Estimation of Industry-level Productivity with Cross-sectional Dependence by Using Spatial Analysis

Jaepil Han *　　Robin C. Sickles †

January 24, 2019

Abstract

We examine aggregate productivity in the presence of inter-sectoral linkages. Cross-sectional dependence is inevitable among industries, in which each sector serves as a supplier to the other sectors. However, the chains of such interconnections cause indirect relationship among industries. Spatial analysis is one of the approaches to address cross-sectional dependence by using *a priori* a specified spatial weights matrix. We exploit the linkage patterns from the input–output tables and use them to assign spatial weights to describe the economic interdependencies. By using the spatial weights matrix, we can estimate the industry-level production functions and productivity of the U.S. from 1947 to 2010. Cross-sectional dependencies are the consequences of indirect effects, and they reflect the interactions among industries linked via their supply chain networks result in larger output elasticities as well as scale effects for the networked production processes. However, productivity growth estimates are reportedly comparable across various spatial and non-spatial model specifications.

Keywords: cross-sectional dependence, spatial panel model, spatial weights matrix, stochastic frontier analysis, industry-level productivity
JEL Classification: C21, C23, C51, O47, R15

1 Introduction

A nation's aggregate productivity can be decomposed into the productivity of each industry and the allocation of factor inputs among industries. These two components may interact with one another. A productivity shock in one sector may result in misallocation of inputs in the other sectors, while the misallocation of inputs in a sector may invoke productivity shocks in the other sectors due to production linkages via the production supply chain. Such linkages have been studied extensively in the literature. Timmer et al. (2014, 2015), and Timmer and Ye (2017) summarized and provided new insights into many of the productivity aspects of supply chain networks. Jones (2012) argued that the effects of misallocation may be amplified through the input–output structure, and that the contagion of the negative effects caused by misallocations can reduce total factor productivity

*Correspondence: Korea Development Institute, Sejong, S. Korea, Email: jaepil.han@kdi.re.kr
†Department of Economics, Rice University, Houston, TX, USA, Email: rsickles@rice.edu

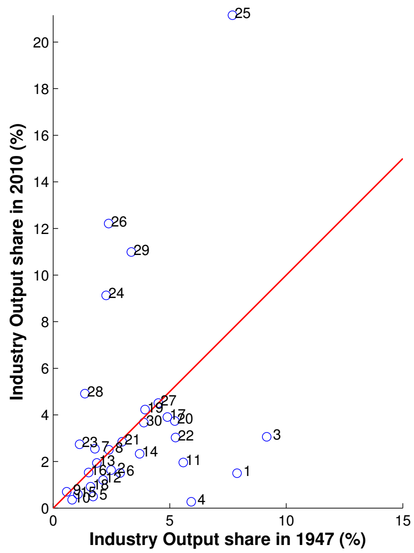
(TFP). Similarly, Acemoglu et al. (2012) believed that microeconomic idiosyncratic shocks can lead to aggregate fluctuations through input–output linkages. Hence, the productivity of each industry may be dependent on the productivity of other industries.

Measurement of industry-level productivity has been examined in many studies by using various methodologies. However, most studies have not considered the possibility of cross-sectional dependencies. Much of the current research has focused on the contributions of industries to aggregate productivity. Growth–accounting techniques and index–number approaches are widely used by many statistical agencies. However, these approaches separately estimate the productivity growth of each industry. Statistical agencies also assume that industries are independent, thus neglecting the possibility of interdependencies. However, interdependency is inevitable, especially among economic industries, in which the outputs of many sectors are used as inputs of others.

Substantial evidence supports the contributions of each industry to aggregate productivity, which changes over time. The contribution of each industry to the economy can be defined as the productivity for the industry weighted by its share in output for the economy as a whole (Jorgenson et al., 2012). Hence, changes in the level of industry productivity and its relative importance in the economy will impact its contribution to aggregate productivity growth. In addition, the relative importance of industries may be affected by the linkages among industries. For example, Figure 1 shows the evolution of the share of industries to gross output of the U.S. by comparing shares from 1947 to 2010[1]. Clearly, the gross output share of each industry has changed substantially over the last six decades. Total manufacturing, e.g., *food, beverages, and tobacco* (Ind3), *textiles, leather, and footwear* (Ind4), and *basic metals and fabricated metal* (Ind11), lost their shares. By contrast, tertiary industries, e.g., *financial intermediation* (Ind24), *real estate, renting, and business activities* (Ind25, Ind26), *education*(Ind 28), and *health and social work* (Ind29), increased their shares. The change in output shares is a result of industrial structural change in general. However, the change can also represent the transition of key sectors. Jorgenson et al. (2012) also argued the possible influential power of several key sectors. For instance, the role of the non-IT industries has shrunk, whereas the contributions of IT-producing and IT-using industries have risen.

The research on productivity measurement has a long history. Modern methods and approaches are usually traced back to the pioneering work of Cobb and Douglas (1928). The ambiguous concept of productivity was clarified owing to contributions of Solow (1957), Griliches (1960), and Jorgenson and Griliches (1967). Solow (1957) measured productivity growth by using the residuals of output growth that have not been explained by capital accumulation or increased labor services. The field has developed in various directions, and index–number approaches have become the main metric for measuring productivity growth (Jorgenson and Griliches, 1967; Diewert, 1976). Direct estimation of the production function by using econometric techniques has also been pursued. Hulten (2001) reported that a number of pitfalls are present in the econometric approach. These pitfalls include the possible odd-shaped isoquants without appropriate *a priori* restrictions, the abundance of parameters to estimate with limited information, and the complication of flexible models. Although the

[1]Source: World KLEMS database (`http://www.worldklems.net/`)



Note: The industries represented by the numbers should refer to the Table C.1

Figure 1: A comparison of Gross-output share of industries: 1947 vs. 2010

econometric approach has its shortcomings, it still offers model flexibility. Such flexibility extends to the econometric setting in which panel data are utilized, thereby employing fixed or random effects specifications to address firm-specific unobserved heterogeneity. However, the econometric approach often does not address spatial interactions. Exceptions in the production function literature can be found in the endogenous growth literature and formal spatial econometric specifications based on both average production/cost models and frontier production/cost models. Models extending the multiplicative spillover effects by framing production in a spatial autoregressive (SAR) setting to address network effects or trade flows among countries have been formulated by Ertur and Koch (2007) and Behrens et al. (2012). General stochastic frontier treatments that do not force efficiency on the productive units—whether they are countries, states, or firms—have been introduced by Druska and Horrace (2004) in the cross-sectional setting and for the panel modeling, as shown in the series of papers by Glass et al. (2013, 2015, 2016a,b) and by Han et al. (2016a,b).

The ignored cross-sectional dependence has caused the standard OLS estimators to become inefficient and the estimated standard errors to become biased (Phillips and Sul, 2003; Chudik and Pesaran, 2013). We expand the standard stochastic frontier model to spatially dependent specifications in the present study.

Models with spatial structures are found in the fields of regional science, economic geography, and urban economics, and such methodology has been applied recently in panel data studies in international economics, public economics, and agricultural economics. The approach captures the structure of cross-sectional correlation with the exogenous spatial weights matrix and estimates the spatial effects via spatial parameters. The spatial weights matrix is usually formed on the basis of spatial (geographical) characteristics. However, the characteristics also can be based on economic or socio-economic distance among units. A distinctive feature of the spatial econometric approach is that the spatial weights matrix is generally specified *a priori* based on an exogenous conceptualization of the structure of spatial dependence. Hence, selecting the right spatial weights matrix is crucial for correct model specification.

We aim to propose a novel approach to create a spatial weights matrix based on economic distance when physical distance does not properly capture the spatial linkages. We examine industry-specific data. Thus, a formal distance metric based on geographical distance has no obvious appeal. Instead, we define economic distance, analogous to a geographic distance by using the supply flows that can be found in the input–output tables. In succeeding sections, we specify spatial production models by utilizing SAR models and spatial Durbin model (SDM) and estimate the model by using Cornwell et al. (1990) type stochastic frontier approaches. By using the weights matrix created from the input–output tables, we examine the production technology and industry-level productivity of the U.S. from 1947 to 2010.

The remainder of the paper is organized as follows. In section 2, we start with a typical production model of the economy and expand the model to associated spatial specifications. We also discuss how to estimate the production technology of the economy and the efficiency score of each industry within the stochastic frontier framework. In section 3, we provide a novel methodology to define a supply chain-based metric for economic distance by using the input–output tables and form the corresponding spatial weights matrix. In section 4 we apply our methodology to estimate a production frontier function by using industry-level data for the U.S. from 1947 to 2010. Finally, we discuss concluding remarks and possible extensions for the future research in section 5.

2 Models

2.1 Production function with heterogeneity in intercepts and slopes

Consider the following production function with Hicks-neutral technological change:

$$Y_{it} = A_{it} F(\mathbf{X}_{it}), \quad i = 1, \cdots, N, \quad t = 1, \cdots, T, \tag{1}$$

where A_{it} is the unobservable productivity term, which differs between economic units and time periods. If we consider a log-linear functional form (e.g., Cobb–Douglas production function) for $F(\cdot)$, then by taking the logarithm of Eq. (1) we derive the following standard linear regression model:

$$y_{it} = \alpha + X_{it}'\beta + \varepsilon_{it},$$

where y_{it} and X_{it} are the logged variables of Y_{it} and $\mathbf{X}_{it}$, α represents the intercept, and ε_{it} is an unobservable term. The ε_{it} term can be decomposed into a productivity term u_{it}, which is known to the firm, or a representation of firms in an industry that is well-informed, but not known to the econometrician, and a statistical noise term v_{it}. Hence, the model becomes

$$y_{it} = \alpha + X_{it}'\beta + u_{it} + v_{it}. \tag{2}$$

Cornwell et al. (1990) extended the model (Eq. (2)) to allow heterogeneity in slopes and intercepts, and they modeled the time-dependent individual effect α_{it} in the following multiplicative form:

$$\alpha_{it} = \alpha + u_{it} = R_t'\delta_i, \tag{3}$$

where R_t is an $L \times 1$ time-varying component that globally affects all individual units, and δ_i denotes $L \times 1$ coefficients that depend on i. The time-dependent individual effect α_{it} can be decomposed into a common time trend $R_t'\delta_0$ and a unit-specific term $R_t'u_i$. After adding the time-invariant fixed effects Z_i for a general model specification the standard log-linear production function, Eq. (2) can be written as

$$y_{it} = X_{it}'\beta + Z_i'\gamma + R_t'\delta_0 + R_t'u_i + v_{it}, \qquad i = 1, \cdots, N, \quad t = 1, \cdots, T, \tag{4}$$

where Z_i is a $J \times 1$ vector, u_i is assumed as *iid* zero mean random variables with covariance matrix Δ, and v_{it} is a random noise following *iid* $N(0, \sigma_v^2)$. If R contains only a constant, for instance, then Eq. (4) is reduced to the standard panel data model.

2.2 Spatial production models

The model in Eq. (4) is misspecified if cross-sectional dependencies exist in the error terms. It is well-known that estimators are inefficient and estimated standard errors are biased if cross-sectional dependencies in the error term are ignored. If the structural model needs to be represented by a SAR model, then ignoring such spatial linkages renders the estimates biased and inconsistent. Spatial analysis is one of the approaches that can explicitly address cross-sectional dependence. Spatial econometric models specify an *a priori* spatial weights matrix that can capture the explicit dependence structure, which is assumed to exist among units. On the basis of the exogenous spatial weights matrix, the spatial approaches essentially include new variables consisting of the weighted averaged variables of neighboring observations. The omission of such spatially correlated variables will cause omitted variable bias for the estimation of the coefficient parameters.

