and Lu, L. (2021d). A spatial panel quantile model with unobserved heterogeneity. *Journal of Econometrics*, forthcoming.

Aquaro, M., Bailey, N., and Pesaran, M. H. (2021). Estimation and inference for spatial models with heterogeneous coefficients: An application to US house prices. *Journal of Applied Econometrics*, 36:18–44.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71:135–171.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77:1229–1279.

Bai, J. and Li, K. (2014). Spatial panel data models with common shocks. *mimeo.,* University of Columbia.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70:191–221.

Bailey, N., Holly, S., and Pesaran, M. H. (2016). A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics*, 31:249–280.

Bosworth, B. and Collins, S. M. (2003). The empirics of growth: An update. *Brookings Papers on Economic Activity*, 2003:113–206.

Campbell, J. Y. and Mankiw, N. G. (1987). Are output fluctuations transitory? *The Quarterly Journal of Economics*, 102:857–880.

Chen, J., Shin, Y., and Zheng, C. (2021). Estimation and inference in heterogeneous spatial panels with a multifactor error structure. *Journal of Econometrics,* Forthcoming.

Chernozhukov, V., Fernandez-Val, I., and Weidner, M. (2020). Network and panel quantile effects via distribution regression. *Journal of Econometrics*.

Cohen, D., Soto, M., et al. (2004). Why are poor countries poor? *Technical Report,* Econometric Society 2004 Latin American Meetings.

Collins, S. M., Bosworth, B. P., and Rodrik, D. (1996). Economic growth in East Asia: accumulation versus assimilation. *Brookings Papers on Economic Activity*, 1996:135–203.

Cui, G., Sarafidis, V., and Yamagata, T. (2020). Iv estimation of spatial dynamic panels with interactive effects: Large sample theory and an application on bank attitude toward risk. *mimeo.,* Monash University.

Dees, S., di Mauro, F., Pesaran, M. H., and Smith, L. V. (2007). Exploring the international linkages of the Euro Area: a global VAR analysis. *Journal of Applied Econometrics*, 22:1–38.

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.

Ertur, C. and Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, 22:1033–1062.

Felipe, J. (1999). Total factor productivity growth in East Asia: A critical survey. *The Journal of Development Studies*, 35:1–41.

Granger, C. W. and Terasvirta, T. (1993). Modelling non-linear economic relationships. *OUP Catalogue*.

Greenwood-Nimmo, M., Nguyen, V., and Shin, Y. (2019). Quantifying informational linkages in a global model of currency spot markets. In Chevallier, J., Goutte, S., Guerreiro, D., Saglio, S., and Sanhaji, B., editors, *Advances in Applied Financial Econometrics: International Financial Markets*, volume 1. Routledge, London.

Greenwood-Nimmo, M., Nguyen, V. H., and Shin, Y. (2012). Probabilistic forecasting of output growth, inflation and the balance of trade in a GVAR framework. *Journal of Applied Econometrics*, 27:554–573.

Greenwood-Nimmo, M., Nguyen, V. H., and Shin, Y. (2021). Measuring the connectedness of the global economy. *International Journal of Forecasting*, 37:899–919.

Härdle, W. K., Wang, W., and Yu, L. (2016). Tenet: Tail-event driven network risk. *Journal of Econometrics*, 192:499–513.

Hirata, H., Kose, M. A., and Otrok, C. (2013). Regionalization vs. Globalization. *mimeo.,* International Monetary Fund.

Ho, C.-Y., Wang, W., and Yu, J. (2013). Growth spillover through trade: A spatial dynamic panel data approach. *Economics Letters*, 120:450–453.

Inklaar, R. and Timmer, M. P. (2013). Capital, labor and TFP in PWT8. 0. *mimeo.,* University of Groningen.

Kim, J.-I. and Lau, L. J. (1994). The sources of economic growth of the east asian newly industrialized countries. *Journal of the Japanese and International Economies*, 8:235–271.

Kose, M. A., Otrok, C., and Whiteman, C. H. (2003). International business cycles: World, region, and country-specific factors. *American Economic Review*, 93:1216–1239.

Kose, M. A., Otrok, C., and Whiteman, C. H. (2008). Understanding the evolution of world business cycles. *Journal of International Economics*, 75:110–130.

Lee, K., Pesaran, M. H., and Smith, R. (1997). Growth and convergence in a multi-country empirical stochastic Solow model. *Journal of Applied Econometrics*, 12:357–392.

LeSage, J. P. and Chih, Y.-Y. (2018). A bayesian spatial panel model with heterogeneous coefficients. *Regional Science and Urban Economics*, 72:58–73.

Lucas, R. E. (1990). Why doesn't capital flow from rich to poor countries? *American Economic Review*, 80:92–96.

Marukawa, T. (2021). Dependence and competition: Trade relationship between Asian countries and China. *Journal of Contemporary East Asia Studies*, 10:246–261.

Masten, M. A. (2018). Random coefficients on endogenous variables in simultaneous equations models. *Review of Economic Studies*, 85:1193–1250.

Mastromarco, C., Serlenga, L., and Shin, Y. (2016). Modelling technical efficiency in cross sectionally dependent stochastic frontier panels. *Journal of Applied Econometrics*, 31:281–297.

Nelson, C. R. and Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: Some evidence and implications. *Journal of Monetary Economics*, 10:139–162.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74:967–1012.

Pesaran, M. H., Schuermann, T., and Weiner, S. M. (2004). Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business & Economic Statistics*, 22:129–162.

Pesaran, M. H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters*, 58:17–29.

Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68:79–113.

Romer, P. (1993). Idea gaps and object gaps in economic development. *Journal of Monetary Economics*, 32:543–573.

Servén, L. and Abate, G. D. (2020). Adding space to the international business cycle. *Journal of Macroeconomics*, 65:103211.

