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

Jaepil Han * Robins C. Sickles †

May 23, 2017

Abstract

This paper examines aggregate productivity in the presence of intersectoral linkages. Cross-sectional dependence is inevitable among industries as each sector serves as supplier to other sectors and the chains of the interconnections cause indirect relationship among industries. Spatial analysis is one of the approaches that address cross-sectional dependence using a priori specified spatial weights matrix. We exploit the linkage patterns from the input-output tables and use them to assign spatial weights that describe economic interdependencies. Using the spatial weights matrix, we estimate industry-level production functions and productivity of the U.S. for the period from 1947 to 2010. Cross-sectional dependencies, which are the consequence of indirect effects reflecting interactions among industries linked via their supply change networks results in larger output elasticities as well as scale effects for the networked production process. Productivity growth estimates, however, are found to be 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

*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

1 Introduction

The aggregate productivity of a nation can be decomposed into the productivity of each industry and the allocation of factor inputs among industries. These two may interact with each other. A productivity shock in one sector may result in misallocation of inputs in other sectors, and/or misallocation of inputs in a sector may invoke productivity shocks in other sectors due to production linkages via the production supply chain. Such linkeages have been studied extensively in the literature. Excellant recent works by Timmer et al. (2014, 2015, 2017) provide new insights and an excellent summary of many of the productivity aspects of supply chain networks. Jones (2012) recognizes the possibility that the effects of misallocation can be amplified through the input-output structure. He points that the contagion of the negative effects caused by misallocations can reduce total factor productivity. Acemoglu et al. (2012) also have a similar idea that microeconomic idiosyncratic shocks can lead to aggregate fluctuations through the input-output linkages. In other words, 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 with a variety of methodologies, but most studies do not take possible cross-sectional dependencies into account. Much of the current research focuses on the contributions of industries to aggregate productivity. Growth accounting techniques and index-number approaches are widely used in many statistical agencies, but these approaches estimate the productivity growth of each industry separately assuming industries are independent and do not consider possible interdependencies among them. However, interdependency is inevitable, especially among industries of an economy since the outputs of many sectors are used as the inputs of many others.

There is substantial evidence that the contributions of each industry to aggregate productivity change over time. As 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) changes in both the level of industry productivity and its relative importance in the economy will impact its contribution to aggregate productivity growth. The relative importance of industries also may be affected by the linkages between industries. For example, fig. 1 displays the evolution of the share of industries to gross output of the United States by comparing the shares in

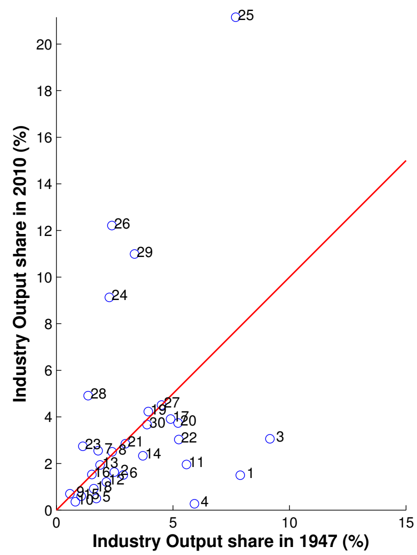


Figure 1: A comparison of Gross-output share of industries: 1947 vs. 2010

the 1947 and 2010[1]. It is clear that 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 share, while the tertiary industries, such as *Financial Intermediation* (Ind24), *Real estate, renting and business activities* (Ind25, Ind26), *Education*, and *Health and social work* (Ind29) increased their share. The change in output shares can be explained as a result of industrial structural change in general, but it also can be interpreted as a transition of key sectors. Jorgenson et al. (2012) also make arguments about the possible influential power of some key sectors, pointing that the role of the non-IT industries that shrunk while the contributions of IT-producing and IT-using industries grew.

Research on productivity measurement has a long history and 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 by Solow (1957), Griliches(1960) and Jorgenson and Griliches (1967). In particular, Solow (1957) measured productivity growth by the residuals of output growth not 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

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

productivity growth measurement (Jorgenson and Griliches, 1967; Diewert, 1976). Direct estimation of the production function using econometric techniques have also been pursued and as Hulten (2001) points out, there exist a number of pitfalls in the econometric approach, such as possible odd shaped isoquants without appropriate a priori restrictions, the abundance of parameters to estimate with limited information, and the complication of flexible models. Even though there are caveats in the econometric approach, it has an advantage in providing for model flexibility. Such flexibility extends to the econometric setting in which panel data is utilized, employing fixed or random effects specifications to address firm specific unobserved heterogeneity. However, the econometric approach often does not address spatial interactions into account. Exceptions to this 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 as well as frontier production/cost models. Models that extend the multiplicative spillover effects by framing production in a spatial autoregressive setting in order to address network effects or trade flows among countries have been formulated by Ertur and Koch (2007) and Behrens et al. (2012). More 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 model in a series of papers by Glass et al. (2013, 2015, 2016a,b) and by Han et al. (2016).

It is well-known that ignored cross-sectional dependence causes standard OLS estimators to be inefficient and the estimated standard errors to be biased.(Phillips and Sul, 2003; Chudik and Pesaran, 2013). We expand the standard stochastic frontier model to spatially dependent specifications in the present paper.

Models with spatial structures are found in fields such as regional science, economic geography and urban economics, but relatively recently the methodology has been applied in many 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 by spatial parameters. The spatial weights matrix is usually formed based on spatial (geographical) characteristics, but 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, the right choice of a spatial weights matrix is crucial for the correct model specifications.

A contribution of this paper is to propose a novel approach for creating a spatial weights matrix based on economic distance when physical distance does not properly capture the spatial linkages. The data we are examining is industry-specific and thus a formal distance metric based on geographical distance has no obvious appeal. Instead, we define economic distance, analogous to a geographic distance, using the supply flows that can be found from Input-Output tables. Below we specify spatial production models utilizing spatial autoregressive models (SAR) and spatial Durbin model (SDM) and estimate the model with Cornwell et al. (1990) type stochastic frontier approaches. We also utilize alternative methods suggested by Levinsohn and Petrin (2003) that are based on a different set of instruments to handle potential endogeneity of the factor inputs. Using the weights matrix created from the Input-Output tables, we examine the production technology and industry-level productivity of the United States during the period 1947 - 2010.

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 using the Input-Output tables and form a spatial weights matrix based on it. In section 4 we apply our methodology to the estimation of a production frontier function using industry-level data for the U.S. during 1947 - 2010. Section 5 gives some concluding remarks and possible extensions for the future research.