Spatial econometric approaches can model possible spatial interactions by determining the type and extent of spatial dependence that exists among economic units. Three types of spatial dependence are usually considered to reflect the interactions that should be modeled by the inclusion of spatially weighted dependent variables (i.e., Spatial Autoregressive Model; SAR), spatially weighted independent variables (i.e., Spatial lagged-X regression model; SLX), or spatial dependence captured in the error term (i.e., Spatial Error Model; SEM). Allowing more than one dependence structure is possible. In the present study, we consider either a spatially weighted dependent variable (i.e., SAR) or both spatially weighted dependent variable and independent variables (i.e., Spatial Durbin Model; SDM) to avoid problems with over-parameterization and to focus on a relatively parsimonious setting, thus allowing us to explore our new approach when constructing the weights matrix.

The SAR specification associated with Eq.(4) is:

$$y_{it} = \rho \sum_{j=1}^{N} w_{ij} y_{jt} + X_{it}'\beta + Z_i'\gamma + R_t'\delta_0 + R_t'u_i + v_{it}, \tag{5}$$

where w_{ij} is the ijth element of $(N \times N)$ spatial weights matrix W_N, to be given exogenously, u_i is assumed as *iid* zero mean random variables with covariance matrix Δ, and v_{it} is a random noise following $N(0, \sigma_v^2)$. The matrix form of Eq.(5) is given by[2]:

$$y = \rho(W_N \otimes I_T)y + X\beta + \boldsymbol{Z}\gamma + \boldsymbol{R}\delta_0 + QU + V, \tag{6}$$

where y and V are $NT \times 1$ vectors, X is an $NT \times K$ matrix, $\boldsymbol{Z} = (Z \otimes \iota_T)$, Z is an $N \times J$ matrix, ι_T is a T dimensional vector of ones, $\boldsymbol{R} = (\iota_N \otimes R)$, $R = (R_1, R_2, \cdots, R_T)'$, $Q = \iota_N \otimes diag(R)$ is an $NT \times LN$ matrix, β is a $K \times 1$ vector, γ is a $J \times 1$ vector, δ_0 is an $L \times 1$ vector, and U is an $LN \times 1$ vector.

The Spatial Durbin specification associated with Eq.(4) is:

$$y_{it} = \rho \sum_{j=1}^{N} w_{ij} y_{jt} + X_{it}'\beta + \sum_{j=1}^{N} w_{ij} X_{jt}'\lambda + Z_i'\gamma + R_t'\delta_0 + R_t'u_i + v_{it}, \tag{7}$$

where w_{ij} is the ijth element of $(N \times N)$ spatial weights matrix W_N, to be given exogenously, u_i is assumed as *iid* zero mean random variables with covariance matrix Δ, and v_{it}, is a random noise following $N(0, \sigma_v^2)$[3]. In addition, the matrix form of Eq.(7) is given by:

$$y = \rho(W_N \otimes I_T)y + X\beta + (W_N \otimes I_T)X\boldsymbol{\lambda} + \boldsymbol{Z}\gamma + \boldsymbol{R}\delta_0 + QU + V, \tag{8}$$

where y and V are $NT \times 1$ vectors, X is an $NT \times K$ matrix, $\boldsymbol{Z} = (Z \otimes \iota_T)$, Z is $N \times J$ matrix, ι_T is a T dimensional vector of ones, $\boldsymbol{R} = (\iota_N \otimes R)$, $R = (R_1, R_2, \cdots, R_T)'$, $Q = \iota_N \otimes diag(R)$ is an $NT \times LN$ matrix, β is a $K \times 1$ vector, γ is a $J \times 1$ vector, δ_0 is an $L \times 1$ vector, and U is an $LN \times 1$ vector.

[2]The observations stacked with t being the fast-running index and i the slow-running index, i.e., $y = (y_{11}, y_{12}, \cdots, y_{1T}, \cdots, y_{N1}, \cdots, y_{NT})'$. The order of observations is very important for writing correct codes. In typical spatial analysis literature, the slower index is over time, the faster index is over individuals.

[3]We may want to specify different spatial correlation structures on dependent variable and independent variables. However, we use the same dependence structure for both variables.

3 Spatial weights matrix: economic distance

The spatial weights matrix is a crucial element in spatial econometric modeling, and it is often specified *a priori* and assumed to be an exogenous conceptualization of the structure of spatial dependence. According to Getis and Aldstadt (2004), a model with a wrong choice of spatial weights matrix is essentially inappropriately-specified. Unlike spatial weight matrices in typical spatial analyses that rely on geographic relationships, we are unable to define the physical distances among units because the aggregate production function is examined at the industry level. Thus we need to define an economic distance measure for the construction of the spatial weights matrix.

3.1 Input–output table and multiplier product matrix

The economic distance measure should be able to exploit the interconnectivity between a pair of individual units. For example, Han et al. (2016b) used a country's relative bilateral trade volume as an economic distance measure. In this paper, we consider the relationship among industries, thus the economic relationships among industries should be initially understood. An input–output table is appropriate for this purpose. Input–output analysis was developed by Leontief in the late 1930s, and its fundamental purpose is to analyze the interdependence of industries in an economy. An input–output table is constructed by using observed data for a particular economic area. Traditionally the area includes a nation or a state, but recently there are attempts to reflect increasing fragmentation of production processes across borders by the World Input-Output Database (WIOD).

An input–output table contains data on the flows of products from each sector to the other sectors. Inter-sectoral flows are measured in monetary terms, that is, accounts are presented in their monetary terms in anticipation of measurement problems that may arise if different products with varying prices are considered in terms of their physical units only. Because the dimension of the input–output table depends on the number of industries, we will explain how we construct the spatial weights matrix by using the input–output table via a hypothetical example with two sectors.

Table 1 presents a hypothetical situation for a two-sector economy. The total output of sector 1 (Y_1) is supposed to be consumed by each sector (z_{11}, z_{12}), household consumers, government, and abroad (F_1). Because the input–output table represents an equilibrium state, total sector output is assumed to be fully utilized by the sectors that require the products as intermediate inputs and by the final consumer including households, government, and abroad. On the other hand, the total outlay of sector 1 (Y_1) is composed of the total value of intermediate inputs purchased from the other sectors and the input factors such as labor and capital. Therefore, the total outlay is assumed to be the sum of intermediate inputs used by a sector and the value-added (VA). After considering the appropriate inventory changes, the total sector output is equal to the total outlays of the industry.

It should be noted that the components of an input–output table cannot directly represent the

Table 1: Input-Output Table for Two-sector Economy

		Demanding Sectors		Final Demand	Total Sector Output
		1	2	(F)	(Y)
Supplying	1	z_{11}	z_{12}	F_1	Y_1
Sectors	2	z_{21}	z_{22}	F_2	Y_2
Payments	Value-added	VA_1	VA_2	VA_F	VA
Total Outlays (Y)		Y_1	Y_2	F	Y

linkage of paired industries; thus, the industries cannot be compared directly with one another because their scales have not been adjusted. Even if the scaling effects are adjusted, the relative importance of a sector to the other sectors may be ambiguous depending on which of total sector outputs or total outlays we are using for the normalization. Suppose that $z_{11} = 2$, $z_{12} = 3$, $z_{21} = 1$, and $z_{22} = 5$. In this example, sector 1 is more important than sector 2 to sector 1 in terms of outlay of intermediate inputs (2/3 vs. 1/3). However, sector 2 is more important than sector 1 to sector 1 in terms of the demander of the production from sector 1 (2/5 vs. 3/5). In fact, these are the input and output direct requirements defined in Eqs.(9) and (10), respectively. The coefficients of input direct requirements represent the amount of inputs that is purchased directly to produce one dollar of output, whereas the coefficients of output direct requirements represent the proportion of outputs of a specific sector that is sold to each purchasing sector. Each of the direct requirements captures direct and indirect relationships, but we lose one of demand-side and supply-side relationships without considering them together.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \frac{z_{11}}{Y_1} & \frac{z_{12}}{Y_2} \\ \frac{z_{21}}{Y_1} & \frac{z_{22}}{Y_2} \end{bmatrix}, \tag{9}$$

$$\bar{A} = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} \\ \bar{a}_{21} & \bar{a}_{22} \end{bmatrix} = \begin{bmatrix} \frac{z_{11}}{Y_1} & \frac{z_{12}}{Y_1} \\ \frac{z_{21}}{Y_2} & \frac{z_{22}}{Y_2} \end{bmatrix}. \tag{10}$$

Meanwhile, from the definitions of the input–output table, we have the following two relations: $Y = Zi + F$, and $Y = i'Z + VA$, where i is a vector of ones. By using the relations and the proportional relationships, we can then establish the relationships among Y, F, A, and $\bar{A}$ as

$$Y = (I - A)^{-1}F, \tag{11}$$

$$Y = (I - \bar{A})^{-1}VA. \tag{12}$$

The inverse matrices in Eqs. (11) and (12), which describe the relationship among final consumption, value-added, and outputs (or outlays), are called *Leontief inverse matrix* and *Ghosh inverse matrix*, respectively. Let us denote *Leontief inverse matrix* as $B = (I - A)^{-1}$ and *Ghosh inverse matrix* as $\bar{B} = (I - \bar{A})^{-1}$. The *Leontief inverse matrix* illustrates how much production will be induced in a certain industry by one unit of increase in demand by another industry. The *Ghosh inverse matrix* is derived when we consider the output direct requirements matrix ($\bar{A}$) and the relationship between intermediate input supply and payments of each sector from the supply side. The basic assumption of the supply-side inverse matrix is that one may expect increasing sales from sector i to each of the other sectors when the output of sector i increases.

The attempts to determine a unified measure for the economic effects of a particular sector on the other sectors have been undertaken, and backward and forward linkages are the most commonly used measures. If industry j increases output, then the demands from industry j on the industries whose products are used as intermediate inputs will increase. The backward linkage measure indicates the interrelation of a particular industry with other industries from which intermediate inputs are purchased. The increased output in industry j suggests increased supplies from industry j to the other industries that use commodity j in their production. Meanwhile, the forward linkage measure indicates the interrelation of a particular industry to other industries to which outputs are sold. Backward and forward linkages are derived from the *Leontief inverse matrix* and the *Ghosh inverse matrix*, respectively. The direct backward ($\mathcal{B}$) and direct forward ($\mathcal{F}$) linkages are defined by the column-sum of *Leontief inverse matrix* and row-sum of *Ghosh inverse matrix*, respectively, i.e.:

$$\mathcal{B}_j = \sum_{i'=1}^{n} b_{i'j}, \qquad \mathcal{F}_i = \sum_{j'=1}^{n} \bar{b}_{ij'},$$

where b_{ij} and $\bar{b}_{ij}$ are the ijth element of *Leontief inverse matrix* B, and *Ghosh inverse matrix* $\bar{B}$, respectively.