Shi, W. and Lee, L.-f. (2017). Spatial dynamic panel data models with interactive fixed effects. *Journal of Econometrics*, 197:323–347.

Shin, Y. and Thornton, M. (2021). Dynamic network analysis via diffusion multipliers. *mimeo.,* University of York.

Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics*, pages 312–320.

Tahari, A., Akitoby, B., and Aka, E. B. (2004). Sources of growth in sub-Saharan Africa. *mimeo.,* International Monetary Fund.

Yang, C. F. (2021). Common factors and spatial dependence: An application to US house prices. *Econometric Reviews*, 40:14–50.

Young, A. (1995). The tyranny of numbers: confronting the statistical realities of the East Asian growth experience. *The Quarterly Journal of Economics*, 110:641–680.

Yu, J., De Jong, R., and Lee, L.-F. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both N and T are large. *Journal of Econometrics*, 146:118–134.

Zhu, X., Pan, R., Li, G., Liu, Y., and Wang, H. (2017). Network vector autoregression. *Annals of Statistics*, 45:1096–1123.

Zhu, X., Wang, W., Wang, H., and Härdle, W. K. (2019). Network quantile autoregression. *Journal of Econometrics*, 212:345–358.

Table 1: Performance of the information criteria in selecting the number of factors

N	T	$h=2$			$h=4$			$h=6$		
		$\hat{r}<r$	$\hat{r}=r$	$\hat{r}>r$	$\hat{r}<r$	$\hat{r}=r$	$\hat{r}>r$	$\hat{r}<r$	$\hat{r}=r$	$\hat{r}>r$
					IC_1					
50	50	0.000	0.867	0.133	0.000	0.843	0.157	0.000	0.830	0.170
	100	0.000	0.970	0.030	0.000	0.965	0.035	0.000	0.983	0.017
	200	0.000	1.00	0.000	0.000	1.00	0.000	0.000	0.998	0.002
100	50	0.000	0.980	0.020	0.000	0.983	0.0170	0.000	0.987	0.013
	100	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
	200	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
200	50	0.000	0.991	0.019	0.000	0.992	0.018	0.000	0.994	0.016
	100	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
	200	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
					IC_2					
50	50	0.000	0.973	0.0270	0.000	0.955	0.045	0.000	0.980	0.020
	100	0.000	0.993	0.0070	0.000	0.993	0.0070	0.000	1.00	0.000
	150	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
100	50	0.000	0.987	0.0130	0.000	0.993	0.0070	0.000	0.993	0.007
	100	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
	200	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
200	50	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
	100	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000
	200	0.000	1.00	0.000	0.000	1.00	0.000	0.000	1.00	0.000

Notes: $r = 2$ is the true number of unobserved factors. Using (12), we estimate r by setting $r_{\max} = 6$.

Table 2: Finite Sample Performance of Individual Estimators ($h = 4$)

		Bias			RMSE			Size			Power		
		$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$
	ϕ^*_{0i}	0.0031	0.0024	0.0026	0.0343	0.0147	0.0070	0.091	0.082	0.071	0.406	0.579	0.769
	ρ_i	-0.0078	-0.0041	-0.0021	0.0048	0.0022	0.0010	0.092	0.075	0.070	0.879	0.980	0.999
$N = 50$	ϕ^*_{1i}	0.0023	0.0013	-0.0006	0.0252	0.0109	0.0051	0.098	0.081	0.071	0.456	0.645	0.842
	β_{1i}	0.0003	-0.0015	-0.0003	0.0212	0.0095	0.0045	0.093	0.078	0.073	0.484	0.683	0.862
	β_{2i}	0.0009	-0.0013	-0.0003	0.0279	0.0127	0.0059	0.088	0.076	0.066	0.386	0.576	0.789
	ϕ^*_{0i}	0.0045	0.0025	0.0023	0.0329	0.0139	0.0065	0.089	0.075	0.066	0.392	0.571	0.760
	ρ_i	-0.0073	-0.0033	-0.0014	0.0047	0.0021	0.0010	0.085	0.067	0.059	0.876	0.977	0.999
$N = 100$	ϕ^*_{1i}	0.0016	0.0006	-0.0008	0.0240	0.0107	0.0049	0.086	0.072	0.060	0.442	0.643	0.838
	β_{1i}	-0.0003	-0.0006	-0.0008	0.0214	0.0092	0.0044	0.086	0.068	0.061	0.481	0.680	0.860
	β_{2i}	0.0013	0.0000	-0.0002	0.0273	0.0121	0.0059	0.082	0.068	0.060	0.379	0.571	0.787
	ϕ^*_{0i}	0.0027	0.0025	0.0018	0.0324	0.0137	0.0064	0.082	0.072	0.058	0.400	0.566	0.759
	ρ_i	-0.0067	-0.0037	-0.0022	0.0047	0.0021	0.0010	0.082	0.065	0.057	0.877	0.979	0.999
$N = 200$	ϕ^*_{1i}	0.0017	0.0005	0.0004	0.0245	0.0104	0.0049	0.087	0.069	0.061	0.448	0.638	0.829
	β_{1i}	-0.0010	-0.0002	-0.0011	0.0205	0.0092	0.0043	0.082	0.066	0.058	0.484	0.674	0.862
	β_{2i}	0.0013	-0.0006	-0.0006	0.0271	0.0123	0.0058	0.078	0.065	0.057	0.381	0.574	0.786

Notes: The simulation results are based on the DGP specified in Section 5. The variance of individual estimator is calculate according to 13, and for each parameter, say ϕ_i for example, the alternative for the power test is $\phi_i^a = \phi_i + 0.2$.