2 Models

2.1 A 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 taking the logarithm of Eq. (1) we get 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 $\mathbf{X}_{it}$, respectively, α represents the intercept, and ε_{it} is the unobservable. The ε_{it} term can be decomposed into a productivity term u_{it}, which is may be known to the firm or about which firms in an industry are well-informed, but not known to the econometrician, and a statistical noise term v_{it}. Then the model becomes

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

Cornwell et al. (1990) extend the model (2) to allow heterogeneity in slopes and intercepts. They model 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 a $L \times 1$ time-varying component that globally affects all individual units, and δ_i is $L \times 1$ the coefficients that depend on i. The time dependent individual effect, α_{it}, can be assumed to be decomposed into a common time trend $R_t'\delta_0$ and a unit-specific term $R_t'u_i$. After adding a time-invariant fixed effects variable Z_i for a more 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, the u_i are assumed to be iid zero mean random variables with covariance matrix Δ, and v_{it}, is a random noise following $iidN(0, \sigma_v^2)$. If R contains only a constant, for instance, Eq. (4) reduces 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 the estimated standard errors are biased if cross-sectional dependence in the error term is ignored. Of course when the structural model should be represented by a spatial autoregressive model, then ignoring such spatial linkages renders estimates biased and inconsistent. Spatial analysis is one of the approaches that can address cross-sectional dependence explicitly. Spatial econometric models specify an a priori spatial weights matrix which captures the explicit dependence structure assumed to exist among units. Based on the exogenous spatial weights matrix, the spatial approaches essentially include new variables consisting of the weighted averaged variables of neighboring observations. Omission of such spatially correlated variables will cause omitted variable bias for the estimation of the coefficient parameters.

The spatial econometric approaches model the possible spatial interactions via determining the type and extent of spatial dependence that exists between economic units. Three types of spatial dependence are usually considered, reflecting interactions that should be modeled by the inclusion of spatially weighted dependent variables (Spatial Autoregressive Model; SAR), spatially weighted independent variables (Spatial lagged X regression model; SLX), or spatial dependence captured in the error term (Spatial Error Model; SEM). Of course, it is possible to allow for more than one dependence structure. In the present paper, we consider either a spatially weighted dependent variable (SAR) or both spatially weighted dependent variable and independent variables (Spatial Durbin Model; SDM) to avoid problems with over-parameterization and to focus on a relatively parsimonious setting to explore our new approach to constructing the weighting matrix.

The Spatial Autoregressive specification associated with (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, which will be given exogenously, the u_i are assumed to be *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 (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 a $NT \times K$ matrix, $\boldsymbol{Z} = (Z \otimes \iota_T)$, Z is $N \times J$ matrix, ι_T is 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 $NT \times LN$ matrix, β is a $K \times 1$, γ is a $J \times 1$, δ_0 is a $L \times 1$ vector, and U is a $LN \times 1$ vector.

The Spatial Durbin specification associated with (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[3], which will be given exogenously, the u_i are assumed to be iid zero mean random variables with covariance matrix Δ, and v_{it}, is a random noise following $N(0, \sigma_v^2)$. Likewise, the matrix form of (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 a $NT \times K$ matrix, $\boldsymbol{Z} = (Z \otimes \iota_T)$, Z is $N \times J$ matrix, ι_T is 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 $NT \times LN$ matrix, β is a $K \times 1$, γ is a $J \times 1$, δ_0 is a $L \times 1$ vector, and U is a $LN \times 1$ vector.

3 Spatial Weights Matrix: Economic Distance

A spatial weights matrix is a crucial element in spatial econometric models. It is often specified a priori and assumed to be an exogenous conceptualization of the structure of spatial dependence. Getis and Aldstadt (2004) point out that a model with a wrong choice of spatial weights matrix is essentially misspecified. Unlike spatial weights matrix in typical spatial analyses relying on geographic relationships, we are not able to define physical distances among units because we

[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 and the faster index is over individuals.

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

examine the aggregate production function 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. Han et al. (2016b) use the relative bilateral trade volume of a country as an economic distance measure. To define economic distance, we should understand the economic relationships among industries. Using the input-output table is appropriate for this purpose. Input-output analysis was developed by Leontief in the late 1930s with the fundamental purpose of analyzing the interdependence of industries in an economy. An input-output table is constructed from 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 other sectors, and the intersectoral flows are measured in monetary terms. The reason why accounts are kept in monetary terms is mainly because there will be measurement problems if we consider different products with different prices only in physical units. Table 1 illustrates a hypothetical table for a two-sector economy. The total output of sector 1 is consumed by each sector (z_{11}, z_{12}), household consumers, government, and abroad(F_1). The last row shows the total value of inputs to each of the industries. Within each industry in the processing sector all of the receipts from sales are paid out for goods and services purchased from other industries. After taking into account appropriate inventory changes, the total sector output is equal to the total outlays made by that industry.

The components of the input-output table cannot directly represent the linkage of a pair of industries and cannot directly be compared to each other because they are not scale-adjusted. Suppose $z_{11} = 1$, $z_{12} = 2$, $z_{21} = 2$, $z_{22} = 3$. In this example, the relative importance of sector 2 in sector 1, 2/3, is greater than the relative importance of sector 2 in sector 2, 3/5, even though z_{22} is larger than z_{12}. However, we can think of a straightforward measure for industrial linkage, which adjust the scaling effects, as the input and output direct requirements as illustrated in Eqs. (9) and (10). The input direct requirements coefficients show the amount of inputs purchased directly to produce one dollar of output. On the other hand, the output direct requirements coefficients show the proportion of outputs of a specific sector, which is sold to each purchasing sector. With the direct requirements, however, we only capture the direct relationship out of four types of relationships: direct, indirect, demand-side, supply-side.

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

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

From the definitions of the direct requirements coefficients, we have proportional relationships between sector outputs (Y) and intermediate inputs(Z): $a_{ij} = \frac{z_{ij}}{Y_j}$ and $\bar{a}_{ij} = \frac{z_{ij}}{Y_i}$. In the input-output table, total output is supposed to be used up by the sectors that require the products as intermediate inputs and final consumers. In addition, the total outlay is the sum of intermediate inputs of each industry. Hence, we have the following two relations: $Y = Zi + F$, and $Y = i'Z + VA$, where i is a vector of ones. Using the relations and the proportional relationships, we have the relationships among Y, F, A, and $\bar{A}$ as

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

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

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

The inverse matrices describing the relationship between intermediate inputs and final consumption, $B = (I - A)^{-1}$ and $\bar{B} = (I - \bar{A})^{-1}$ are called *Leontief inverse matrix* and *Ghosh inverse matrix*, respectively. The *Leontief inverse matrix* illustrates how much production will be induced in what industry by a one unit increase in demand by a certain industry. The *Ghosh inverse matrix* is derived when we use the output direct requirements matrix($\bar{A}$) and the relationship between intermediate inputs supply and payments of each sector from supply-side. The basic assumption of the supply-side inverse matrix is that one might expect increasing sales from sector i to each of other sectors when output of sector i increases. Even if the inverse matrices include both of direct and indirect intermediate flows required to produce a dollar of output, they are interpreted differently depending on which matrix and element is considered.