Even if the backward and forward linkages are appropriate for measuring the economic influences of expansion and shrinkage of industries as a whole, they do not provide consistent measures for linkages between two industries. Moreover, a large backward linkage does not guarantee a large forward linkage and vice versa. Consider the hypothetical supply flows of intermediate inputs illustrated in Figure 2. Sector A supplies for itself, and thus, its expansion cannot induce output growth in the other sectors in terms of forward linkage. However, Sector A can influence the other sectors in terms of backward linkage because it uses products from all sectors. Accordingly, Sonis and Hewings (1999) proposed an index called the multiplier product matrix (MPM) to connect the properties of backward and forward linkages as

$$MPM = \frac{1}{V} \mathcal{F} \cdot \mathcal{B}, \tag{13}$$

where

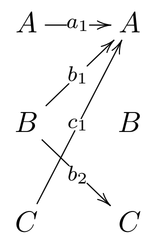


Figure 2: Hypothetical supply flows of intermediate inputs

$$V = \sum_{j=1}^{n} \mathcal{B}_j = \sum_{i=1}^{n} \mathcal{F}_i. \tag{14}$$

MPM provides a measure of the relationships among industries, thus allowing them to be organized into a rank-sized hierarchy. MPM measures the impacts of a certain industry on the other industries. However, MPM is not symmetric because the forward linkage of a particular industry mostly differs from its backward linkage (e.g., Figure 4). To be utilized as a proxy for economic distance, we modify MPM to be symmetric by taking the Euclidean norm of its elements of the MPM, such that

$$m_{ij}^E = m_{ji}^E = \sqrt{m_{ij}^2 + m_{ji}^2}. \tag{15}$$

3.2 Economic distance and weights matrix

We utilize the MPM defined in the previous section to create a spatial weights matrix based on economic distance. Geographic distance is not a meaningful spatial concept that can link the supply chains embedded in the sectoral flows among industries. At first glance, one may perceive the MPM values as useful in filling the weighting elements of the matrix. However, MPM values have opposite interpretations as those of the weights matrix based on typical geographical distance. Moreover, the differences in distances may not be significant across industries. Thus, to address these issues, we define an economic distance metric that is analogous to physical distance and introduce a distance-decay function to give more weights on closely related industries.

We first define a measure of economic distance between industry i and j as

$$d_{ij} \equiv \max_{i'} m_{i'j}^E - m_{ij}^E, \tag{16}$$

where m_{ij}^E is the element of MPM, as discussed in the previous section[4]. Once we obtain economic distance, we can assign weights according to distance. Several schemes are used to assign weights, such as contiguous neighbors, inverse distances, lengths of shared borders divided by perimeter,

[4]The diagonals of MPM are non-zero and in fact are mostly the largest element in their columns (or rows). Hence we need to set the diagonals to zero afterward to satisfy the regularity assumption A1.

bandwidth, centroid distance, and k-nearest neighbors. The widely used weighting scheme for spatial weights matrices should exclude observations that are beyond threshold distance d^*, that is,

$$w_{ij} = \begin{cases} 1 & \text{if } d_{ij} < d^*; \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

However, the spatial weighting function presented in Eq. (17) suffers from discontinuity. One way to circumvent this problem is to assume a continuous function denoted by d_{ij}. For continuous weighting schemes, the application of a distance-decay function or a decline function allows more weights to be designated to closely associated units than to the areas far from one another (Brunsdon et al., 1996; McMillen, 2003). A negative exponential function is suggested by Brunsdon et al. (1996) as follows:

$$w_{ij} = e^{-\eta d_{ij}^2}, \tag{18}$$

where η is the spatial scale parameter that determines the degree of distance-decay. The larger the value of η, the more abrupt is the cut-off of influence of distant economic units. The degree of distance-decay varies between different concepts of distance, groups of economic units, and estimation approaches.

We suggest Eq. (18) as the spatial weights elements to reflect input–output relationship among industries. The spatial weights matrix should be appropriately modified to follow the standard assumptions in the literature as follows:

Assumption A1 *W values are row-normalized non-stochastic spatial weights matrices with zero diagonals.*

Assumption A2 *$I_N - \rho W$ is invertible for all $\rho \in \Lambda$, where the parameter space Λ is compact and ρ is in the interior of Λ.*

Assumption A3 *W values are uniformly bound in both row and column sums in absolute value. In addition, $(I_N - \rho W)^{-1}$ values are uniformly bound in $\rho \in \Lambda$.*

4 Estimation

The estimation of Eqs. (6) and (8) involves a nonlinear two-step procedure: Step 1) estimation of scale parameter η to obtain the optimal weights matrix, and Step 2) estimation of the model parameters. The estimation of the exponent parameter η is crucial because it determines the degree of interaction. For the second step, we estimate model parameters with QMLE. We also estimate the relative efficiency scores via the approach suggested by Schmidt and Sickles (1984) and Cornwell et al. (1990).

Production functions are typically estimated by using various parametric, nonparametric, and semi-parametric techniques. A standard approach to production function estimation is to adhere to the average production technology instead of the best-practice technology, which is accomplished in the stochastic frontier literature by neglecting the assumption that all producers are cost or profit efficient. Minimal differences, if any differences exist at all, usually appear in the estimates of the basic production model parameters, such as in output elasticities, among others. However, the stochastic frontier analysis (SFA) approach can decompose the Solow-type residual into two components that utilize either the TFP and its change over time or the TFP growth. The identification of the decomposition of TFP growth into separate efficiency and technical change components is based on the assumption that the average production function represents the maximum level of output given the levels of inputs on the average. Shifts in this average level of productivity over time, which are usually represented as a common trend by using either a time variable or a time index, indicates technical change. Inefficiency is interpreted as the productivity of a unit at a specific time period relative to the average best-practice production frontier, and it typically includes a one-sided term (negative) that represents the short-fall in a firm's average production relative to a benchmark set by the most efficient firm. One-sided distributions, such as half-normal, truncated normal, exponential, or gamma distribution, are often used in parametric models. Schmidt and Sickles (1984) and Cornwell et al. (1990) suggested the avoidance of strong distributional assumptions by utilizing the structure of a panel production frontier. Schmidt and Sickles (1984) assumed inefficiency to be time-invariant and unit-specific, while Cornwell et al. (1990) relaxed the time-invariant assumption by introducing a flexibly parametrized function of time, thereby replacing individual fixed effects. In the present study, we follow the work of Cornwell et al. (1990) for estimation. This method allows us to estimate time-varying efficiency without requiring further distributional assumptions on the one-sided efficiency term.

According to Cornwell et al. (1990), the non-spatial model (Eq. (4)) can be estimated via three techniques: within transformation, generalized least squares, and efficient instrumental variable approach. However, the extended models (Eqs.(5) and (7)) have several difficulties in estimation because they include additional spatially correlated variables. A quasi-maximum likelihood estimation (QMLE) is used in our analysis. QMLE can provide robust standard errors against misspecification of the error distributions (Yang, 2013). QMLE enables us to minimize the number of parameters to be estimated via the concentrated likelihood function instead of using the full likelihood function. We typically substitute the closed-form solutions of a set of parameters into the likelihood function, and the resultant concentrated likelihood function becomes a function of spatial coefficient parameters only. The optimization with the concentrated likelihood is known to give the same maximum likelihood estimates after maximizing the full likelihood (LeSage and Pace, 2009). We will outline the estimation procedure briefly. The details are presented in Appendix A.

We first find η by minimizing the mean-squared error as

$$\hat{\eta} = \underset{\eta \in \mathcal{E}}{\operatorname{argmin}} \frac{1}{N} \sum_{i} \left[y_i - y^*_{\neq i}(\eta) \right]^2 ,$$

where $\mathcal{E} = \{\eta \in \mathbb{R} : \eta > 0\}$, and $y^*_{\neq i}(\eta)$ is the fitted value of y_i with the observations for industry i omitted from the estimation using a distance-decay of η. We can find closed-form solutions for the parameters, except for the spatial autoregressive parameter ρ, by using the first-order conditions of the likelihood functions of Eqs. (5) and (7). The spatial parameters of λ are the coefficients of the spatially weighted independent variables. We treat the spatially weighted independent variables as additional regressors. The substitution of the closed-form solutions into the likelihood functions gives the concentrated likelihood functions with ρ as the only unknown variable. However, $\hat{\rho}$ can be obtained by maximizing the concentrated likelihood functions. Hence, all other parameters can be found by using $\hat{\rho}$. The details of the derivation of the asymptotic distribution of the estimated parameters are presented in Appendix A.

Once we obtain the estimates of the parameters β, ρ, δ_i, and σ_v^2, we can recursively solve for an estimate of α_{it}, although we cannot separately identify δ_0 and u_i. By using the estimate of α_{it}, we can obtain the relative inefficiency measure following the studies of Schmidt and Sickles (1984) and Cornwell et al. (1990). In particular, from Eq. (3), we know the estimate of α_{it} is

$$\hat{\alpha}_{it} = R_t'\hat{\delta}_i.$$

The estimates of the frontier intercept α_t and the time-dependent relative inefficiency measure u_{it} can be derived as[5]:

$$\hat{\alpha}_t = \max_j(\hat{\alpha}_{jt}), \tag{19}$$

$$\hat{u}_{it} = \hat{\alpha}_t - \hat{\alpha}_{it}. \tag{20}$$

5 Empirical applications: industry-level productivity of U.S.

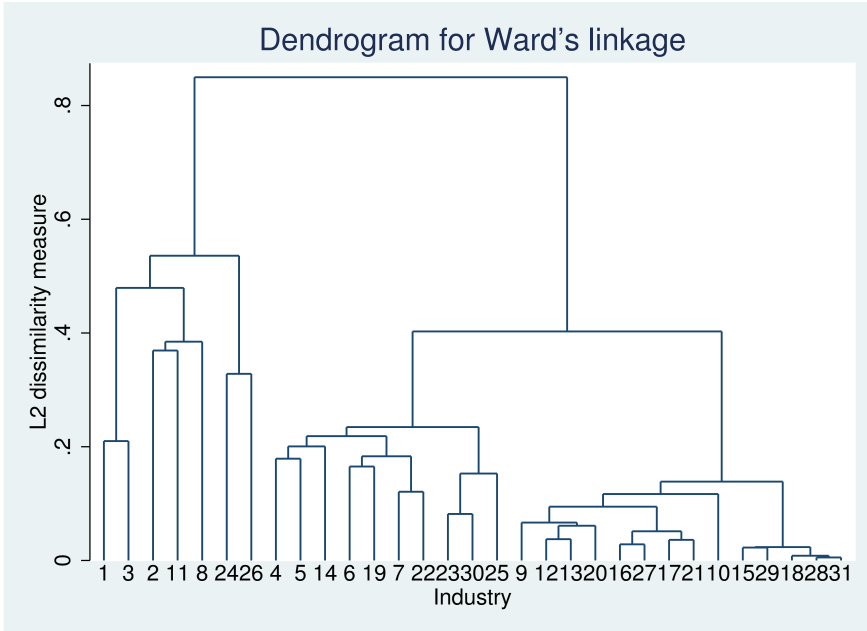
We examine the industry-level productivity of the U.S. from 1947 to 2010 by using the classical Cobb–Douglas production technique.

5.1 Data

The databases employed for the empirical application have been sourced from the U.S. accounts of World KLEMS and World Input–Output Database [6]. The usual difficulties of using datasets from different sources include maintaining the consistency of data formation, such as industry classification levels or units. Maintaining the conformability of industry classification and the target period across the datasets is essential because the U.S. national accounting system, which provides raw data, has shifted to the North American Industry Classification System (NAICS) in

[5] Hence, the relative efficiency score can be written as $\widehat{EFF}_{it} = e^{-\hat{u}_{it}}$.

[6] The databases are open-sources on the following web pages: `http://www.worldklems.net/`, and `http://www.wiod.org/`.



Note: The industries represented by the numbers should refer to the Table C.1.

Figure 3: Cluster analysis on intermediate input to industry output ratio

2003[7]. The World KLEMS dataset covers data from 1947 to 2010 and uses NAICS. The WIOD is also based on NAICS, and it covers data from 1995 to 2011. Both databases are mapped into the 31 industry classifications of the International Standard Industrial Classification (ISIC) for "All Economic Activities" (Rev. 3) for easy comparison[8].