Table 3: Regional FEVD Effects by Continent Using $\boldsymbol{W}_{tra}$

	$h=0$						$h=2$						$h=5$					
	North America	Europe	Asia	Oceania	South America	Africa	North America	Europe	Asia	Oceania	South America	Africa	North America	Europe	Asia	Oceania	South America	Africa
Regional Connectedness Matrix																		
North America	0.9967	0.0005	0.0019	0.0000	0.0002	0.0000	0.996	0.0006	0.0022	0.0000	0.0004	0.0000	0.996	0.0006	0.0022	0.0000	0.0004	0.0000
Europe	0.0027	0.9943	0.0037	0.0000	0.0002	0.0001	0.0036	0.9932	0.0041	0.0000	0.0002	0.0001	0.0036	0.9932	0.0041	0.0001	0.0002	0.0001
Asia	0.0004	0.0002	0.9993	0.0001	0.0000	0.0000	0.0007	0.0003	0.9991	0.0001	0.0000	0.0000	0.0007	0.0003	0.999	0.0001	0.0000	0.0000
Oceania	0.0003	0.0001	0.0022	0.9892	0.0000	0.0000	0.0005	0.0001	0.0036	0.9823	0.0000	0.0000	0.0006	0.0001	0.0037	0.9819	0.0000	0.0000
Africa	0.0006	0.0001	0.0006	0.0000	0.9986	0.0000	0.0012	0.0001	0.0012	0.0000	0.9971	0.0000	0.0012	0.0001	0.0013	0.0000	0.997	0.0000
South America	0.0001	0.0007	0.0012	0.0000	0.0000	0.998	0.0004	0.0018	0.0028	0.0000	0.0001	0.995	0.0004	0.0018	0.0028	0.0000	0.0001	0.9949
Regional Effects																		
RSI	0.0026	0.0067	0.0007	0.0026	0.0012	0.002	0.0032	0.008	0.0011	0.0043	0.0026	0.005	0.0032	0.008	0.0011	0.0044	0.0026	0.0051
RSO	0.0041	0.0015	0.0097	0.0001	0.0004	0.0001	0.0064	0.0029	0.0138	0.0002	0.0007	0.0001	0.0066	0.0029	0.0139	0.0002	0.0007	0.0001
RNE	0.0015	-0.0052	0.009	-0.0025	-0.0008	-0.0019	0.0032	-0.0051	0.0128	-0.0041	-0.0019	-0.0049	0.0034	-0.0051	0.0129	-0.0042	-0.0019	-0.0049
EM	0.0026	0.0067	0.0007	0.0026	0.0012	0.002	0.0032	0.008	0.0011	0.0043	0.0026	0.005	0.0032	0.008	0.0011	0.0045	0.0026	0.0051
SI	0.1418	-0.4969	0.8582	-0.2396	-0.0785	-0.1851	0.2022	-0.3176	0.7978	-0.2578	-0.1192	-0.3054	0.2078	-0.3154	0.7922	-0.2613	-0.1199	-0.3034

Table 4: Regional FEVD Effects by Sub-Continent Using $\boldsymbol{W}_{tra}$

	Northern America	Western Europe	Other Europe	Eastern Asia	Other Asia	Pacific Islands	South America	Central America	Africa	South-Eastern Asia
					Regional Connectedness Matrix					
					$h = 0$					
Northern America	0.9843	0.0004	0	0.0041	0.0001	0	0.0001	0.0029	0	0.0002
Western Europe	0.0042	0.9924	0.0006	0.0049	0.0009	0	0.0002	0.0002	0.0001	0.0004
Other Europe	0.0011	0.0038	0.9896	0.0013	0.0001	0	0	0	0	0.0001
Eastern Asia	0.0004	0	0	0.9994	0	0	0	0	0	0.0002
Other Asia	0.0004	0.0007	0	0.0009	0.9979	0	0	0	0	0.0001
Pacific Islands	0.0008	0.0001	0	0.0054	0	0.9892	0	0	0	0.0007
South America	0.0008	0.0001	0	0.001	0	0	0.9986	0.0001	0	0
Central America	0.0134	0.0003	0	0.0017	0	0	0.0002	0.9896	0	0.0001
Africa	0.0002	0.0008	0	0.0014	0.0003	0	0	0	0.9981	0.0001
South-Eastern Asia	0.0006	0.0001	0	0.0032	0	0.0001	0	0	0	0.996
RSI	0.0078	0.0117	0.0064	0.0007	0.0022	0.0072	0.002	0.0157	0.0027	0.004
RSO	0.0219	0.0063	0.0007	0.0238	0.0015	0.0002	0.0005	0.0033	0.0001	0.002
RNE	0.0141	-0.0054	-0.0056	0.0231	-0.0007	-0.0069	-0.0015	-0.0124	-0.0026	-0.002
EM	0.0078	0.0116	0.0064	0.0007	0.0022	0.0072	0.002	0.0156	0.0027	0.004
SI	0.3782	-0.1454	-0.1516	0.6218	-0.0188	-0.1863	-0.0406	-0.3324	-0.0704	-0.0546
					$h = 5$					
Northern America	0.9825	0.0005	0	0.0045	0.0001	0	0.0001	0.0031	0	0.0003
Western Europe	0.0056	0.9911	0.0008	0.0053	0.0011	0.0001	0.0002	0.0003	0.0001	0.0005
Other Europe	0.0012	0.004	0.9889	0.0013	0.0001	0	0	0	0	0.0001
Eastern Asia	0.0009	0.0001	0	0.9986	0	0.0001	0	0	0	0.0004
Other Asia	0.0005	0.0008	0	0.0011	0.9974	0	0	0	0	0.0001
Pacific Islands	0.0014	0.0002	0	0.0082	0	0.9833	0	0	0	0.0012
South America	0.0015	0.0001	0	0.0021	0	0	0.9972	0.0002	0	0
Central America	0.0172	0.0004	0	0.0019	0	0	0.0003	0.9868	0	0.0001
Africa	0.0005	0.0019	0	0.003	0.0007	0	0.0001	0	0.9956	0.0003
South-Eastern Asia	0.0007	0.0001	0	0.0043	0	0.0001	0	0	0	0.9948
RSI	0.0086	0.0138	0.0067	0.0015	0.0027	0.0111	0.004	0.0199	0.0064	0.0052
RSO	0.0294	0.0081	0.0009	0.0316	0.0021	0.0003	0.0008	0.0038	0.0002	0.0029
RNE	0.0208	-0.0057	-0.0058	0.0301	-0.0006	-0.0108	-0.0033	-0.0161	-0.0063	-0.0024
EM	0.0087	0.0137	0.0068	0.0015	0.0027	0.0111	0.004	0.0198	0.0064	0.0052
SI	0.4086	-0.1114	-0.1147	0.5914	-0.0116	-0.2116	-0.064	-0.317	-0.1235	-0.0462