There have been attempts to make a unified measure for economic effects of a particular sector on other sectors. Backward linkage and forward linkage are the mostly used measures. If industry j increases output, there will be increased demands from industry j on the industries whose products are used as intermediate inputs. Backward linkage indicates the interrelation of a particular industry to those industries from which it purchases intermediate inputs. On the other hand, increased output in industry j means increased supplies from industry j for the industries which use the commodity j in their production. Forward linkage indicates the interrelation of a particular industry to those industries to which it sells its output. Backward linkage and forward linkage are derived from the *Leontief inverse matrix* and the *Ghosh inverse matrix*, respectively. The direct backward linkage($\mathcal{B}$) and direct forward linkage($\mathcal{F}$) 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 $\bar{b}_{ij}$ is 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 economic influences of expansion and shrinkage of industries as a whole, they do not provide consistent measures for linkages between industries. In other words, a large backward linkage does not guarantee a large forward linkage and vice versa. Consider the following hypothetical supply flows of intermediate inputs illustrated in figure 2. Sector A supplies for itself, so its expansion cannot induce output growth in other sectors in terms of forward linkage. However, sector A can have an influence on other sectors in terms of backward linkage since it uses products from all sectors.

To this end, Sonis and Hewings (1999) proposed an index called the Multiplier Product Matrix (MPM), which connects the properties of backward and forward linkages as follows:

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

where

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

MPM provides a measure of the relationships among industries that allows them to be organized into a rank-sized hierarchy. MPM measures the impacts of an industry on other industries. MPM is not symmetric because the forward linkage of a particular industry is mostly different than the backward linkage of the industry (see, for example, fig. 4). In order to utilize the MPM to proxy for economic distance we modify it to be symmetric by taking the Euclidean norm of the elements of the MPM, i.e.,

$$m_{ij}^E = m_{ji}^E = \sqrt{m_{ij}^2 + m_{ji}^2}. \tag{13}$$

3.2 An Economic Distance and Weights matrix

We utilize the multiplier product matrix defined in the previous section for creating a spatial weights matrix based on economic distance as geographic distance is not a meaningful spatial

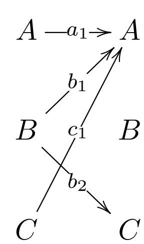


Figure 2: Hypothetical supply flows of intermediate inputs

concept linking the supply chains embedded in the sectoral flows among industries. Although at first glances one may think that values of the multiplier product matrix can be used to fill in the weighting matrix elements in this matrix have the opposite interpretations to 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 as well as introduce a distance-decay function to give more weight on a more closely related industry.

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

$$d_{ij} \equiv \max_{i'} m^E_{i'j} - m^E_{ij},$$

where m^E_{ij} is the element of multiplier product matrix MPM developed in the previous section[4]. Once we have economic distance we can assign weights according to the distance. There are several schemes to assign weights, such as contiguous neighbors, inverse distances, lengths of shared borders divided by the perimeter, bandwidth, centroid distance, and k nearest neighbors. A widely using weighting scheme for a spatial weights matrix excludes observations that are further than a threshold distance d^*, that is,

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

However, the spatial weighting function (3.2) suffers the problem of discontinuity. One way to circumvent this is to assume a continuous function, d_{ij}. For continuous weighting schemes, application of distance decay function or distance decline function is suggested which allows more weights for nearby units than areas that are far from one another (Brunsdon et al., 1996; LeSage, 2003; McMillen, 2003). An negative exponential function is suggested by Brunsdon et al. (1996) as follows:

$$w_{ij} = exp(-\eta d_{ij}^2),$$

where η is the spatial scale parameter which 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 distance-decay

[4]The diagonals of Multiplier Product Matrix 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.

varies between different concepts of distance, for different groups of economic units, and also for different estimation approaches.

Conditions that the spatial weights matrix W should satisfy that mirror the standard assumptions in this literature are:

Assumption A1 *W's 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's are uniformly bounded in both row and column sums in absolute value. Also $(I_N - \rho W)^{-1}$'s are uniformly bounded in $\rho \in \Lambda$.*

4 Estimation

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

Typically, production functions can be estimated by a variety of parametric, nonparametric, and semi-parametric techniques. A standard approaches to production function estimation is to estimate the average production technology, not the best-practice technology, which is accomplished in the stochastic frontier literature by not maintaining the assumption that all producers are cost or profit efficient. Little if any differences usually appear in the estimates of the basic production model parameters, such as output elasticities, etc. but the stochastic frontier analysis (SFA) approach does allow one to decompose the Solow-type residual into two components that make up TFP and its change over time, or TFP growth. 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 levels of inputs on average. Shifts in this average level of productivity over time, usually represented as a common trend using a time variable or a time index represents 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 typically involves the inclusion of a one-sided term (negative) that represents this short-fall in a firm's average production relative to a benchmark set by the most efficient firm. For parametric models, one-sided distributions such as half-normal, truncated normal, exponential, or gamma distribution are often used. Schmidt and Sickles (1984) and Cornwell et al. (1990) provide approaches to avoid the strong distributional assumptions by utilizing the structure of a panel production frontier. Schmidt and Sickles (1984) assume inefficiency to be time-invariant and unit-specific, while Cornwell et al. (1990) relax the time-invariant assumption by introducing a flexibly parametrized function of time replacing the individual fixed effects. Here, we follow Cornwell et al. (1990) for estimation, which allows us to estimate time-varying efficiency but does not require further distributional assumptions on the one-sided efficiency term.

In Cornwell et al. (1990), the non-spatial model (4) is estimated by three techniques: within transformation, generalized least squares, and efficient instrumental variable approach. The extended models (5) and (7), however, have some difficulties in estimation because they include additional spatially correlated variables. A quasi-maximum likelihood estimation (QMLE) is used in our analysis below. QMLE can provide robust standard errors against misspecification of the error distributions (Yang, 2013). The idea of QMLE is that we can minimize the number of parameters to estimate by using 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 from maximizing the full likelihood(LeSage and Pace, 2009). We will outline the estimation procedure briefly. Details can be found in the Appendix A.

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

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

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 to parameters except for the spatial autoregressive parameter ρ using the first order conditions of the likelihood functions of Eqs. (5) and (7). The spatial parameters, λ, are the coefficients for the spatially weighted independent variables. We treat the spatially weighted independent variables as additional regressors. Substitution of the closed-form solutions into the likelihood functions gives the concentrated likelihood functions that have ρ as the only unknown. $\hat{\rho}$ can be obtained by maximizing the concentrated likelihood functions, hence all other parameters can be found using $\hat{\rho}$. Details of derivation of the asymptotic distribution of the estimated parameters are in the Appendix A.

Once we obtain the estimates of the parameters β, ρ, δ_i and σ_v^2, we are able to recursively solve for an estimate of α_{it}, even though we cannot identify δ_0 and u_i, separately. Using the estimate of α_{it}, we can obtain the relative inefficiency measure following Schmidt and Sickles (1984) and Cornwell et al. (1990). Specifically, 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]

$$\hat{\alpha}_t = \max_j (\hat{\alpha}_{jt}), \tag{14}$$

$$\hat{u}_{it} = \hat{\alpha}_t - \hat{\alpha}_{it}. \tag{15}$$

5 An Empirical Application: Industry-level productivity of the U.S.

We examine the industry-level productivity of the U.S. in the period of 1947 - 2010 assuming classical Cobb-Douglas production technique.