Our study aims to examine industry-level productivity growth in the context of interdependent spatial linkages among industries. Hence, we narrow our focus to a relatively small sample to show the strong similarities in their production processes. We wish to minimize the extent to which the unobservable differences in the production technology are interpreted as either spillovers or inefficiency. We acknowledge that this decision may alter to some extent the potential linkages among the remaining sectors. However, we feel that the tradeoff is warranted. We select the sample of sectors with hierarchical cluster analysis based on sector-specific average products of inputs. Figure 3 illustrates a dendrogram of clusters based on Ward's minimum variance method (Ward, J. H., Jr., 1963). In the dendrogram, the seven industries on the left are less similar to the other industries. Moreover, the dissimilarity level is high among the seven industries. Hence, we exclude these items from the analyzed sample. The industry classifications used in this study are listed in Table 2[9].

One of the advantages of the KLEMS data is that the variables are quality adjusted via the

[7]In addition, the national accounts systems of each country are based on different international classification, and keeping concordance with each other is difficult.

[8]However, the two databases have slightly different industry definitions. That is, *Textiles, Textile, Leather and Footwear* and *Transport and Storage* industries are finely defined in World Input–Output Database, which requires us to aggregate the finer industries to meet the number of industries of data-sets.

[9]The complete industry classifications and the excluded industries can be found in Appendix C.

Table 2: Industry classifications and codes

No.	Industry	ISIC Rev. 3
	TOTAL MANUFACTURING	
1	Textiles, Textile, Leather and Footwear	17t19
2	Wood and Products of Wood and Cork	20
3	Pulp, Paper: Paper: Printing and Publishing	21t22
4	Coke, Refined Petroleum and Nuclear Fuel	23
5	Rubber and Plastics	25
6	Other Non-Metallic Mineral	26
7	Machinery, Nec	29
8	Electrical and Optical Equipment	30t33
9	Transport Equipment	34t35
10	Manufacturing, Nec; Recycling	36t37
11	ELECTRICITY, GAS AND WATER SUPPLY	E
12	CONSTRUCTION	F
	WHOLESALE AND RETAIL TRADE	G
13	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	50
14	Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	51
15	Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	52
16	HOTELS AND RESTAURANTS	H
	TRANSPORT AND STORAGE AND COMMUNICATION	I
17	Transport and Storage	60t63
18	Post and Telecommunications	64
	REAL ESTATE, RENTING AND BUSINESS ACTIVITIES	K
19	Real Estate Activities	70
20	PUBLIC ADMIN AND DEFENCE; COMPULSORY SOCIAL SECURITY	L
21	EDUCATION	M
22	HEALTH AND SOCIAL WORK	N
23	OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES	O

Törnqvist index approach. In data production, the variables often reflect heterogeneity due to quality differences in the components of the aggregated quantity indices. Coelli et al. (2005), among other researchers, summarized several possible options to incorporate the variation in the quality of goods and services[10]. Fortunately, the KLEMS dataset reflects the quality differences in the variables. Labor force heterogeneity has been addressed by using the approach of Jorgenson et al. (1987). Aggregate labor services are assumed to be a translog function of their individual components, which are based on the market equilibrium conditions that equate the supply of each type of factor input to the sum of demands for those inputs by all sectors. This assumption is connected to the index–number approach of Diewert (1978). We collect the output measures (gross output and value added) and input measures (capital and labor services; intermediate inputs) formed with the Törnqvist index from the KLEMS database. The summary statistics of the real variables and volume indices are listed in Table 3. Between-industry standard deviations are larger

[10] i) quality-augmented measures, ii) using some numerical weights to goods and services of different qualities, iii) two-stage approach: obtain productivity or efficiency measure from unadjusted measures and regress them by using various quality measures. iv) the model of Battese and Coelli (1995).

Table 3: Summary Statistics

	Variable		Mean	Std. Dev.	Min	Max	Observations
Real Variables	GO	overall	3,662.64	3,732.45	56.28	23,976.29	$N = 1,472$
		between		2,865.96	712.41	11,156.59	$n = 23$
		within		2,463.59	-4,802.55	16,482.34	$T = 64$
	VA	overall	2,190.05	2,674.45	3.09	15,683.36	$N = 1,472$
		between		2,301.49	299.35	8,290.85	$n = 23$
		within		1,443.16	-3,712.55	9,921.11	$T = 64$
	CAP	overall	794.75	1,768.99	6.12	14,102.48	$N = 1,472$
		between		1,602.41	60.31	8,001.58	$n = 23$
		within		819.51	-4,524.85	6,895.65	$T = 64$
	LAB	overall	1,597.84	1,654.70	117.55	12,030.32	$N = 1,472$
		between		1,456.21	194.82	6,138.22	$n = 23$
		within		841.60	-2,772.77	7,489.94	$T = 64$
	II	overall	1,556.97	1,359.73	46.04	9,085.89	$N = 1,472$
		between		839.64	401.78	3,309.06	$n = 23$
		within		1,083.54	-1,309.28	7,416.73	$T = 64$
Index Variables	GO_{QI}	overall	57.51	31.85	1.16	152.14	$N = 1,472$
		between		16.17	30.01	109.50	$n = 23$
		within		27.64	-1.98	142.85	$T = 64$
	VA_{QI}	overall	59.21	34.10	0.14	189.65	$N = 1,472$
		between		19.22	20.97	99.31	$n = 23$
		within		28.45	-6.99	227.89	$T = 64$
	CAP_{QI}	overall	50.76	33.54	2.10	134.97	$N = 1,472$
		between		13.09	35.22	81.50	$n = 23$
		within		31.00	-3.14	130.40	$T = 64$
	LAB_{QI}	overall	87.54	42.09	15.95	282.09	$N = 1,472$
		between		35.31	53.24	213.66	$n = 23$
		within		24.05	-56.07	155.97	$T = 64$
	II_{QI}	overall	57.05	35.53	2.24	214.43	$N = 1,472$
		between		20.61	32.61	134.53	$n = 23$
		within		29.26	-35.51	198.73	$T = 64$

than the within-industry standard deviations in all real variables except for intermediate inputs. However, the within-industry standard deviations are larger than the between-industry standard deviations in all index variables except for labor services. This finding implies that the quality-adjusted measures have somewhat different distributional shapes relative to the quality-unadjusted measures.

We use the input–output table to exploit the patterns of intermediate inputs flows and create spatial weights matrices. Figure 5 shows the linkages between industries. We find that industries with strong backward and forward linkages have large coefficients for MPM. By contrast, industries with weak backward and forward linkages, such as *retail trade, excluding motor vehicles and motorcycles, repair of household goods*(Ind15), and *real estate activities*(Ind19) have small MPM coefficients. *Public admin and defense, compulsory social security, and health* and *social work* have highly weak linkages with other industries due to their almost complete lack of forward linkage.

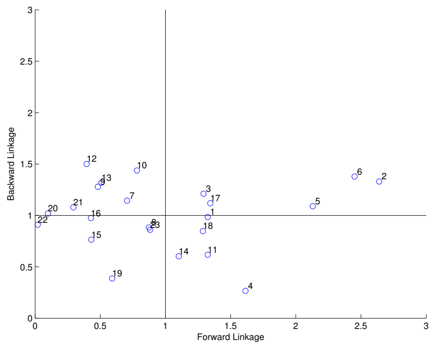


Figure 4: Forward and Backward linkage

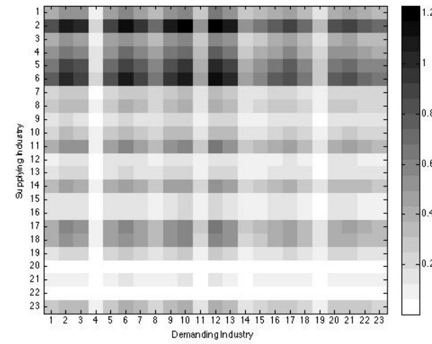


Figure 5: Multiplier product matrix

5.2 Empirical findings

We assume a quadratic functional form for α_{it} following the work of Cornwell et al. (1990), i.e, $\alpha_{it} = \theta_{i1}+\theta_{i2}t+\theta_{i3}t^2$, to model sector-specific efficiency change and a time dummy and its interactive time dummy to control for different innovations in different industries. Additionally, the quadratic term of time in the efficiency term can help capture any nonlinearity in efficiency changes during the relatively lengthy sample period. To control for possible endogeneity problem between the input factor levels and productivity, we also use one-year lagged variables as instruments for the factor inputs as well as various control function approaches. These had minimal impact on the results and are available upon request.

5.2.1 Output elasticity and returns to scale

The estimation results are presented in Table 4. The dependent variable is the gross-output index[11]. The first two pairs of columns in Table 4 show the results for CSSW and CSSG without spatial specifications. All coefficient estimates for factor inputs are statistically significant. The coefficients can be interpreted as output elasticities in the models. The input with the largest coefficient(elasticity) is labor, whereas the input with the smallest coefficient is capital. Returns to scale are nearly 0.87 for both models. We can estimate the parameters for *intercept*, $Time$, and $Time^2$ in the CSSG model. The intercept term is negative but insignificantly different from zero. $Time$ is approximately 0.014, but $Time^2$ has no effects. Thus, the growth rate of the economy is 1.4% on the average.

The succeeding four pairs of columns show the results under spatial specifications. The coefficients estimates no longer represent output elasticities when spatial dependent terms are added. The output elasticity of a factor input for a given time t involves pre-multiplication of the inverse matrix, $(I_N - \rho W)^{-1}$, on the coefficient estimates β. Hence, the output elasticities for SAR and SDM take the following forms:

[11] The estimation results using the value-added as a dependent variable are given in Appendix B.