Table 5: Regional FEVD Effects by Income Using $\boldsymbol{W}_{tra}$

	$h=0$				$h=2$				$h=5$			
	High income	Upper-middle income	Lower-middle income	Low income	High income	Upper-middle income	Lower-middle income	Low income	High income	Upper-middle income	Lower-middle income	Low income
						Regional Connectedness Matrix						
High income	0.9965	0.004	0.0001	0.0000	0.9956	0.0051	0.0001	0.0000	0.9956	0.0051	0.0001	0.0000
Upper-middle income	0.0024	0.9969	0.0001	0.0000	0.0036	0.9954	0.0002	0.0000	0.0037	0.9953	0.0002	0.0000
Lower-middle income	0.0017	0.0013	0.9959	0.0000	0.0029	0.0024	0.9928	0.0000	0.0029	0.0024	0.9927	0.0000
Low income	0.0002	0.0007	0.0006	0.9986	0.0002	0.0008	0.0007	0.9984	0.0002	0.0008	0.0007	0.9984
						Regional Effects						
RSI	0.0041	0.0026	0.003	0.0014	0.0052	0.0038	0.0053	0.0016	0.0053	0.0039	0.0054	0.0016
RSO	0.0043	0.006	0.0008	0.0000	0.0067	0.0082	0.001	0.0000	0.0068	0.0083	0.001	0.0000
RNE	0.0002	0.0034	-0.0022	-0.0014	0.0015	0.0044	-0.0044	-0.0016	0.0016	0.0044	-0.0044	-0.0016
EM	0.0041	0.0026	0.003	0.0014	0.0052	0.0038	0.0053	0.0016	0.0053	0.0039	0.0054	0.0016
SI	0.047	0.953	-0.6211	-0.3789	0.255	0.745	-0.7356	-0.2644	0.2645	0.7355	-0.7371	-0.2629

Table 6: Regional TFP Effects by Continent Using $\boldsymbol{W}_{tra}$

	$h=0$						$h=2$						$h=5$					
	North America	Europe	Asia	Oceania	South America	Africa	North America	Europe	Asia	Oceania	South America	Africa	North America	Europe	Asia	Oceania	South America	Africa
Regional Connectedness Matrix																		
North America	1.2479	0.051	0.0823	0.0043	0.0204	0.0007	1.7715	0.0836	0.1222	0.0082	0.0234	-0.0007	1.8805	0.092	0.1319	0.0095	0.0251	-0.0006
Europe	0.1286	1.2388	0.1233	0.0076	0.0149	0.0087	0.2177	1.6444	0.1524	0.0123	0.019	0.0104	0.2319	1.68	0.1522	0.0131	0.019	0.0103
Asia	0.0691	0.0445	1.0326	0.0106	0.006	0.0024	0.0541	0.0359	1.1352	0.0071	0.0025	0.0016	0.0465	0.0339	1.1352	0.0052	0.0022	0.0016
Oceania	0.0303	0.0143	0.0583	0.9548	0.002	0.0006	0.0739	0.0313	0.1259	1.5085	0.0036	0.0011	0.0895	0.036	0.1436	1.6298	0.0038	0.0012
South America	0.0424	0.0119	0.025	0.0005	1.0418	0.0008	0.1111	0.0212	0.0485	0.0016	1.2371	0.0008	0.1241	0.0216	0.0488	0.0018	1.2455	0.0007
Africa	0.0156	0.0331	0.0483	0.0011	0.003	0.9605	-0.0219	-0.0299	-0.0036	-0.0018	-0.0048	1.0503	-0.0287	-0.0349	-0.0038	-0.0024	-0.0052	1.056
Regional Effects																		
RSI	0.1587	0.2831	0.1325	0.1056	0.0807	0.101	0.2367	0.4118	0.1013	0.2358	0.1832	-0.0621	0.258	0.4266	0.0894	0.2741	0.197	-0.075
RSO	0.286	0.1547	0.3373	0.0241	0.0464	0.0132	0.4348	0.1422	0.4454	0.0273	0.0438	0.0132	0.4634	0.1486	0.4727	0.0272	0.045	0.0132
RNE	0.1273	-0.1284	0.2048	-0.0815	-0.0343	-0.0879	0.1981	-0.2696	0.3441	-0.2085	-0.1394	0.0752	0.2054	-0.278	0.3833	-0.2469	-0.152	0.0882
EM	0.1128	0.186	0.1137	0.0996	0.0719	0.0952	0.1178	0.2002	0.0819	0.1352	0.129	-0.0558	0.1206	0.2025	0.073	0.144	0.1366	-0.0663
SI	0.3834	-0.3868	0.6166	-0.2454	-0.1033	-0.2646	0.3209	-0.4366	0.5573	-0.3376	-0.2258	0.1219	0.3035	-0.4107	0.5663	-0.3647	-0.2245	0.1302