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

5.1 Data

The databases employed for the empirical application are the US accounts of the World KLEMS database and World Input-Output database[6]. The usual difficulties of using data sets from different sources include maintaining the consistency of the data formation, such as industry classification levels or the units. Maintaining conformability of industry classification and the target period across the datasets is essential as the U.S. national accounts system, which provides the raw data, shifted to the North American Industry Classification System (NAICS) in 2003[7]. The World KLEMS data set covers data for the period 1947-2010 and uses the NAICS. The World Input-Output database is also based on NAICS, and it covers the period 1995-2011. Both databases are mapped into the 31 industry classifications of the International Standard Industrial Classification (ISIC) of All Economic Activities, Rev.3 for easy comparisons.[8]

Since the purpose of our paper is to examine industry level productivity growth in a context of interdependent spatial linkages among industries, we narrow our focus to a smaller sample, which display relatively strong similarities in their production processes, as we wish to minimize the extent to which unobserved differences in the production technology are interpreted as either spillovers or inefficiency. We acknowledge that this may alter somewhat the potential linkages among the remaining sectors but feel the trade-off is warranted. We selected the sample of sectors based on hierarchical cluster analysis that was based on sector specific average products of the inputs. Figure 3 illustrates a dendrogram of clusters based on Ward's minimum variance method(Ward, J. H., Jr., 1963). The dendrogram shows that the seven industries on the left are less similar to the other industries and also that dissimilarity level is high among those seven industries and we exclude them from the sample we analyze. The industry classifications we use are listed in table 2[9].

One of the advantages of the KLEMS data is that the variables are quality-adjusted using the Törnqvist index approach. A main feature of production data is that the variables often reflect heterogeneity due to quality differences in the components of the aggregated quantity indices. Coelli

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

[7]Also, the national accounts systems of each country are based on different international classification and it is hard work to keep concordance with each other.

[8]However, the two database have slightly different industry definitions. That is, *Textiles, Textile, Leather and Footwear* and *Transport and Storage* industries are more 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.

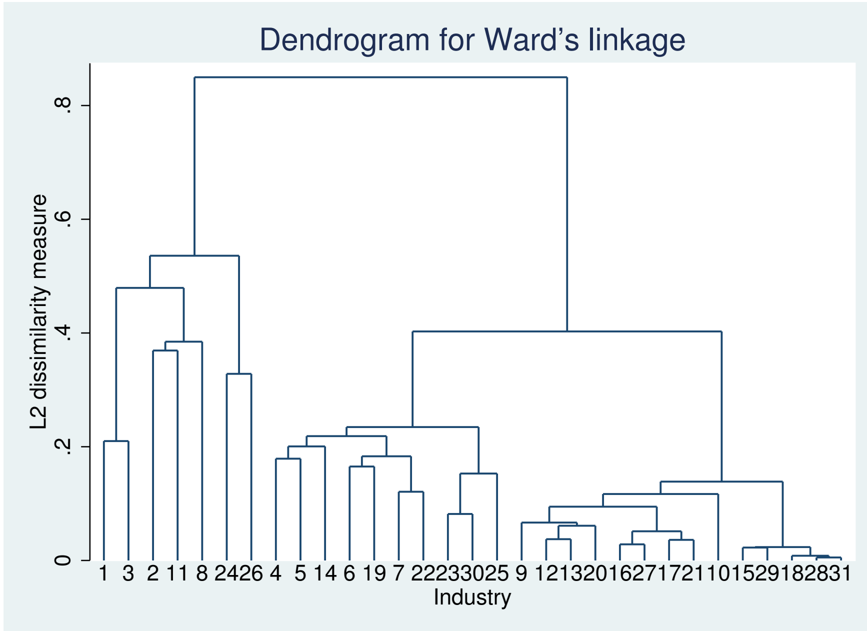


Figure 3: Cluster analysis on intermediate input to industry output ratio

et al. (2005) , among others, 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 using the approach of Jorgenson et al. (1987). It is assumed that aggregate labor services are a translog function of their individual components, 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 collected the output measures (gross output and value added) and input measures(capital and labor services, and intermediate inputs) formed in 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 than the within standard deviations in all real variables except for intermediate inputs, but the within standard deviations are larger than the between standard deviations in all index variables except for labor services. This implies that the quality-adjusted measures have somewhat different distributional shapes than the quality-unadjusted measures.

We use the Input-Output table for exploiting the patterns of intermediate inputs flows and

[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 on a variety of quality measures. iv) The model of Battese and Coelli (1995).

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

creating spatial weights matrices. Fig. 5 shows the linkages between industries. We find that industries with both strong backward and forward linkages have larger coefficients for the multiplier product matrix. Industries with weak linkages in both backward and forward, such as *Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Good*(Ind15) and *Real Estate Activities*(Ind19) show small MPM coefficients. *Public Admin and Defense; Compulsory Social Security*, and *Health and Social Work* have very weak linkages with other industries due to their almost complete lack of a Forward linkage.

5.2 Empirical findings

We assume a quadratic functional form for α_{it} following 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 interactive time dummy

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	$T = 64$

to control for different innovations in different industries. Additionally, the quadratic term of time in the efficiency term helps capturing any nonlinearities in efficiency changes during the relatively lengthy sample period. To control for possible endogeneity problem between the input factor levels and productivity, we also used 1-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

Estimation results are given 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

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

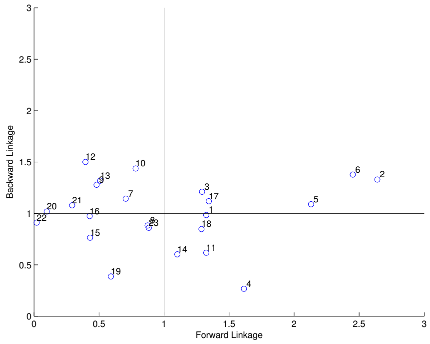


Figure 4: Forward and Backward linkage

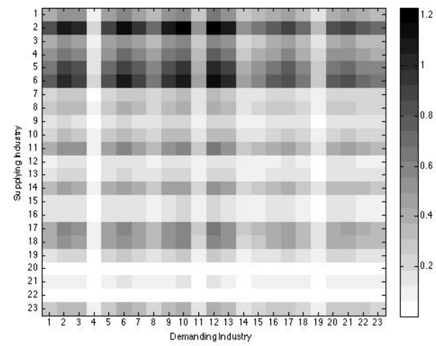


Figure 5: Multiplier product matrix

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	
$Adjusted R^2$	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	

interpreted as output elasticities in the models. The input with the biggest coefficient (elasticity) is labor while the coefficient of capital is the smallest. Returns to scale are about 0.87 for both models. We can estimate the parameters for $intercept$, $Time$, and $Time^2$ in CSSG model. The intercept term is negative but insignificantly different from zero. $Time$ is about 0.014, but $Time^2$ has no effects, which implies the growth rate of the economy is about 1.4% on average.