Table 4: CSS vs SARCSS vs SDMCSS

	Non-spatial				SAR				SDM			
	CSSW		CSSG		CSSW		CSSG		CSSW		CSSG	
	Coef.	std.err	Coef.	std.err	Coef.	std.err	Coef.	std.err	Coef.	std.err	Coef.	std.err
$log(K)$	0.144	0.024	0.146	0.022	0.102	0.022	0.107	0.021	0.089	0.024	0.096	0.023
$log(L)$	0.454	0.025	0.446	0.024	0.389	0.024	0.388	0.023	0.407	0.025	0.405	0.024
$log(I)$	0.270	0.014	0.278	0.014	0.255	0.014	0.261	0.013	0.260	0.014	0.267	0.013
$Intercept$	-	-	-0.369	0.125	-	-	-0.067	0.131	-	-	-0.052	0.142
$Time$	-	-	0.014	0.002	-	-	0.001	0.003	-	-	0.001	0.003
$Time^2$	-	-	0.000	0.000	-	-	0.000	0.000	-	-	0.000	0.000
σ_v^2	0.005	0.000	0.005	0.000	0.004	0.000	0.004	0.000	0.004	0.000	0.004	0.000
Spatial Parameters												
$W \cdot log(Y)(\rho)$	-	-	-	-	0.384	0.031	0.374	0.024	0.492	0.049	0.492	0.044
$W \cdot log(K)(\lambda_1)$	-	-	-	-	-	-	-	-	0.053	0.054	0.047	0.035
$W \cdot log(L)(\lambda_2)$	-	-	-	-	-	-	-	-	-0.131	0.040	-0.131	0.053
$W \cdot log(I)(\lambda_3)$	-	-	-	-	-	-	-	-	-0.092	0.044	-0.097	0.040
η	-	-	-	-	1.496		1.505		1.259		1.158	
Elasticities												
θ_K	0.144	0.024	0.146	0.022	0.168	0.037	0.171	0.023	0.281	0.090	0.279	0.046
θ_L	0.454	0.025	0.446	0.024	0.635	0.044	0.623	0.040	0.545	0.089	0.539	0.089
θ_I	0.270	0.014	0.278	0.014	0.415	0.030	0.418	0.024	0.331	0.070	0.337	0.069
R^2	0.675		0.681		0.719		0.722		0.713		0.717	
$AdjustedR^2$	0.658		0.664		0.704		0.707		0.697		0.701	
$loglikelihood$	-		-		1963.634		1929.299		1971.500		1937.191	
RTS	0.867		0.870		1.219		1.211		1.158		1.155	

$$\text{(SAR)} \qquad \frac{\partial y_t}{\partial X_{k,t}} = (I_N - \rho W)^{-1}(\beta_k I_N), \tag{21}$$

$$\text{(SDM)} \qquad \frac{\partial y_t}{\partial X_{k,t}} = (I_N - \rho W)^{-1}(\beta_k I_N + \lambda_k W). \tag{22}$$

The results are based on a time invariant spatial weights matrix. Given that the output elasticities are given via the $(N \times N)$ matrix, the averages are taken along the diagonals to obtain the mean direct effects and along the column (or row) sums to obtain the mean indirect effects as proposed by LeSage and Pace (2009). The sum of the mean direct and indirect effects can then be defined as the average total effect of each factor input on output. The total effects can be interpreted as average total output elasticities of the factor inputs. Hence, the total effects are calculated along with the regression results. The corresponding distribution of total effects should be obtained separately because the significance of β estimates does not guarantee the significance of output elasticities of SAR and SDM, as given by Eqs. (21) and (22), and also because they are functions of the estimates of ρ and λ. As such, we follow the algorithms suggested by LeSage and Pace (2009). This algorithm involves drawing parameter estimates D times based on their estimated covariance structure and computing the mean and standard deviation of the direct, indirect, and total effects.

All estimates of parameters and elasticities for SAR and SDM are statistically significant at the 1% level except for *intercept*. The spatial scale parameter η is estimated to be in the range of 1.158 to 1.505. The coefficient of the spatially lagged dependent variable, ρ, is estimated to be in

Table 5: Direct, Indirect, and Total Elasticity

		Direct		Indirect		Total	
		Elasticity	asy. t-stat	Elasticity	asy. t-stat	Elasticity	asy. t-stat
SARCSSW	Capital	0.104***	4.617	0.064***	4.113	0.168***	4.594
	Labor	0.393***	16.570	0.242***	7.597	0.635***	14.278
	Intermediate	0.257***	18.244	0.158***	7.238	0.415***	13.874
SARCSSG	Capital	0.108***	6.168	0.063***	9.142	0.171***	7.319
	Labor	0.393***	16.980	0.230***	9.242	0.623***	15.686
	Intermediate	0.264***	20.269	0.154***	9.404	0.418***	17.501
SDMCSSW	Capital	0.095***	3.975	0.187**	2.102	0.281***	3.143
	Labor	0.409***	15.820	0.136	1.563	0.545***	6.110
	Intermediate	0.262***	18.698	0.069	1.011	0.331***	4.724
SDMCSSG	Capital	0.100***	4.433	0.179***	3.649	0.279***	6.097
	Labor	0.408***	17.095	0.131	1.503	0.539***	6.038
	Intermediate	0.269***	19.883	0.068	1.034	0.337***	4.899

Note: *, **, *** denote that we reject the null hypotheses of constant returns to scale at the 5%, 1%, and 0.1% levels, respectively.

the range of 0.374 to 0.492. For the coefficients of spatially weighted independent variables, λ_1 is estimated to be positive, whereas λ_2 and λ_3 are negative.

The output elasticities of each factor inputs are calculated as θ_K, θ_L, and θ_I, which are the total effects of the factor inputs. Meanwhile, the coefficient estimates in the first three rows represent the direct effects. In comparing the results of SARCSS and SDMCSS with the results of non-spatial CSS, we found that all output elasticities are estimated larger under spatial specifications. In particular, SARCSS estimates the largest elasticities out of the three specifications. In comparing SARCSS with SDMCSS, we also found that the output elasticity of capital is larger. However, the output elasticities of labor and intermediate inputs are estimated to be smaller with SDM. This discrepancy leads to relatively small returns to scale in SDMCSS. We can compare the goodness-of-fit of SARCSS and SDMCSS by using the likelihood ratio test, given that SAR is nested in SDM. The LR test statistics are 15.732 and 15.784 for within-estimation and GLS-estimation, respectively. Thus, adding the spatially weighted independent variables results in a statistically significant improvement in model fit.

One advantage of spatial analysis is that we can estimate separately the direct and indirect effects of regressors on a dependent variable. Table 5 shows the direct, indirect, and total output elasticities of each factor input. Most effects are statistically significant, except for the indirect effects of intermediate inputs under SDM. The sum of the direct effects is between 0.76 and 0.78, which implies possible decreasing scale economies. However, the returns to scale are increased to 1.2 when the indirect effects are considered. The indirect effects are estimated to be approximately 37% of the total effects when only a spatially lagged dependent variable is included in the model. If we consider spatially lagged terms for dependent and independent variables, then the portions of the indirect effects will vary. Interestingly, the indirect effects of capital service have increased to 67%, whereas the indirect effects of labor and intermediate inputs have decreased to 25% and

21%, respectively.

Glass et al. (2015) defined internal, external, and total returns to scale with the direct, indirect, and total effects from a spatial production function of European countries to determine if the returns to scale measures show constant returns to scale. In our analysis, the internal and external returns to scale by themselves do not show evidence of constant returns to scale. However, they are large enough to make the total returns to scale to become increasing returns to scale (SAR) or constant returns to scale (SDM), as shown in Table 6.

Table 6: Internal, External, and Total Returns to Scale

Model	Internal RTS	asy. t-stat	External RTS	asy. t-stat	Total RTS	asy. t-stat
SARCSSW	0.754***	−7.576	0.464***	−11.239	1.219***	3.396
SARCSSG	0.765***	−7.919	0.446***	−16.804	1.211***	4.822
SDMCSSW	0.766***	−6.706	0.392***	−4.796	1.158	1.260
SDMCSSG	0.777***	−6.858	0.378***	−6.431	1.155	1.577

Note: *, **, *** denote that we reject the null hypotheses of constant returns to scale at the 5%, 1%, and 0.1% levels, respectively.

5.2.2 Efficiency analysis

The average efficiency scores of each industry are shown in Table 7. The total average efficiency scores are approximately 0.70 for gross output productivity. The efficiency ranking only varies marginally even if the coefficient estimates in Table4 and the efficiency scores differ across the models. The *construction* sector (Ind12) is found to be the most efficient, whereas the *electrical and optical equipment* sector (Ind8) is the least efficient industry on the average for all models. This result is inconsistent with expectations because the *electrical and optical equipment* industry is one of the driving forces of economic growth in the total manufacturing sectors of many countries. To check if the relative efficiency scores are reasonable, we compare these scores with the TFP measure in the KLEMS database, which is computed on the basis of the growth accounting approach. TFP is also computed industry-by-industry by using the quality-adjusted Törnqvist index variables. Hence, the measure actually does not consider interactions among industries. In the last two columns of Table 7, we present the period average of industry in terms of TFPs. We observe that *construction* has the largest average TFP, and *electrical and optical equipment* has the smallest, which is consistent with our results[12].

Jorgenson et al. (2012) highlighted the performance of industries from the perspective of innovation, which is regarded the engine for long-run economic growth. They note the important role of IT-producing industries, including software and hardware manufacturing industries, and IT-service-producing industries. In particular, they found that these industries have substantial

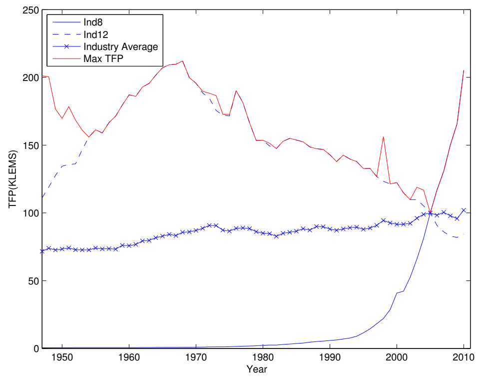
[12] Note that TFPs are calculated using TFP growth rates obtained from the growth accounting approach, which set the TFP level of the reference year to be 1. The growth accounting approach applies the methodology to a dataset from each industry independently. Hence, the comparison of TFP levels across industries is not provided. Because the approach sets the TFP level of the reference year(2005) as one, we may have the period average TFP that is greater than one depending on the fluctuations in TFP levels.

Table 7: Efficiency Scores

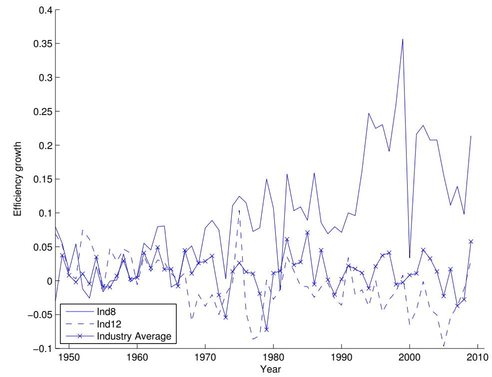
	non-spatial CSS				SAR				SDM				KLEMS	
	Within		GLS		Within		GLS		Within		GLS		Growth Accounting	
	Eff. Score	Rank	Eff. Score	Rank	Eff. Score	Rank	Eff. Score	Rank	Eff. Score	Rank	Eff. Score	Rank	Avg. TFP	Rank
Ind01	0.664	14	0.664	15	0.755	11	0.748	11	0.736	11	0.731	11	0.491	21
Ind02	0.789	8	0.793	7	0.801	7	0.803	7	0.787	7	0.790	7	1.098	7
Ind03	0.855	4	0.858	4	0.870	4	0.871	3	0.853	4	0.856	4	1.234	3
Ind04	0.596	19	0.597	19	0.637	18	0.635	18	0.628	18	0.626	18	0.342	22
Ind05	0.636	17	0.638	17	0.639	17	0.640	17	0.638	16	0.639	16	0.778	14
Ind06	0.692	12	0.693	12	0.736	12	0.734	12	0.730	12	0.728	12	0.781	13
Ind07	0.693	11	0.697	11	0.715	13	0.717	13	0.705	13	0.707	13	1.046	8
Ind08	0.293	23	0.294	23	0.293	23	0.294	23	0.290	23	0.291	23	0.211	23
Ind09	0.662	15	0.665	14	0.681	15	0.682	15	0.672	15	0.673	15	0.812	12
Ind10	0.583	20	0.585	20	0.598	20	0.598	20	0.592	20	0.592	20	0.545	19
Ind11	0.842	5	0.843	5	0.881	2	0.878	2	0.876	3	0.873	3	1.357	2
Ind12	0.973	1	0.973	1	0.974	1	0.974	1	0.972	1	0.973	1	1.503	1
Ind13	0.651	16	0.651	16	0.646	16	0.647	16	0.636	17	0.637	17	0.629	17
Ind14	0.501	21	0.503	21	0.488	22	0.490	22	0.485	22	0.487	22	0.532	20
Ind15	0.619	18	0.621	18	0.609	19	0.611	19	0.602	19	0.605	19	0.63	16
Ind16	0.792	6	0.792	8	0.781	9	0.782	9	0.786	8	0.786	9	1.126	6
Ind17	0.675	13	0.674	13	0.691	14	0.690	14	0.691	14	0.690	14	0.714	15
Ind18	0.496	22	0.499	22	0.491	21	0.494	21	0.486	21	0.489	21	0.591	18
Ind19	0.783	9	0.787	9	0.765	10	0.768	10	0.775	10	0.777	10	0.892	11
Ind20	0.759	10	0.763	10	0.814	6	0.813	6	0.810	6	0.809	6	0.951	10
Ind21	0.791	7	0.794	6	0.781	8	0.784	8	0.785	9	0.787	8	1.002	9
Ind22	0.864	3	0.868	3	0.839	5	0.843	5	0.845	5	0.848	5	1.205	5
Ind23	0.884	2	0.883	2	0.871	3	0.870	4	0.878	2	0.877	2	1.223	4
Average	0.700		0.702		0.711		0.712		0.707		0.707		0.856	

