KIEP Working Paper 16-11

Korea Institute for International Economic Policy

A New Measure of Inter-industry Distance and Its Application to the U.S. Regional Growth

YOON Yeo Joon and WHANG Un Jung

KIEP Korea Institute for International Economic Policy

WORKING PAPER 16-11

A New Measure of Inter-industry Distance and Its Application to the U.S. Regional Growth

YOON Yeo Joon and WHANG Un Jung

KIEP Korea Institute for International Economic Policy

KOREA INSTITUTE FOR
INTERNATIONAL ECONOMIC POLICY (KIEP)

Building C, Sejong National Research Complex, 370,
Sicheong-daero, Sejong-si, Korea
Tel: 82-44-414-1251 Fax: 82-44-414-1144
URL: http://www.kiep.go.kr

HYUN Jung Taik, President

KIEP Working Paper 16-11
Published December 30, 2016 in Korea by KIEP
ISBN 978-89-322-4262-0 94320
978-89-322-4026-8(set)
Price USD 3

Executive Summary

We propose a new measure of inter-industry 'distance'. This is constructed à la Antras *et al.* (2012). While they measure the distance of an industry from its final use – what they call 'downstreamness' of an industry – we measure the distance between a pair of industries. Our proposed index is a measure of input-output linkages between industries that incorporates a 'distance' flavor. Our measure distinguishes the number of vertical production stages that an industry's product goes through until it is finally used by another industry by assigning larger weights to the value of input use with longer production chains. Hence our measure contains more information on the relation between two industries along the vertical production chain. We use this index to construct an aggregate measure of 'industry connectedness' of regions in the U.S. It measures the degree of industrial linkages of a region. We then empirically establish that each region's labor productivity is positively associated with the 'industry connectedness'. The result contributes to the large literature of agglomeration economies that the industrial linkage is one of the main sources of agglomeration economies and productivity growth, as emphasized by Marshall (1920). It also suggests that our index can serve as an alternative measure of the industrial linkages.

Keywords: Inter-industry Distance, Regional Growth, Input-Output Linkages
JEL Classification: F43, F63, O11

Contributors

YOON Yeo Joon

He is an Associate Research Fellow in Americas Team at the Korea Institute for International Economic Policy. His research interests mainly lie in American economic history with a particular emphasis on the history of its trade and related policies. He received his Ph.D. degree in Economics from the University of Warwick in U.K.

WHANG Unjung

Received his Ph.D. degree in Economics from the University of Colorado in Boulder, Colorado. Currently, he is an Associate Research Fellow in the Department of International Trade at the Korea Institute for International Economic Policy.
His research interests mainly include a model with both firms and product quality heterogeneity, structural changes, wage inequality, foreign direct investment, and global value chains. His work has been published in journals related to international economics, such as *Review of International Economics, World Economy, and Open Economies Review.*

Contents

Tables

Figures

1. Introduction

We propose a new measure of inter-industry 'distance'. This is constructed *à la* Antras et al. (2012). While they measure the distance of an industry *from its final use* - what they call 'downstreamness' of an industry - we measure the distance *between a pair of industries*. Our proposed index basically is a measure of input-output linkages between industries that incorporates a 'distance' flavor.

Backward and forward linkages among industries are important concepts in international trade, industrial organization, economic geography and many other fields in economics. Understanding how intensively outputs produced by a particular industry is used as intermediate inputs to other industries is crucial for analyzing the global value chain, the agglomeration of industries and the regional economic development.

In general there are four ways to look at the linkage relations between a pair of industries, say industry i and j.

(1) the relative importance of input from i in the total input to produce j (j's point of view)

(2) the relative importance of i's output as j 's input in i's total output (i's point of view)

(3) the relative importance of input from j in the total input to produce i (i's point of view)

(4) the relative importance of j's output as i's input in j's total output (j's point of view)