The next four pairs of columns show the results under the spatial specifications. The coefficients estimates no longer represent output elasticities when there are additional spatial dependent terms. 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 have the following forms:

$$\text{(SAR)} \qquad \frac{\partial y_t}{\partial X_{k,t}} = (I_N - \rho W)^{-1}(\beta_k I_N), \tag{16}$$

$$\text{(SDM)} \qquad \frac{\partial y_t}{\partial X_{k,t}} = (I_N - \rho W)^{-1}(\beta_k I_N + \lambda_k W). \tag{17}$$

The results are based on a time invariant spatial weights matrix. Since the output elasticities are given in a form of an $(N \times N)$ matrix, LeSage and Pace (2009) suggest that one takes averages along the diagonals to obtain the mean direct effects and along the column (or row) sums to obtain the mean indirect effects. Then the sum of the mean direct effect and the mean indirect effect can 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, we should calculate the total effects in addition to 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 given in Eqs. (16) and (17) since they are also functions of the estimates of ρ and λ. We follow algorithms that LeSage and Pace (2009) suggested. This 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% significance level except for the *intercept*. The spatial scale parameter η is estimated in a range of 1.158 to 1.505. The coefficient of the spatially lagged dependent variable, ρ, is estimated in a range of $0.374 to 0.492$. For the coefficients of spatially weighted independent variables, λ_1 is estimated positive while 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, while the coefficient estimates in the first three rows represent the direct effects. Comparing the results of SARCSS and SDMCSS with the results of non-spatial CSS, all output elasticities are estimated larger under spatial specifications. In particular, SARCSS estimates the largest elasticities out of the three specifications. Comparing SARCSS with SDMCSS, we found that the output elasticity of capital is larger, but the output elasticities of labor and intermediate inputs are estimated to be smaller with SDM. The discrepancy results in smaller returns to scale in SDMCSS. We can compare the goodness-of-fit of SARCSS and SDMCSS using

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 likelihood ratio test as SAR is nested in SDM. The LR test statistics are 15.732 and 15.784 for within-estimation and GLS-estimation, respectively, which implies that 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 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 direct effects sum to between 0.76 and 0.78, which implies possible decreasing scale economies. Returns to scale, however, increase up to 1.2 when the indirect effects are considered. The indirect effects are estimated about 37% of the total effects when only a spatially lagged dependent variable is included in the model. If we consider spatially lagged terms for both dependent and independent variables, the portions of the indirect effects vary. Interestingly, the indirect effects of capital service increases up to 67% while the indirect effects of labor and intermediate inputs decrease to 25% and 21%, respectively.

Glass et al. (2015) define internal, external, and total returns to scale with the direct, indirect, and total effects from a spatial production function of European countries and test if the returns to scale measures show constant returns to scale. In our analysis, the internal and external returns to scale do not show evidences of constant returns to scale by themselves, but they are large enough to make the total returns to scale to be 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 averaged efficiency score of each industry is illustrated in table 7. The total average efficiency scores are about 0.70 for gross output productivity. Even if the coefficient estimates in tables 4 and the efficiency scores show differences across the models, the efficiency ranking only varies marginally. The *Construction* sector (Ind12) turns out to be the most efficient and *Electrical and Optical Equipment* sector (Ind8) is the least efficient industry on average for all models. This result is at odds with what one would expects since the *Electrical and Optical Equipment* industry is one of the driving forces for economic growth in the total manufacturing sector of many countries. We check if the relative efficiency scores are reasonable by comparing with the total factor productivity(TFP) measure provided by the $KLEMS$ database, which is computed based on the growth accounting approach. TFP is computed industry-by-industry 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 $TFPs$, and we observe that *Construction* has the largest average TFP and *Electrical and Optical Equipment* has the smallest, which is consistent to our results[12].

Jorgenson et al. (2012) stress the performance of industries where innovation takes place as engines for long-run growth in an economy. They note the important role of IT-producing industries including software and hardware manufacturing industries as well as IT-service-producing industries. In particular, they found a substantial contribution of these industries in economic growth

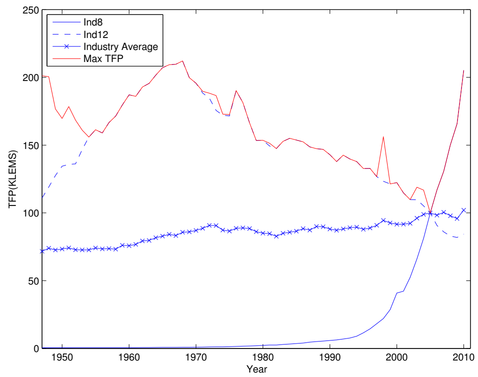
[12]Note that TFPs are calculated from TFP growth rates obtained by the growth accounting approach, which set the TFP level of the reference year to be 1. The growth accounting approach applies the methodology to a dataset from each industry independently. Hence, the comparison of TFP levels across industries 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

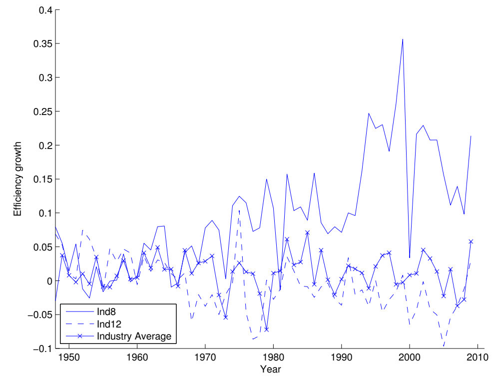
	non-spatial CSS				SAR				SDM				KLEMS	
	Within		GLS		Within		GLS		Within		GLS		Growth Accounting	
	Eff. Score	Rank	Eff. Score	Rank	Eff. Score	Rank	Eff. Score	Rank	Eff. Score	Rank	Eff. Score	Rank	Avg. TFP	Rank
Ind01	0.664	14	0.664	15	0.755	11	0.748	11	0.736	11	0.731	11	0.491	21
Ind02	0.789	8	0.793	7	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	

during the investment boom of 1995-2000. The sample period of 1947-2010 is long enough to have changes in the relative importance of different industries in the growth process. We thus examine the most and least efficient industries further by scrutinizing the variation of the relevant productivity and efficiency measures in fig. 6. We compare total factor productivity computed by the growth accounting approach and the efficiency scores estimated by the spatial stochastic frontier approach. We also create a relative measure of TFP defined by the relative size to the largest TFP of each year. The left panels of fig. 6 illustrate the evolution of the productivity measure of *Electrical and Optical Equipment*(Ind8), *Construction*(Ind12), and the industry average. The right panels are the growth rate of productivity or efficiency. Panels (a) and (b) are drawn from $KLEMS$ data, and panels (c) and (d) are the relative TFP generated for a comparison purpose. The solid red line of panel (a) indicates the maximum TFP of each year. It is almost identical to the solid blue line, which is a TFP series of *Construction* sector. Even if we observe that the TFP of the *Construction* industry falls after the late-60s, the industry is the most efficient in terms of total factor productivity. Meanwhile, the productivity of *Electrical and Optical Equipment* industry soars sharply after the mid-90s, in which so called IT investment boom has been started. Normalizing the TFP measure relative to the most productive industry, we can obtain the panels (c) and (d), which can be compared to the relative efficiency scores from the spatial stochastic frontier