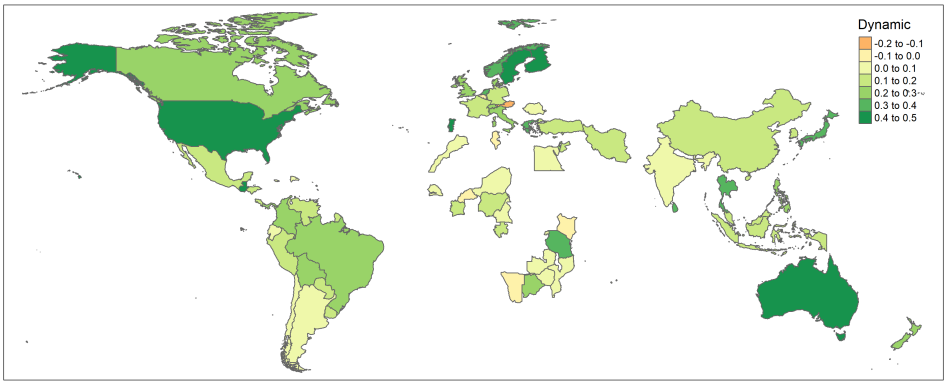
Table 7: Regional TFP Effects by Sub-Continent Using $\boldsymbol{W}_{tra}$

	Northern America	Western Europe	Other Europe	Eastern Asia	Other Asia	Pacific Islands	South America	Central America	Africa	South-Eastern Asia
					Regional Connectedness Matrix					
					$h = 0$					
Northern America	1.5427	0.0263	0.0063	0.0939	0.0073	0.0051	0.0068	0.0263	0.0008	0.0091
Western Europe	0.1725	1.1838	0.0421	0.128	0.0294	0.0076	0.0141	0.0135	0.0082	0.0174
Other Europe	0.0955	0.1099	1.0193	0.0668	0.0112	0.0054	0.0055	0.0045	0.0018	0.0072
Eastern Asia	0.0353	0.0064	0.0014	1.186	0.0031	0.0042	0.0011	0.0022	0.0002	0.0079
Other Asia	0.0778	0.0693	0.0098	0.0515	0.9712	0.0044	0.0044	0.0043	0.0024	0.011
Pacific Islands	0.0833	0.0173	0.0045	0.1323	0.0059	0.9548	0.002	0.0031	0.0006	0.0206
South America	0.0578	0.013	0.0024	0.0336	0.0066	0.0005	1.0418	0.0085	0.0008	0.003
Central America	0.3965	0.0432	0.0073	0.0932	0.0099	0.0031	0.0182	0.91	0.0001	0.0074
Africa	0.0198	0.0355	0.0039	0.0433	0.0189	0.0011	0.003	0.0013	0.9605	0.0063
South-Eastern Asia	0.1079	0.0291	0.0048	0.1957	0.0122	0.0173	0.005	0.0056	0.0013	0.7817
RSI	0.1819	0.4327	0.3079	0.0618	0.2349	0.2696	0.1262	0.5789	0.1332	0.3789
RSO	1.0464	0.3499	0.0825	0.8383	0.1046	0.0488	0.0601	0.0693	0.0162	0.0899
RNE	0.8645	-0.0828	-0.2254	0.7764	-0.1303	-0.2209	-0.0661	-0.5096	-0.117	-0.2889
EM	0.1055	0.2677	0.232	0.0495	0.1948	0.2202	0.1081	0.3888	0.1218	0.3264
SI	0.5268	-0.0505	-0.1373	0.4732	-0.0794	-0.1346	-0.0403	-0.3106	-0.0713	-0.1761
					$h = 0$					
Northern America	2.4347	0.0468	0.0127	0.1561	0.0118	0.0104	0.0115	0.0445	0.0012	0.0146
Western Europe	0.3099	1.5305	0.0714	0.1815	0.0436	0.0134	0.0201	0.021	0.011	0.0238
Other Europe	0.0935	0.0866	1.4317	0.0501	0.008	0.005	0.004	0.0038	0.0012	0.0051
Eastern Asia	-0.0132	-0.0021	-0.0006	1.345	-0.0002	-0.0027	-0.0018	-0.0007	-0.0003	-0.0006
Other Asia	0.1161	0.0557	0.009	0.0395	1.1334	0.0067	0.0044	0.0051	0.0021	0.0119
Pacific Islands	0.1908	0.0344	0.0102	0.2677	0.011	1.4314	0.0034	0.0061	0.001	0.0392
South America	0.1426	0.0219	0.0047	0.0645	0.0119	0.0015	1.2126	0.0171	0.0009	0.0048
Central America	0.6328	0.0615	0.0119	0.1223	0.0121	0.0051	0.0182	1.0958	-0.0015	0.0093
Africa	-0.0201	-0.0216	-0.0017	-0.0024	0.0051	-0.0014	-0.0037	-0.0013	1.0393	0.0003
South-Eastern Asia	0.055	0.0135	0.0027	0.0718	0.0051	0.0124	0.0021	0.0026	0.0006	0.8961
RSI	0.3095	0.6956	0.2573	-0.0221	0.2505	0.5639	0.2698	0.8719	-0.0468	0.1658
RSO	1.5076	0.2967	0.1203	0.9512	0.1084	0.0504	0.0582	0.0982	0.0162	0.1084
RNE	1.1981	-0.3989	-0.137	0.9733	-0.1421	-0.5135	-0.2116	-0.7737	0.0629	-0.0575
EM	0.1128	0.3125	0.1524	-0.0162	0.181	0.2826	0.182	0.4424	-0.0426	0.1562
SI	0.5362	-0.1786	-0.0613	0.4356	-0.0636	-0.2298	-0.0947	-0.3463	0.0282	-0.0257