contribution to economic growth during the investment boom of 1995 to 2000. The sampling period of 1947 to 2010 is long enough to see changes in the relative importance of different industries in the growth process. Thus, we examine the most and least efficient industries further by scrutinizing the variation of the relevant productivity and efficiency measures in Figure 6. We compare the TFP values, which are computed with the growth accounting approach and the efficiency scores estimated by the spatial stochastic frontier approach. We also create a relative measure of TFP defined as the relative size to the largest TFP of each year. The left panels of Figure 6 illustrate the evolution of the productivity measure of *electrical and optical equipment* (Ind8), *construction* (Ind12), and the industry average. The right panels present the growth rate of productivity or efficiency. Panels (a) and (b) are drawn from KLEMS data, whereas Panels (c) and (d) are generated from TFP for comparison. The solid red line of Panel (a) represents the maximum TFP of each year. The red line is almost identical to the solid blue line, which is a TFP series of the *construction* sector. Even if we observe that the TFP of the *construction* industry falls after the late 1960s, it is the most efficient in terms of TFP. By contrast, the productivity of the *electrical and optical equipment* industry soars sharply after the mid-1990s when the so-called IT investment boom has started. By normalizing the TFP measure relative to the most productive industry, we can obtain Panels (c) and (d), which can be compared with the relative efficiency scores from the spatial stochastic frontier approach. In comparing Panels (c) and (e), the relative efficiency behaves somewhat similarly to the relatively smooth lines by using the spatial stochastic frontier approach. The variations in efficiency changes are averaged in the regression-based approach. The fluctuations in efficiency are not captured. However, the comparison with the TFP measure from the growth accounting approach, which does not explicitly consider randomness, provides a benchmark that allows us to confirm if our methodology not only provides an explicit role for spatial supply chain

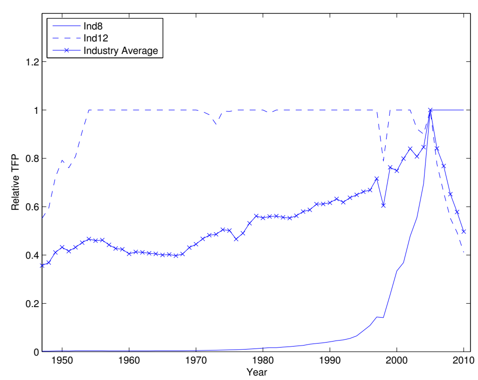
Figure 6: Efficiency and Efficiency Growth



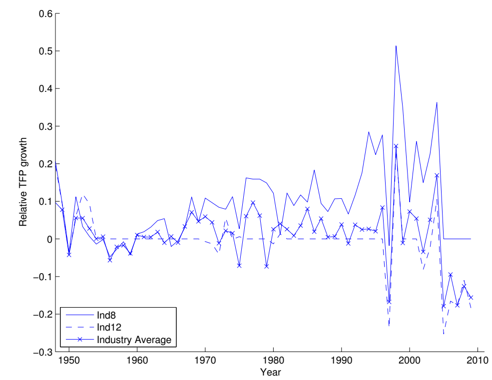
(a) TFP (KLEMS)



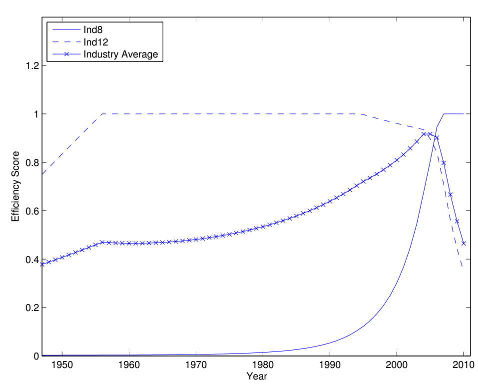
(b) TFP Growth



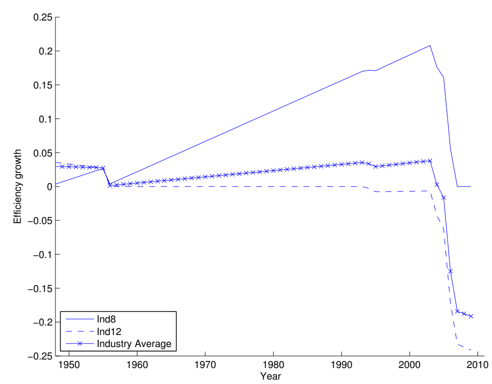
(c) Relative TFP



(d) Relative TFP Growth



(e) Relative Efficiency Score



(f) Relative Efficiency Growth

linkages but also provides summary measures of productivity growth that are consistent with those generated using established approaches of statistical agencies in the U.S. and in the other developed and developing countries.

We then examine how the industry-average efficiencies evolve over time (Figure 7). Two local

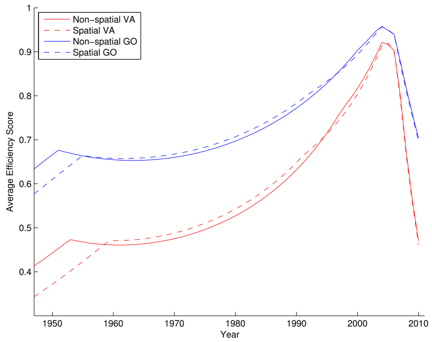


Figure 7: Efficiency Scores (Industry Average)

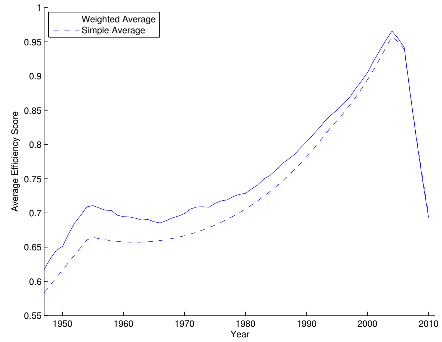
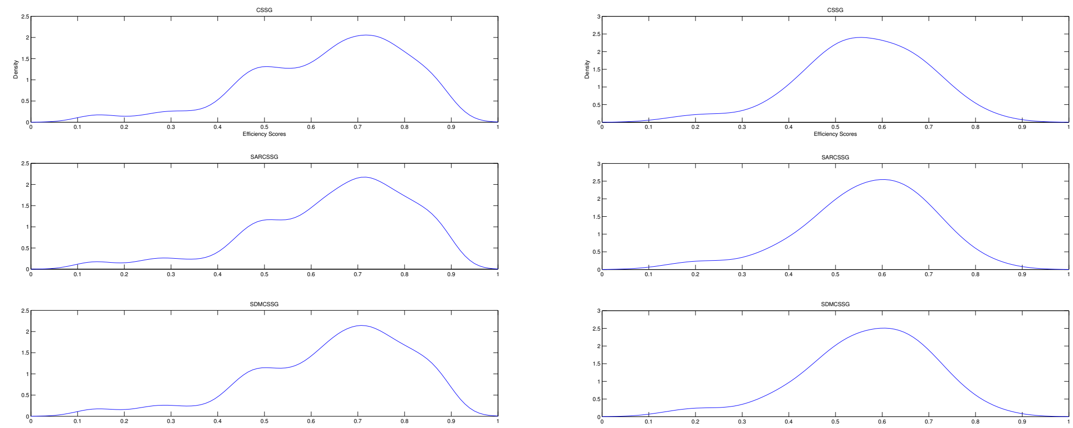


Figure 8: Weighted average of efficiency scores

peaks are observed, and the behavior of the average efficiency scores moves similarly as those in the spatial and non-spatial settings. In particular, the global peak of the efficiency scores is observed in 2004. However, we discern that the size and significance of the small peaks in the 1950s depend on the models, i.e., the adoption of spatial specification and value-added productivity measure shifts the peak to the right relative to their counterparts. We interpret that the U.S. industry-average relative efficiency has converged to almost full efficiency in 2004 in terms of value-added and gross-output measures. However, the average relative efficiency score plummeted after the global peak. The relative efficiencies of all industries, except the *electrical and optical equipment* sector, fell after 2004. The implication of this finding is in line with the importance of IT-producing and IT-using industries, as stressed by Jorgenson et al. (2012), who defined the contribution of each industry to productivity as the productivity growth rate of each industry weighted by the ratio of the industry's output to aggregate output. Despite the high productivity growth of *electrical and optical equipment*, the industry cannot drive the productivity growth of the total economy, as its output ratio is only approximately 4% of the total economy. Instead of the simple arithmetic average of efficiency scores, we draw the weighted average (Figure 8), thereby following the weighting scheme used by Jorgenson et al. (2012). The solid line of Figure 8 shows the weighted average. The weighted average efficiency score is larger than the simple average (dotted line). However, the differences decrease over time, and the two averages are essentially the same after efficiency begins to decline.

We also check the empirical density of the estimated efficiency scores. Figures 9a and 9b compare the kernel density plots of the efficiency scores. The kernel density estimates in Figure 9a, which are drawn from the estimates of uit, appear to coincide with the typically assumed one-sided inefficiency distribution, such as Half-normal, Truncated normal, Exponential, or Gamma distributions. Meanwhile, the kernel densities of period-average efficiency scores for each industry in Figure 9b are symmetric and have dispersed shapes.

Figure 9: Kernel density of efficiency scores



(a) Total efficiency scores

(b) Period-averaged efficiency scores

6 Conclusions

In this study, we examine how to measure industry-level productivity with cross-sectional dependence. We propose a method for choosing the appropriate weights matrix when no explicit distance concept is available. In particular, we construct a unified measure to characterize the linkage between two industries, in which the linkage includes direct and indirect demand-side and supply-side effects. An economic distance measure that is analogous to geographic distance is defined by the relative size of our linkage measure to the measure of the most closely linked industry. We also specify a spatial production model by expanding a traditional Cobb–Douglas production function with two basic spatial specifications, namely, SAR and SDM. We estimate the model by using the CSS-type frontier production approach, which allows us to parsimoniously estimate time-varying efficiency levels.

The spatial approach allows us to measure indirect effects that inevitably occur due to spatial interdependency. As a result, we found that the total estimated output elasticities of factor inputs are larger than those from a non-spatial specification. The indirect effects are approximated as 25% and 21% of the total elasticities for SARCSS and SDMCSS, respectively. We can therefore conclude that the U.S. economy has increasing returns to scale in the last six decades when only the spatially weighted dependent variable is included in the model. However, the returns to scale is not significantly increasing if we additionally assume that the factor inputs have cross-sectional dependence.