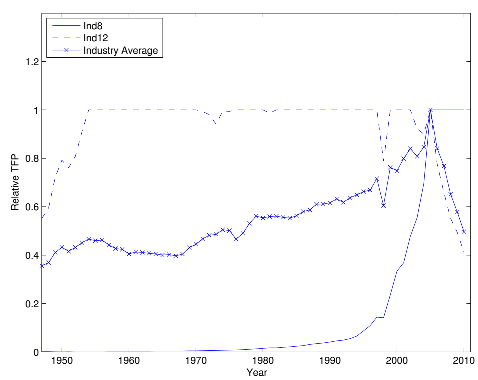
Figure 6: Efficiency and Efficiency Growth



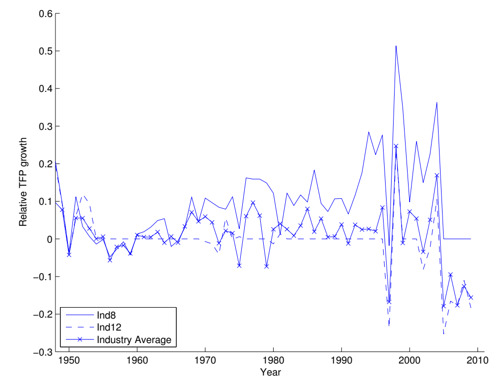
(a) TFP (KLEMS)



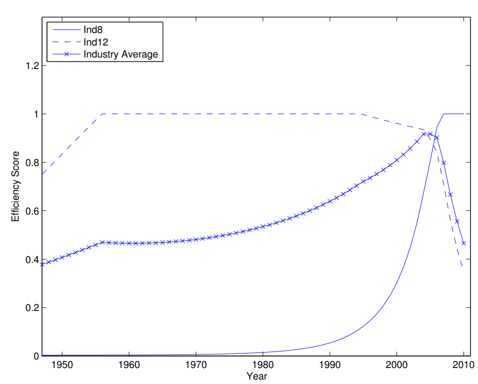
(b) TFP Growth



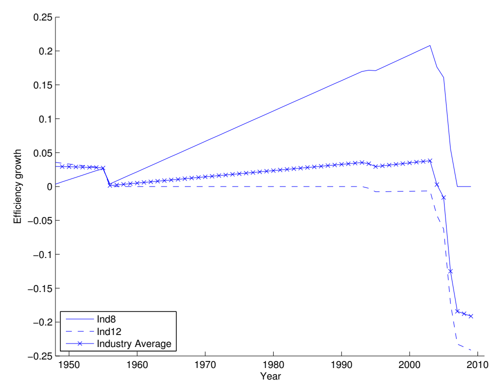
(c) Relative TFP



(d) Relative TFP Growth



(e) Relative Efficiency Score



(f) Relative Efficiency Growth

approach. Comparing the panels (c) and (e), the relative efficiency behaves quite similarly to the smoother lines for the spatial stochastic frontier approach. The variations in efficiency changes are averaged out in the regression-based approach. Likewise, the fluctuations in efficiency growth is

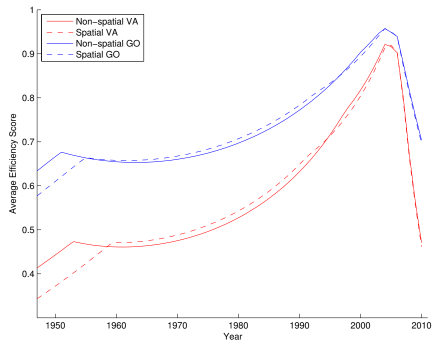


Figure 7: Efficiency Scores (Industry Average)

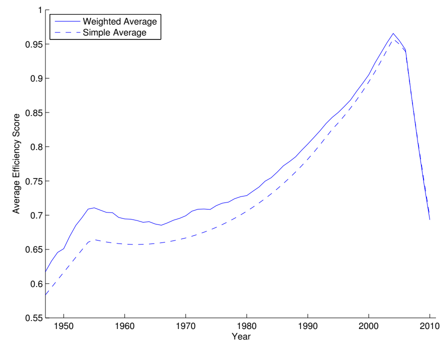


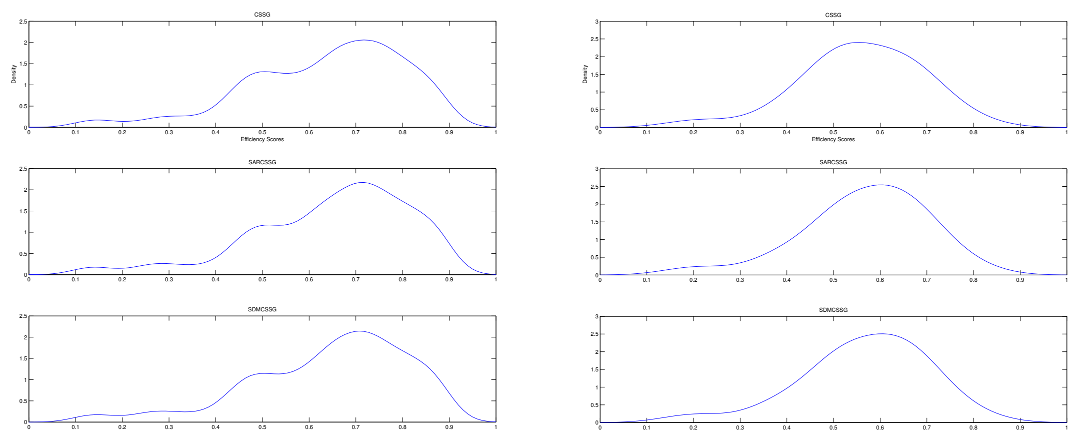
Figure 8: Weighted average of efficiency scores

not captured as well. However, the comparison with TFP measure from the growth accounting approach, which does not explicitly consider randomness, provide a benchmark that allows us to confirm that our methodology not only provides a more explicit role for spatial supply chain linkages but also provides summary measures of productivity growth that are consistent with those generated using established approaches of statistical agencies in the US and in other developed and developing countries.