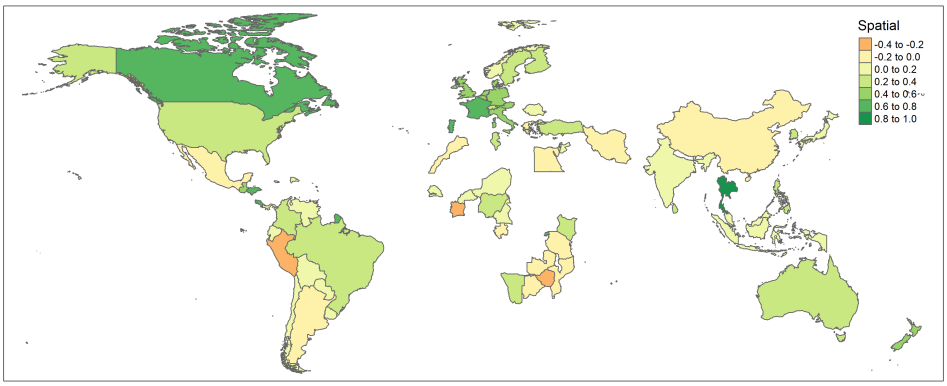
Table 8: Regional TFP Effects by Income Using $\boldsymbol{W}_{tra}$

	$h=0$				$h=2$				$h=5$			
	High income	Upper-middle income	Lower-middle income	Low income	High income	Upper-middle income	Lower-middle income	Low income	High income	Upper-middle income	Lower-middle income	Low income
						Regional Connectedness Matrix						
High income	1.3016	0.1056	0.0196	0.0005	1.768	0.1478	0.0238	0.0005	1.8286	0.1528	0.0241	0.0005
Upper-middle income	0.1256	1.0091	0.0094	0.0001	0.1781	1.184	0.0124	0.0003	0.1935	1.1988	0.0128	0.0004
Lower-middle income	0.0806	0.0461	0.8782	-0.0013	0.0137	0.0148	0.9841	-0.003	0.005	0.0134	0.9925	-0.003
Low income	0.0216	0.0226	0.0248	0.963	0.0131	0.0115	0.0134	0.993	0.0125	0.0112	0.0131	0.9932
						Regional Effects						
RSI	0.1257	0.1351	0.1253	0.069	0.1722	0.1908	0.0256	0.0381	0.1775	0.2066	0.0154	0.0369
RSO	0.2278	0.1742	0.0537	-0.0007	0.2049	0.1741	0.0496	-0.0021	0.211	0.1774	0.0501	-0.0021
RNE	0.1021	0.0391	-0.0716	-0.0697	0.0328	-0.0167	0.0241	-0.0401	0.0335	-0.0292	0.0347	-0.039
EM	0.088	0.1181	0.1246	0.0668	0.0887	0.1388	0.0252	0.0369	0.0885	0.147	0.0151	0.0358
SI	0.7229	0.2771	-0.507	-0.493	0.5764	-0.2936	0.4236	-0.7064	0.491	-0.4284	0.509	-0.5716

Figure 1: The Spatial Patterns of Individual Coefficients Using $\boldsymbol{W}_{tra}$