Although the coefficient estimates vary across the model specifications, the relative efficiency scores can be estimated comparably. For instance, in comparison with the TFP measure from the growth accounting approach, our regression-based approach has smoother efficiency estimates in terms of level and growth. Moreover, we observe that the *electrical and optical equipment* sector is the least efficient industry on the average in over 60 years, even if its TFP growth is more rapid than those of the other industries. Meanwhile, the *construction* sector is the most efficient industry on the average in terms of productivity level, but it has slightly lower growth than that of the industry average.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Battese, G. E. and T. J. Coelli (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics 20*(2), 325–332.

Behrens, K., C. Ertur, and W. Koch (2012). 'Dual' Gravity: Using Spatial Econometrics to control for multilateral resistance. *Journal of Applied Econometrics 27*, 773–794.

Brunsdon, C., A. S. Fortheringham, and M. E. Charlton (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis 28*(4), 281–298.

Chudik, A. and M. H. Pesaran (2013). A review and comparision of tests of cross-section independence in panels. In B. H. Baltagi (Ed.), *The Oxford Handbook on Panel Data*. Oxford University Press.

Cobb, C. W. and P. H. Douglas (1928). A Theory of Production. *The American Economic Review 18*(1), 139–165.

Coelli, T. J., D. Rao, O. Prasada, and G. E. Battese (2005). *An Introduction to Efficiency and Productivity Analysis*. Springer.

Cornwell, C., P. Schmidt, and R. C. Sickles (1990). Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics 46*(1), 185–200.

Diewert, W. E. (1976). Exchange superlative index numbers. *Journal of Econometrics 4*, 115–145.

Diewert, W. E. (1978). Superlative index numbers and consistency in aggregation. *Econometrica 46*(4).

Druska, V. and W. C. Horrace (2004). Generalized moments estimation for spatial panel data: Indonesian rice farming. *American Journal of Agricultural Economics 86*, 185–198.

Elhorst, J. P. (2014). *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*. Springer.

Ertur, C. and W. Koch (2007). Growth, Technological interdependence and Spatial Externalities: Theory and Evidence. *Journal of Applied Econometrics 22*, 1033–1062.

Getis, A. and J. Aldstadt (2004). Constructing the spatial weights matrix using a local statistic. *Geographical Analysis 36*(2), 90–104.

Glass, A., K. Kenjegalieva, and J. Paez-Farrell (2013). Productivity growth decomposition using a spatial autoregressive frontier model. *Economics Letters 119*, 291–295.

Glass, A., K. Kenjegalieva, and R. C. Sickles (2015). Returns to scale and curvature in the presence of spillovrs: evidence from european countries. *Oxford Economic Papers 67*(4), 1–24.

Glass, A., K. Kenjegalieva, and R. C. Sickles (2016a). Returns to Scale and Curvature in the Presence of Spillovers: Evidence from European Countries. *Oxford Economic Paper 68*, 40–63.

Glass, A., K. Kenjegalieva, and R. C. Sickles (2016b). A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers. *Journal of Econometrics 190*, 289–300. Rice University.

Griliches, Z. (1960). Estimates of the aggregate U.S. farm supply function. *American journal of Agricultural Economics 42*(2), 282–293.

Han, J., D. Ryu, and R. C. Sickles (2016a). How to measure spillover effects of public capital stock: A spatial autoregressive stochastic frontier model. In T. B. Fomby, R. C. Hill, I. Jeliazkov, J. C. Escanciano, and E. Hillebrand (Eds.), *Spatial Econometrics: Qualitative and Limited Dependent Variables*, Volume 37, Chapter 9, pp. 259–294. Emerald.

Han, J., D. Ryu, and R. C. Sickles (2016b). Spillover effects of public capital stock using spatial frontier analyses: A first look at the data. In W. H. Greene, R. Sickles, L. Khalaf, M. Veall, and M.-C. Voia (Eds.), *Productivity and Efficiency Analysis: Proceedings from the 2014 North American Productivity Workshop*. New York: Springer Publishing.

Hulten, C. R. (2001). Total factor productivity: A short biography. In C. R. Hulten, E. R. Dean, and M. J. Harper (Eds.), *New developments in productivity analysis*. Chicago: University of Chicago Press.

Jones, C. I. (2012). Misallocation, economic growth, and Input-Output economics. In D. Acemoglu, M. Arellano, and E. Dekel (Eds.), *Advances in Economics and Econometrics*. New York: Cambridge University Press.

Jorgenson, D. W., Gollop, and B. M. Fraumeni (1987). *Productivity and U.S. Economic Growth*. Amsterdam: North-Holland.

Jorgenson, D. W. and Z. Griliches (1967). The explanation of productivity change. *The Review of Economic Studies 34*(3), 249–283.

Jorgenson, D. W., M. S. Ho, and J. D. Samuels (2012). A prototype industry-level production account for the United States, 1947-2010. presentation to the Final World Input-Output Database Conference, Groningen, The Newtherlands.

LeSage, J. and R. K. Pace (2009). *Introduction to Spatial Econometrics*. Boca Raton, Florida: Chapman & Hall/CRC.

McMillen, D. P. (2003). Identifying sub-centres using contiguity matrices. *Urban Studies 40*(1), 57–69.

Phillips, P. C. B. and D. Sul (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal 6*, 217–259.

Pitt, M. and L. Lee (1981). The measurement and sources of technical inefficiency in the indonesian weaving industry. *Journal of Development Economics 9*, 43–64.

Schmidt, P. and R. C. Sickles (1984). Production frontiers and panel data. *Journal of Business and Economic Statistics 2(4)*, 367–374.

Solow, R. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics*.

Sonis, M. and G. J. D. Hewings (1999). Economic landscapes: multiplier product matrix analysis for multiregional Input-Output systems. *Hitotsubashi Journal of Economics 40*, 59–74.

Timmer, M., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2015). An Illustrated User Guide to the World Input-Output Database: The Case of Global Automotive Production. *Review of International Economics*.

Timmer, M., A. A. Erumban, B. Los, R. Stehrer, and G. J. de Vries (2014). Slicing Up Global Value Chains. *Journal of Economic Perspectives 28*(2), 99–118.

Timmer, M. and X. Ye (2017). Productivity and substitution patterns in global value chains. In E. G. Tatje, C. A. K. Lovell, and R. C. Sickles (Eds.), *Oxford Handbook of Productivity*, Chapter 21. New York: Oxford University Press.

Ward, J. H., Jr. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association 58*(201), 236–244.

Yang, Z. (2013). Quasi-maximum likelihood estimation for spatial panel data regressions. Research Collection School of Economics.

A A derivation of estimation procedure

Quasi-Maximum Likelihood Estimator for ρ

Let $\psi = (\beta, \gamma, \rho, \sigma_v^2)'$. The log-likelihood function of Eq.(5) is:

$$logL(\psi, \delta_i; y) = -\frac{NT}{2} log(2\pi\sigma_v^2) + Tlog|I_N - \rho W| - \frac{1}{2\sigma_v^2} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(y_{it} - \rho \sum_{j=1}^{N} w_{ij} y_{jt} - X'_{it}\beta - Z'_i\gamma - R'_t\delta_i \right)^2, \quad (23)$$

The first order condition of maximizing Eq. (23) with respect to δ_i is

$$\frac{\partial logL}{\partial \delta_i} = \frac{1}{\sigma_v^2} \sum_{i=1}^{N} \sum_{t=1}^{T} R_t \left(y_{it} - \rho \sum_{j=1}^{N} w_{ij} y_{jt} - X'_{it}\beta - Z'_i\gamma - R'_t\delta_i \right) = 0. \quad (24)$$

By solving for (24), we can obtain

$$\hat{\delta}_i = (R_t R'_t)^{-1} R_t \left(y_{it} - \rho \sum_{j=1}^{N} w_{ij} y_{jt} - X'_{it}\beta - Z'_i\gamma \right). \quad (25)$$

Substituting (25) into the log-likelihood function, (23), we obtain the concentrated likelihood function

$$logL(y; \beta, \rho, \sigma_v^2) = -\frac{NT}{2} log(2\pi\sigma_v^2) + Tlog|I_N - \rho W| - \frac{1}{2\sigma_v^2} \tilde{V}'\tilde{V}, \quad (26)$$

where $\tilde{V} = M_Q y - \rho M_Q (W_N \otimes I_T) y - M_Q X\beta$, and $M_Q = I_{NT} - Q(Q'^{-1}Q$[13].

Within Estimator

Assuming $T \geq L$, the projections onto the column space of Q and the null space of Q are denoted by $P_Q = Q(Q'^{-1}Q'$ and $M_Q = I_{NT} - P_Q$, respectively[14]. Let's suppose the true value of ρ is known, say ρ^*. By pre-multiplying M_Q on (6), we have the within-transformed model

$$M_Q y = \rho^* M_Q (W_N \otimes I_T) y + M_Q X\beta + \tilde{V}. \quad (27)$$

And the estimates of $\beta(\rho^*)$ and of $\sigma_v^2(\rho^*)$ are derived by

$$\hat{\beta}_W(\rho^*) = (X' M_Q X)^{-1} X' M_Q (y - \rho^* (W_N \otimes I_T) y), \quad (28)$$

$$\hat{\sigma}_v^2(\rho^*) = \frac{1}{N(T-L) - K} e(\rho^*)'^*), \quad (29)$$

[13] Note that the variables in Z do not vary over time. Hence γ cannot be identified because $M_Q\boldsymbol{Z} = 0$.

[14] Q need to be a full column rank matrix for estimation of the individual δ_i.

respectively, where $e(\rho^*) = y - \rho^*(W_N \otimes I_T)y - X\hat{\beta}_W(\rho^*)$. By substituting the closed form solutions for the parameters $\beta(\rho^*)$ and $\sigma_v^2(\rho^*)$ to Eq. (26), we can concentrate out β and σ_v^2, and the concentrated log-likelihood function with single parameter ρ is of the form:

$$logL(y;\rho) = C - \frac{NT}{2} log\left[e(\rho)'e(\rho)\right] + Tlog|I_N - \rho W|, \tag{30}$$

where C is a constant term that is not a function of ρ. By maximizing the concentrated log-likelihood function Eq.(30) with respect to ρ, we can obtain the optimal solution for ρ. Even if there is no closed-form solution for ρ, we can find a numerical solution because the equation is concave in ρ. Finally, the estimators for β and σ^2 can be calculated by plugging $\rho^* = \hat{\rho}$ in Eq. (28) and Eq. (29).

The asymptotic variance-covariance matrix of parameters (β, ρ, σ^2) is given by:

$$\begin{aligned} &Asy.Var(\beta, \rho, \sigma_v^2) \\ &= \begin{bmatrix} \frac{1}{\sigma_v^2}\tilde{X}'\tilde{X} & \frac{1}{\sigma_v^2}\tilde{X}'(W^* \otimes I_T)\tilde{X}\beta & \mathbf{0} \\ - & T \cdot tr(W^*W^* + W^{*'}W^*) + \frac{1}{\sigma_v^2}\beta'\tilde{X}'(W^{*'}W^* \otimes I_T)\tilde{X}\beta & \frac{T}{\sigma_v^2}tr(W^*) \\ - & - & \frac{NT}{2\sigma_v^4} \end{bmatrix}^{-1}, \end{aligned} \tag{31}$$

where $\tilde{X} = M_Q X$ and $W^* = W(I_N - \rho W)^{-1}$.

Generalized least squares estimator

Alternatively, we can estimate Eq. (6) by generalized least squares(GLS). Denote the variance-covariance matrix of the composite error $\varepsilon = QU + V$ as $cov(\varepsilon) = \Omega$.