We next examine how the industry-average efficiencies evolve over time in Fig. 7. The two local peaks are observed and the behavior of the average efficiency scores moves in similar for both spatial and non-spatial setting. Specifically, the global peak of the efficiency scores is observed in the 2004, but we discern that the size and significance of the small peaks in the 1950s depend on the models: the adoption of spatial specification and value-added productivity measure shift the peak to the right relative to their counterparts. We interpret that the U.S. industry-average relative efficiency converged to almost fully efficient status at around 2004 in both value-added and gross-output measures. However, the average relative efficiency score plummeted after the global peak. Actually, the relative efficiencies of all industries but *Electrical and Optical Equipment* sector fall after 2004. The implication of this is in line with the importance of IT-producing and IT-using industries stressed by Jorgenson et al. (2012). They define 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 high productivity growth in *Electrical and Optical Equipment*, the industry could not drive the productivity growth of total economy as its output ratio is only about 4% of total economy. Instead of the simple arithmetic average of efficiency scores, we draw the weighted average in fig. 8, following the weighting scheme used in Jorgenson et al. (2012). The solid line of fig. 8 shows the weighted average. The weighted average efficiency score is larger than the simple average (dotted line) but the differences decrease as time goes by, and the two averages are essentially the same after efficiency begins to decline.

Finally, we check the empirical density of the estimated efficiency scores. Fig. 9a and 9b compare the kernel density plots of the efficiency scores. The kernel density estimates in fig. 9a, drawn from the estimates of u_{it}, look coincident to the typically assumed one-sided inefficiency distribution, such as Half-normal, Truncated normal, Exponential, or Gamma. Meanwhile, the kernel densities of period-average efficiency scores for each industry in Fig. 9b are more symmetric, having more dispersed shapes.

Figure 9: Kernel density of efficiency scores



(a) Total efficiency scores

(b) Period-averaged efficiency scores

6 Conclusions

In this paper, we have examined how to measure industry-level productivity with cross-sectional dependence. We propose a method for choosing an appropriate weights matrix when there is no explicit distance concept. Specifically, we construct a unified measure characterizing the linkage between a pair of industries, which includes direct and indirect demand and supply side effects. An economic distance measure that is analogous to a geographic distance is defined by the relative size of our linkage measure to the measure of the most closely linked industry. We also have specified a spatial production model by expanding a traditional Cobb-Douglas production function with two basic spatial specifications: SAR and SDM. We estimate the model 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 take about 25% and 21% of total elasticities for SARCSS and SDMCSS, respectively. We could conclude that the U.S. economy has increasing returns to scale for the last six decades when only spatially weighted dependent variable is included in the model. However, the returns to scale is not significantly increasing if we additionally assume the factor inputs also show cross-sectional dependence.

Although the coefficient estimates vary across the model specifications, the relative efficiency scores are estimated comparably. Compared to the TFP measure from the growth accounting approach, our regression-based approach has the smoother efficiency estimates in both level and growth. For a noticeable result, we observe that *Electrical and Optical Equipment* is the least efficient industry on average over sixty years, even if it has a much more rapid growth in TFP than any other industry. On the other hand, the *Construction* sector is the most efficient industry on average in terms of its productivity level while it has slightly lower growth than 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.

Adetutu, M., A. J. Glass, K. Kenjegalieva, and R. C. Sickles (2015). The affects of efficiency and TFP growth on pollution in europe: A multistage spatial analysis. *Journal of Productivity Analysis 43*, 307–326.

Aigner, D., C. Lovell, and P. Schmidt (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics 6(1)*, 21–37.

Amsler, C., A. Prokhorov, and P. Schmidt (2014). Using copulas to model time dependence in stochastic frontier models. *Economic Reviews 33*(5-6), 497–522.

Anselin, L., J. Le Gallo, and H. Jayet (2006). Spatial panel econometrics. In L. Matyas and P. Sevestre (Eds.), *The econometrics of panel data, fundamentals and recent developments in theory and practice* (3 ed.)., pp. 901–969. Kluwer, Dordrecht.

Balk, B. M. (2009). On the relation between gross output and value added based productivity measures: The importance of the domar factor. *Macroeconomic Dynamics 13*, 241–267.

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.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2014). The impact of regional and sectoral productivity change on the U.S. economy. NBER Working Paper, 20168.

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.

Chudik, A., M. H. Pesaran, and E. Tosetti (2011). Weak and strong cross-section dependence and estimation of large panels. *Econometrics Journal 14*, 45–90.

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.

Denny, M., M. Fuss, C. Everson, and L. Waverman (1981). Estimating the effects of diffusion of technological innovations in telecommunications: The production structure of Bell Canada. *The Canadian Journal of Economics 14*(1), 24–43.

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.

Farrell, M. J. (1957). The measurement of productive efficiency. *Jornal of the Royal Statistical Society. Series A (General) 120*(3), 253–290.

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

Greene, W. (1980a). Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics 13*, 27–56.

Greene, W. (1980b). On the estimation of a flexible frontier production model. *Jornal of Econometrics 13*, 101–115.

Greene, W. (1982). Maximum likelihood estimation of stochastic frontier production models. *Journal of Econometrics 18*, 285–289.

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.

Horrace, W. C. (2005). Some results on the multivariate truncated normal distribution. *Journal of Multivariate Analysis 94*(1), 209–221.

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.

Kneller, R. and P. A. Stevens (2003). The specification of the aggregate production function in the presence of inefficiency. *Economics Letters 81*, 223–226.

Kumar, S. and R. Russell (2002). Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence. *American Economic Review 92(3)*, 527–548.

Kumbhakar, S. C. and H.-J. Wang (2005). Estimation of growth convergence using a stochastic production frontier approach. *Economics Letters 88*, 300–305.

Lee, L.-F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica 72*(6), 1899–1925.

LeSage, J. and R. Pace (2009). *Intoroduction to Spatial Econometrics*. CRC Press Taylor & Francis Group.

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

Levinsohn, J. and A. Petrin (2003). Estimating production function using inputs to control for unobservables. *The Review of Economic Studies 70*(2), 317–341.

Miller, R. E. and P. D. Blair (1985). *Input-Output Analysis: Foundations and Extentions*. Prentice Hall.

OECD (2001). *Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity*. Paris: Organisation for Economic Co-operation and Development.

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

Pires, J. O. and F. Garcia (2012). Productivity of nations: A stochastic frontier approach to TFP decomposition. *Economics Research International*.

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.

Prucha, I. R. and M. I. Nadiri (1986). A comparision of alternative methods for the estimation of dynamic factor demand mmodel under non-static expectations. *Journal of Econometrics 33*, 187–211.

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.

Stevenson, R. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics 13*, 58–66.

Timmer, M., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2014). The world Input-Output database: Contents, concepts and applications. University of Groningen.

Timmer, M., T. van Moergastel, E. Stuivenwold, G. Wpma, M. O'Mahony, and M. Kangasniemi (2007). EU KLEMS growth and productivity accounts version 1.0. University of Groningen.

Timmer, M. P., 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. P., 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. P. 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.

Van Biesebroeck, J. (2003, November). Revisiting Some Productivity Debates. NBER Working Papers 10065, National Bureau of Economic Research, Inc.