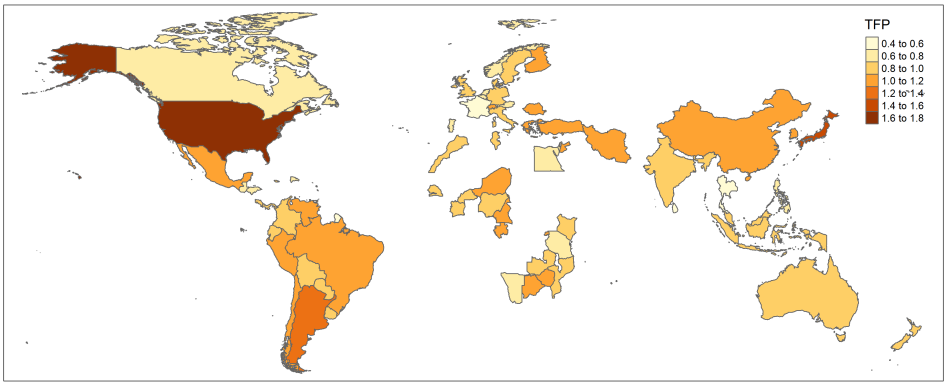


(a) Individual Estimation for ρ_i



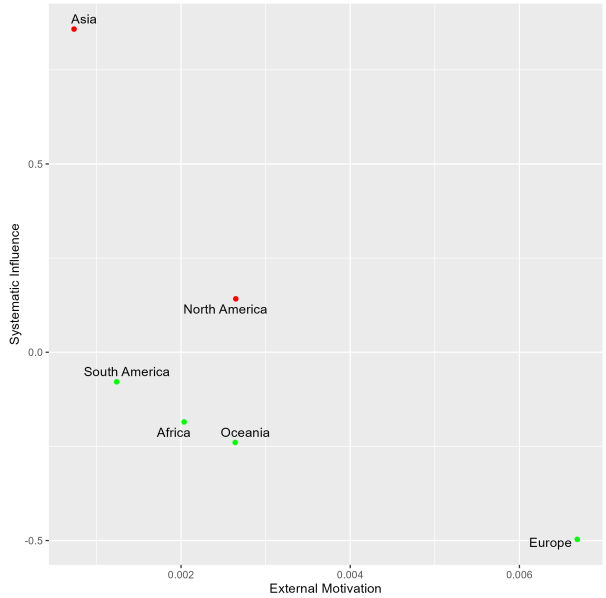
(b) Individual Estimation for ϕ_{i0}

(c) Individual Estimation for ϕ_{i1}

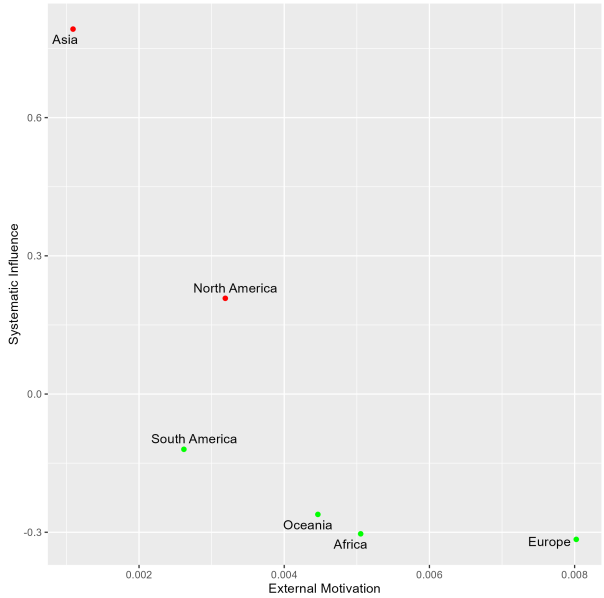


(d) Individual Estimation for β_i

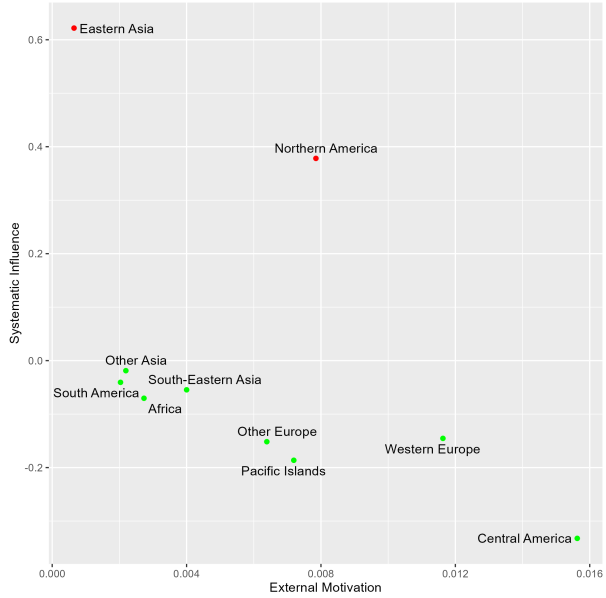
Figure 2: Regional/Group Network Analysis: FEVD