The GLS estimator is the SAR estimator applied to the following transformed equation:

$$\begin{aligned} \sigma_v\Omega^{-1/2}y = {} & \rho\sigma_v\Omega^{-1/2}(W_N \otimes I_T)y + \sigma_v\Omega^{-1/2}X\beta + \sigma_v\Omega^{-1/2}\boldsymbol{Z}\gamma \\ & + \sigma_v\Omega^{-1/2}\boldsymbol{R}\delta_0 + \sigma_v\Omega^{-1/2}\varepsilon, \end{aligned} \tag{32}$$

where $\varepsilon = QU + V$, $\Omega = cov(\varepsilon) = \sigma_v^2 I_{NT} + Q(I_N \otimes \Delta)Q'$. The estimation procedure of Eq. (32) is same as the procedure for within-estimation. Let $\eta = (\beta, \gamma, \delta_0)$. Assuming we know the true value of $\rho = \rho^*$, the GLS estimators of $\eta(\rho^*)$ are

$$\begin{aligned} \hat{\eta}_G(\rho^*) &= [(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})]^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(y - \rho^*(W_N \otimes I_T)y) \\ &= [(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})]^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}y \\ &\quad - \rho^*[(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})]^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(W_N \otimes I_T)y. \end{aligned} \tag{33}$$

Hence, the GLS estimators of η can be represented as a difference of OLS estimators of regressing $\widetilde{y}$ on $(\widetilde{X}, \widetilde{\boldsymbol{Z}}, \widetilde{\boldsymbol{R}})$ and regressing $\widetilde{(W_N \otimes I_T)}y$ on $(\widetilde{X}, \widetilde{\boldsymbol{Z}}, \widetilde{\boldsymbol{R}})$ premultiplied by the spatial autoregressive coefficient ρ^*, where tilde represents GLS transformation. Ω can be estimated by:

$$\hat{\Omega}(\rho^*) = \hat{\sigma}_v^2 I_{NT} + Q(I_N \otimes \hat{\Delta}(\rho^*))Q'. \tag{34}$$

Following Cornwell et al. (1990), Δ can be estimated as

$$\hat{\Delta}(\rho^*) = \frac{1}{N}\sum_{i=1}^{N}\left[(R'^{-1}R'e_i e_i' R(R'^{-1} - \hat{\sigma}_v^2(R'^{-1}\right], \tag{35}$$

where $e_i = M_R y - \rho^* M_R(W_N \otimes I_T)y - M_R X\hat{\beta}_W(\rho^*)\Big|_i$, which represents the IV residuals for individual i, and $M_R = \boldsymbol{R(R'^{-1}R'}$ is the projection onto the column space of $\boldsymbol{R}$.

Consider the likelihood function of Eq. (32). Since $\sigma_v \Omega^{-1/2}\varepsilon$ has mean zero and variance σ_v^2, the likelihood function is written in the form of

$$logL(\eta, \rho; y) = -\frac{NT}{2}log(2\pi\sigma_v^2) + Tlog|I_N - \rho W| - \frac{1}{2}\varepsilon'^{-1}\varepsilon, \tag{36}$$

where $\varepsilon = y - \rho(W_N \otimes I_T)y - X\beta - \boldsymbol{Z}\gamma - \boldsymbol{R}\delta_0$. Substitution of Eq. (33) and Eq. (34) into Eq. (36) gives a concentrated likelihood function as follows:

$$logL(\rho; y) = C - \frac{NT}{2}log\left[e(\rho)'e(\rho)\right] + Tlog|I_N - \rho W|, \tag{37}$$

where $e(\rho) = \hat{\sigma}_v\hat{\Omega}(\rho)^{-1/2}y - \rho\hat{\sigma}_v\hat{\Omega}(\rho)^{-1/2}(W_N \otimes I_T)y - \hat{\sigma}_v\hat{\Omega}(\rho)^{-1/2}X\hat{\beta} - \hat{\sigma}_v\hat{\Omega}(\rho)^{-1/2}\boldsymbol{Z}\hat{\gamma} - \hat{\sigma}_v\Omega\hat{(\rho)}^{-1/2}\boldsymbol{R}\hat{\delta}_0$, and C is a constant term that is not a function of ρ. Finally, as is the case of within estimator, we can obtain the estimators for η and Ω using the estimate of ρ from Eq. (37).

Implementation

For the implementation, we combine the procedure suggested by Elhorst (2014) and a typical two-stage approach of FGLS. In this section we discuss the implementation of within estimator and then turn to the implementation of the GLS estimator. The implementation consists of the following steps.

[Within Estimator]
From Eq. (28), it is easy to show that $\hat{\beta}_W = b_0 - \rho^* b_1$, where b_0 and b_1 are the OLS estimators of regressing $M_Q y$ and $M_Q(W_N \otimes I_T)y$ on $M_Q X$, respectively. Similarly, the estimated residuals from Eq. (27), $e(\rho^*)$, can be expressed as $e(\rho^*) = e_0 - \rho^* e_1$, where e_0 and e_1 are the associated OLS residuals to b_0 and b_1, respectively. Hence the first step is obtaining b_0, b_1, e_0, and e_1. Second, we

maximize Eq. (30) with respect to ρ after replacing $e(\rho) = e_0 - \rho e_1$, i.e.,

$$\max_{\rho} logL(y|\rho) = C - \frac{NT}{2} log\left[(e_0 - \rho e_1)'(e_0 - \rho e_1)\right] + Tlog|I_N - \rho W|. \quad (38)$$

Third, by replacing $\rho^* = \hat{\rho}$ in Eq. (28) and (29) gives the within estimator, $\hat{\beta}_W$, and the estimated variance, $\hat{\sigma}^2$. Finally, the asymptotic variance-covariance matrix of parameters $(\hat{\beta}_W, \hat{\rho}, \hat{\sigma}_v)$ can be calculated by Eq. (31).

[GLS Estimator]
Unlike the within estimator case, we are unable to find the separate OLS estimators of regressing $\sigma_v \Omega^{-1/2} y$ and $\sigma_v \Omega^{-1/2}(W_N \otimes I_T)y$ on $\sigma_v \Omega^{-1/2}(X, \boldsymbol{Z}, \boldsymbol{R})$ in advance of having $\hat{\rho}$, even if Eq. (33) is expressed as a subtraction of two terms. This is because the feasible Ω is obtainable only after we have a value for ρ. Instead of following the steps of within estimator, we can obtain $\hat{\rho}$ by simply maximizing the concentrated log-likelihood function (37). Once we have $\hat{\rho}$, Eq. (35), Eq. (34), Eq. (33) give $\Delta(\hat{\rho})$, $\Omega\hat{\rho}$, and $\eta_G(\hat{\rho})$ in order.

B Estimation results with Value-added dependent variable

Table B.1: CSS vs SARCSS vs SDMCSS: Value-added

	Non-spatial				SAR				SDM			
	CSSW		CSSG		CSSW		CSSG		CSSW		CSSG	
	Coef.	std.err	Coef.	std.err	Coef.	std.err	Coef.	std.err	Coef.	std.err	Coef.	std.err
$log(K)$	0.159	0.057	0.199	0.052	0.110	0.056	0.155	0.052	0.064	0.061	0.121	0.055
$log(L)$	0.388	0.057	0.400	0.053	0.302	0.056	0.325	0.053	0.210	0.061	0.245	0.057
$Intercept$	-	-	-0.661	0.259	-	-	-0.383	0.265	-	-	-0.119	0.299
$Time$	-	-	0.023	0.005	-	-	0.011	0.005	-	-	0.001	0.007
$Time^2$	-	-	0.000	0.000	-	-	0.000	0.000	-	-	0.000	0.000
$sigma_v^2$	0.029	0.000	0.029	0.000	0.028	0.001	0.028	0.001	0.027	0.001	0.027	0.001
Spatial Parameters												
$W \cdot log(Y)(\rho)$	-	-	-	-	0.386	0.051	0.364	0.043	0.204	0.068	0.208	0.066
$W \cdot log(K)(\lambda_1)$	-	-	-	-	-	-	-	-	0.250	0.119	0.196	0.070
$W \cdot log(L)(\lambda_2)$	-	-	-	-	-	-	-	-	0.538	0.123	0.496	0.119
η	-	-	-	-	1.670		1.276		1.170		1.700	
Elasticities												
θ_K	0.159	0.057	0.199	0.052	0.182	0.093	0.241	0.043	0.399	0.136	0.398	0.036
θ_L	0.388	0.057	0.400	0.053	0.493	0.099	0.517	0.084	0.942	0.131	0.931	0.126
R^2	0.189		0.218		0.253		0.274		0.295		0.310	
$Adjusted R^2$	0.148		0.176		0.215		0.236		0.258		0.272	
$loglikelihood$	-		-		576.111		542.115		587.847		554.150	
$Returns to scale$	0.547		0.599		0.675		0.758		1.341		1.329	

Table B.2: Direct, Indirect, and Total Elasticity: Value-added

		Direct		Indirect		Total	
		Coef.	asy. t-stat	Coef.	asy. t-stat	Coef.	asy. t-stat
SARCSSW	Capital	0.113**	1.984	0.069**	1.813	0.182**	1.963
	Labor	0.305***	5.378	0.188***	3.510	0.493***	4.978
SARCSSG	Capital	0.156***	4.325	0.085***	8.183	0.241***	5.583
	Labor	0.331***	6.308	0.186***	4.371	0.517***	6.141
SDMCSSW	Capital	0.070	1.199	0.329***	2.383	0.399***	2.938
	Labor	0.215***	3.408	0.727***	5.457	0.942***	7.182
SDMCSSG	Capital	0.125**	2.321	0.273***	4.411	0.398***	10.931
	Labor	0.251***	4.388	0.680***	5.303	0.931***	7.412

Note *p<.05, **p<.01, ***p<.001

C Industry classifications and ISIC Rev.3 codes

Table C.1: Industry classifications and codes

No.	Industry	ISIC Rev. 3
1	Agriculture, Hunting, Forestry and Fishing	AtB
2	MINING AND QUARRYING	C
	TOTAL MANUFACTURING	
3	Food, Beverages and Tobacco	15t16
4	Textiles, Textile, Leather and Footwear	17t19
5	Wood and Products of Wood and Cork	20
6	Pulp, Paper, Paper , Printing and Publishing	21t22
7	Coke, Refined Petroleum and Nuclear Fuel	23
8	Chemicals and Chemical Products	24
9	Rubber and Plastics	25
10	Other Non-Metallic Mineral	26
11	Basic Metals and Fabricated Metal	27t28
12	Machinery, Nec	29
13	Electrical and Optical Equipment	30t33
14	Transport Equipment	34t35
15	Manufacturing, Nec; Recycling	36t37
16	ELECTRICITY, GAS AND WATER SUPPLY	E
17	CONSTRUCTION	F
	WHOLESALE AND RETAIL TRADE	G
18	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	50
19	Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	51
20	Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	52
21	HOTELS AND RESTAURANTS	H
	TRANSPORT AND STORAGE AND COMMUNICATION	I
22	Transport and Storage	60t63
23	Post and Telecommunications	64
24	FINANCIAL INTERMEDIATION	J
	REAL ESTATE, RENTING AND BUSINESS ACTIVITIES	K
25	Real Estate Activities	70
26	Renting of M&Eq and Other Business Activities	71t74
27	PUBLIC ADMIN AND DEFENCE; COMPULSORY SOCIAL SECURITY	L
28	EDUCATION	M
29	HEALTH AND SOCIAL WORK	N
30	OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES	O