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

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:

$$logL(\psi, \delta_i; y) = -\frac{NT}{2} log(2\pi\sigma_v^2) + Tlog|I_N - \rho W| - \frac{1}{2\sigma_v^2} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(y_{it} - \rho \sum_{j=1}^{N} w_{ij} y_{jt} - X_{it}'\beta - Z_i'\gamma - R_t'\delta_i \right)^2, \quad (18)$$

The first order condition of maximizing Eq. (18) with respect to δ_i is

$$\frac{\partial logL}{\partial \delta_i} = \frac{1}{\sigma_v^2} \sum_{i=1}^{N} \sum_{t=1}^{T} R_t \left(y_{it} - \rho \sum_{j=1}^{N} w_{ij} y_{jt} - X_{it}'\beta - Z_i'\gamma - R_t'\delta_i \right) = 0. \quad (19)$$

By solving for (19), we can obtain

$$\hat{\delta}_i = (R_t R_t')^{-1} R_t \left(y_{it} - \rho \sum_{j=1}^{N} w_{ij} y_{jt} - X_{it}'\beta - Z_i'\gamma \right). \quad (20)$$

Substituting (20) into the log-likelihood function, (18), we obtain the concentrated likelihood function

$$logL(y; \beta, \rho, \sigma_v^2) = -\frac{NT}{2} log(2\pi\sigma_v^2) + Tlog|I_N - \rho W| - \frac{1}{2\sigma_v^2} \tilde{V}'\tilde{V}, \quad (21)$$

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

$$M_Q y = \rho^* M_Q (W_N \otimes I_T) y + M_Q X \beta + \tilde{V}. \quad (22)$$

[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.

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

$$\hat{\beta}_W(\rho^*) = (X'M_QX)^{-1}X'M_Q(y - \rho^*(W_N \otimes I_T)y), \tag{23}$$

$$\hat{\sigma}_v^2(\rho^*) = \frac{1}{N(T-L)-K}e(\rho^*)'^*), \tag{24}$$

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. (21), we can concentrate out β and σ_v^2, and the concentrated log-likelihood function with single parameter ρ is of the form:

$$logL(y;\rho) = C - \frac{NT}{2}log\left[e(\rho)'e(\rho)\right] + Tlog|I_N - \rho W|, \tag{25}$$

where C is a constant term that is not a function of ρ. By maximizing the concentrated log-likelihood function Eq.(25) 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. (23) and Eq. (24).

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

$$Asy.Var(\beta, \rho, \sigma_v^2)$$

$$= \begin{bmatrix} \frac{1}{\sigma_v^2}\tilde{X}'\tilde{X} & \frac{1}{\sigma_v^2}\tilde{X}'(W^* \otimes I_T)\tilde{X}\beta & \mathbf{0} \\ - & T \cdot tr(W^*W^* + W^{*\prime}W^*) + \frac{1}{\sigma_v^2}\beta'\tilde{X}'(W^{*\prime}W^* \otimes I_T)\tilde{X}\beta & \frac{T}{\sigma_v^2}tr(W^*) \\ - & - & \frac{NT}{2\sigma_v^4} \end{bmatrix}^{-1}, \tag{26}$$

where $\tilde{X} = M_QX$ 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:

$$\sigma_v\Omega^{-1/2}y = \rho\sigma_v\Omega^{-1/2}(W_N \otimes I_T)y + \sigma_v\Omega^{-1/2}X\beta + \sigma_v\Omega^{-1/2}\boldsymbol{Z}\gamma$$
$$+ \sigma_v\Omega^{-1/2}\boldsymbol{R}\delta_0 + \sigma_v\Omega^{-1/2}\varepsilon, \tag{27}$$

where $\varepsilon = QU + V$, $\Omega = cov(\varepsilon) = \sigma_v^2 I_{NT} + Q(I_N \otimes \Delta)Q'$. The estimation procedure of Eq. (27) 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

$$\begin{aligned}\hat{\eta}_G(\rho^*) &= [(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})]^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(y - \rho^*(W_N \otimes I_T)y) \\ &= [(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})]^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}y \\ &\quad - \rho^*[(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})]^{-1}(X, \boldsymbol{Z}, \boldsymbol{R})'^{-1}(W_N \otimes I_T)y.\end{aligned} \tag{28}$$

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:

$$\hat{\Omega}(\rho^*) = \hat{\sigma}_v^2 I_{NT} + Q(I_N \otimes \hat{\Delta}(\rho^*))Q'. \tag{29}$$

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

$$\hat{\Delta}(\rho^*) = \frac{1}{N}\sum_{i=1}^{N}\left[(R'^{-1}R'e_i e_i' R(R'^{-1} - \hat{\sigma}_v^2(R'^{-1}\right], \tag{30}$$

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

$$logL(\eta, \rho; y) = -\frac{NT}{2}log(2\pi\sigma_v^2) + Tlog|I_N - \rho W| - \frac{1}{2}\varepsilon'^{-1}\varepsilon, \tag{31}$$

where $\varepsilon = y - \rho(W_N \otimes I_T)y - X\beta - \boldsymbol{Z}\gamma - \boldsymbol{R}\delta_0$. Substitution of Eq. (28) and Eq. (29) into Eq. (31)

gives a concentrated likelihood function as follows:

$$logL(\rho; y) = C - \frac{NT}{2} log \left[e(\rho)' e(\rho) \right] + Tlog|I_N - \rho W|, \tag{32}$$

where $e(\rho) = \hat{\sigma}_v \hat{\Omega}(\rho)^{-1/2} y - \rho \hat{\sigma}_v \hat{\Omega}(\rho)^{-1/2} (W_N \otimes I_T) y - \hat{\sigma}_v \hat{\Omega}(\rho)^{-1/2} X \hat{\beta} - \hat{\sigma}_v \hat{\Omega}(\rho)^{-1/2} \boldsymbol{Z} \hat{\gamma} - \hat{\sigma}_v \hat{\Omega(\rho)}^{-1/2} \boldsymbol{R} \hat{\delta}_0$, 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. (32).

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. (23), 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. (22), $e(\rho^*)$, 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. (25) with respect to ρ after replacing $e(\rho) = e_0 - \rho e_1$, i.e.,

$$\max_{\rho} logL(y|\rho) = C - \frac{NT}{2} log \left[(e_0 - \rho e_1)'(e_0 - \rho e_1) \right] + Tlog|I_N - \rho W|. \tag{33}$$

Third, by replacing $\rho^* = \hat{\rho}$ in Eq. (23) and (24) 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. (26).

[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. (28) 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 (32). Once we have $\hat{\rho}$, Eq. (30), Eq. (29), Eq. (28) give $\Delta(\hat{\rho})$, $\Omega\hat{\rho}$, and $\eta_G(\hat{\rho})$ in order.

B Estimation results with Value-added dependent variable

Table 8: 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	
$AdjustedR^2$	0.148		0.176		0.215		0.236		0.258		0.272	
$loglikelihood$	-		-		576.111		542.115		587.847		554.150	
$Returnstoscale$	0.547		0.599		0.675		0.758		1.341		1.329	

Table 9: 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 10: 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