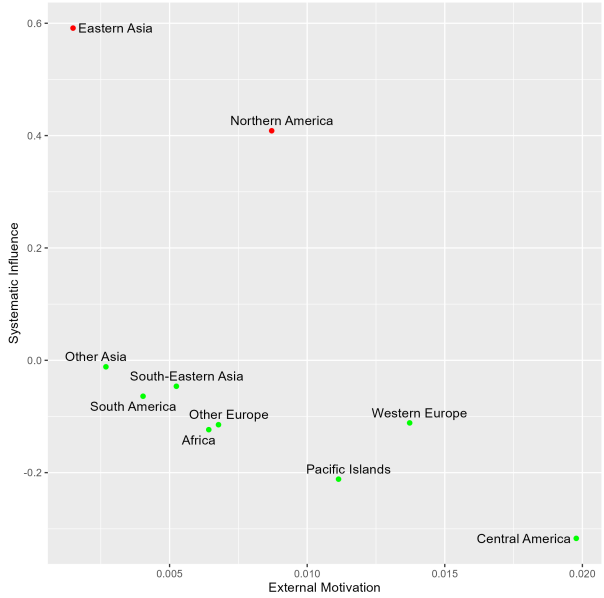
(a) Grouping by Continent, $h = 0$



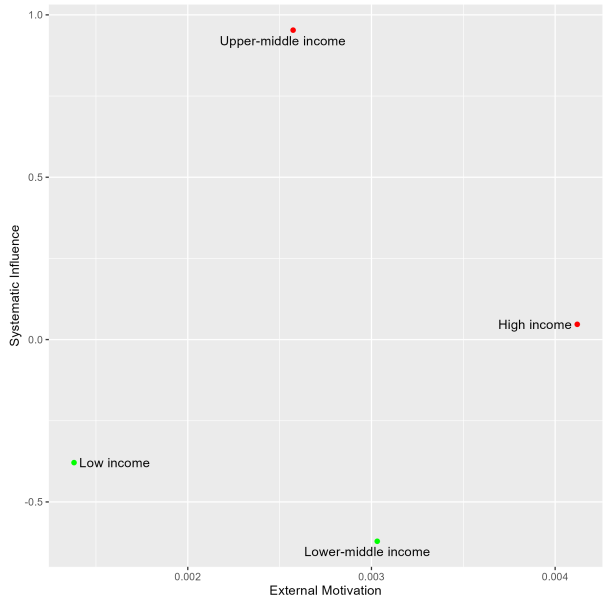
(b) Grouping by Continent, $h = 5$



(c) Grouping by Subregion, $h = 0$



(d) Grouping by Subregion, $h = 5$



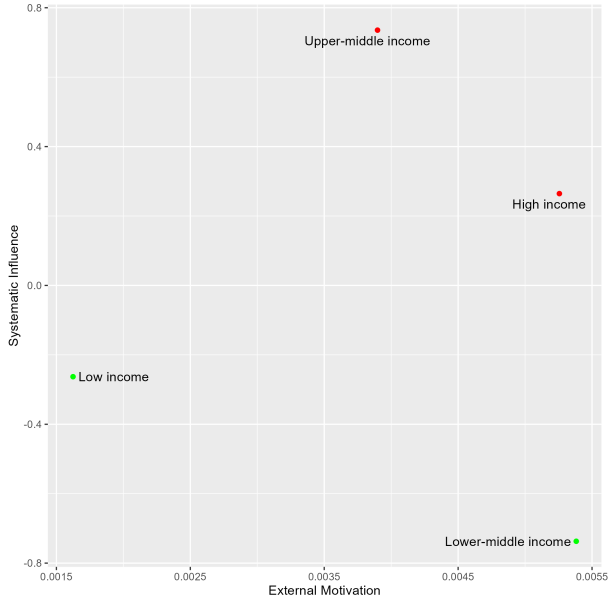
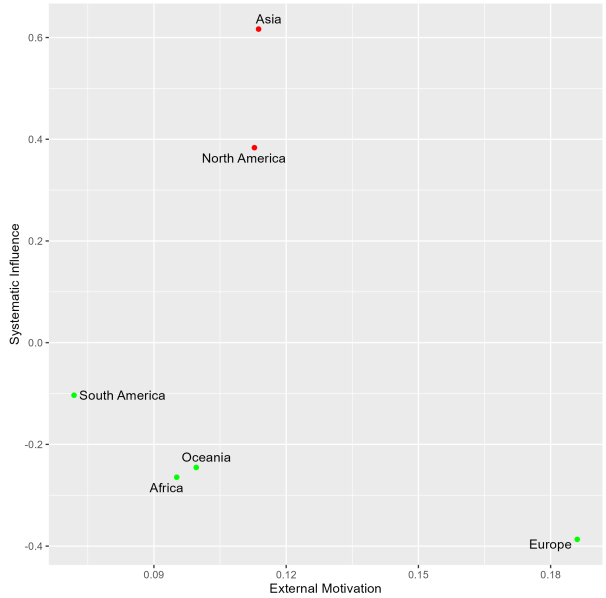
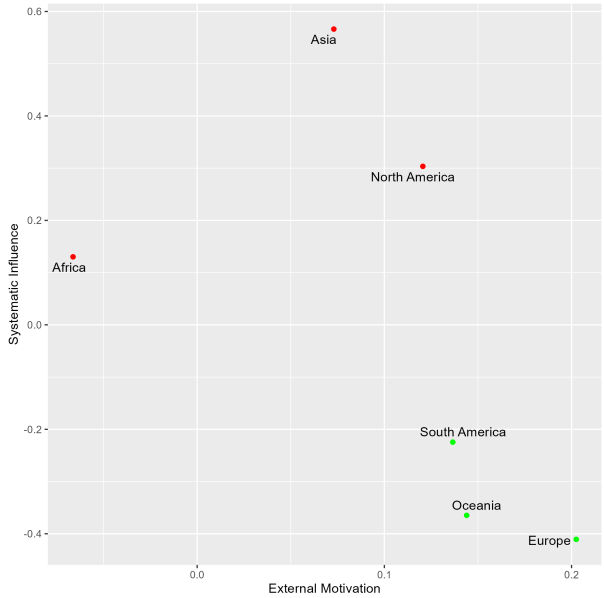


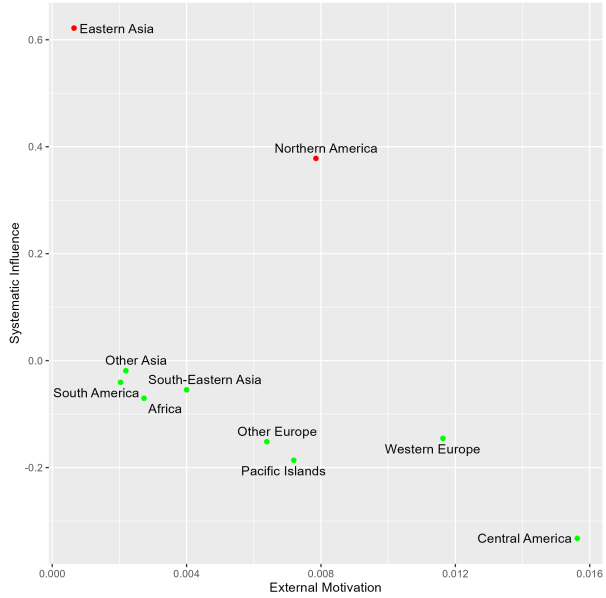
Figure 3: Regional/Group Network Analysis: TFP



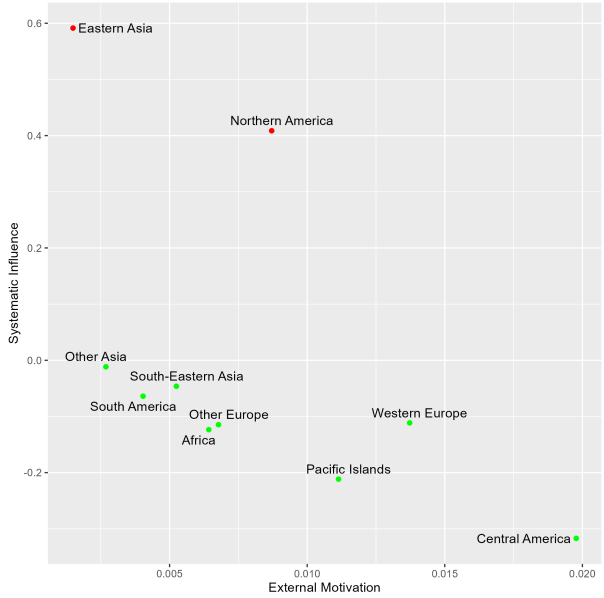
(a) Grouping by Continent, $h = 0$



(b) Grouping by Continent, $h = 5$



(c) Grouping by Subregion, $h = 0$



(d) Grouping by Subregion, $h = 5$