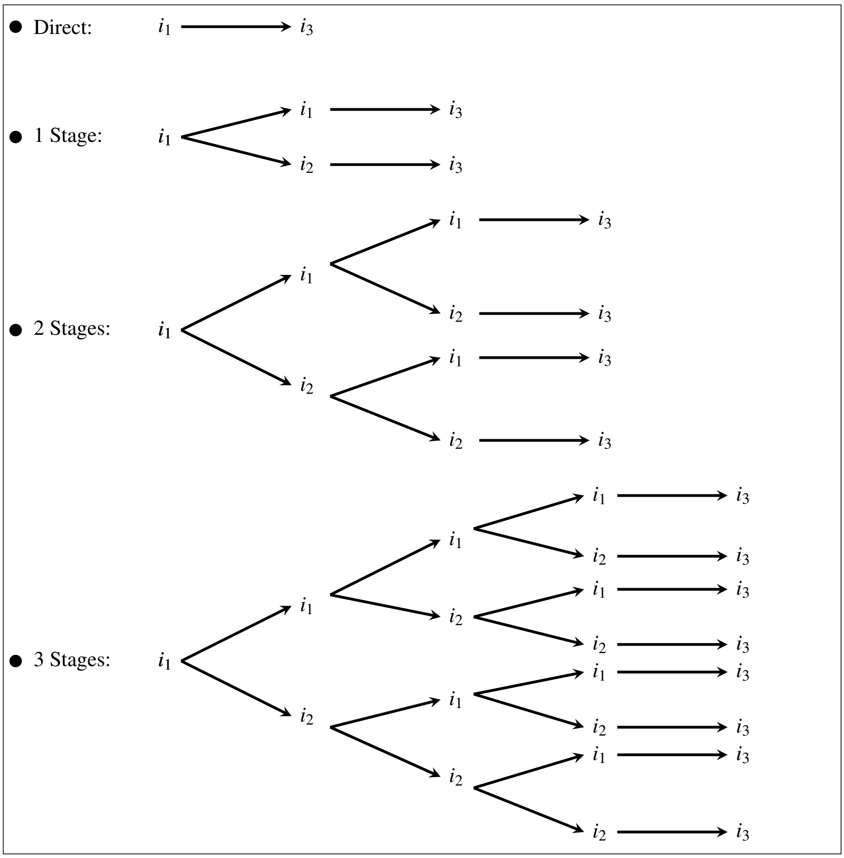
We call (1) and (3) the upstream measure from j to i and the upstream measure from i to j, respectively. They are *upstream* in that they measure the importance of its upstream industry from a downstream industry's point of view. Likewise we call (2) and (4) the downstream measure from i to j and the downstream measure from j to i, respectively which measure the importance of its downstream industry from an upstream industry's perspectives. According to this classification our measure is related to (2) and (4).

One can construct these measures with basic information provided by the input-output (I-O) table. The direct requirement table in the U.S. I-O table provides the share of industry i's input *directly* required to produce one unit of output in industry j. But this only considers the direct channel. The total requirement table makes up this shortcoming by considering the indirect channels as well. A ith-row, jth-column element in the total requirement table denotes the total amount of i required to produce one unit of j's final output. Using the total requirement table one can construct the downstream measures ((1) and (3)) that takes the indirect channels into account[1]. But it does not distinguish, within the indirect channels, whether the relation

[1] See Rasmussen (1956), Alfaro and Charlton (2009) and di Giovanni and Levchenko (2010).

between industry *i* and *j* is characterized by a long vertical production chain versus a short chain. Our measure, on the other hand, distinguishes this by the number of vertical production stages *i*'s product goes through until it is finally used by *j* and assign larger weights to the value of input use with longer production chains. In this sense our measure has a flavor of distance and contains more information on the relation between two industries along the vertical production chain.

Figure 1. Illustrative Example of Input-output Linkages



For illustrative purpose assume that there are only 3 industries in the economy and think of the channels that an output from industry 1 goes through until it is used by industry 3 as an intermediate input. As illustrated in Figure 1, product from industry 1 can be directly consumed by industry 3 as input or it can be used after being embedded in products from other industries. For example industry-1-product can be embedded in industry-2-product and then used as an intermediate input for industry-3-product in which case it goes through a single stage until it is used by industry 3. It can also go through multiple number of stages to reach industry 3 as Figure 1 shows. Theoretically it can go through infinite number of stages. Our index considers these indirect channels and distinguishes them by assigning different weights to each stage.

We use this index to construct an aggregate measure of 'industry connectedness' of regions in the U.S. It measures the degree of industrial linkages of a region. We then empirically establish that each region's labor productivity is positively associated with the 'industry connectedness'. The result contributes to the large literature of agglomeration economies that the industrial linkage is one of the main sources of agglomeration economies and productivity growth, as emphasized by Marshall (1920).[2] It also suggests that our index can serve as an alternative measure of the industrial linkages. The remainder of the paper proceeds as follows. In Section 2 we describe how we construct the measure and the natures of our measure more in details. In Section 3 we perform the empirical analysis using our measure. In Section 4 we conclude the paper.

2. Measure of Inter-Industry Distance

The I-O table can be used for analyzing how industries are connected to each other through intermediate uses. The input-output linkages across industries can be expressed in vector-matrix notation, which is as follows.

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad F = \begin{bmatrix} F_1 \\ \vdots \\ F_n \end{bmatrix},$$

where Y_i is the gross output of sector i. a_{ij} is the sector i's output consumed by sector j as intermediate inputs (that is, the inter-industry intermediate sales by sector i to sector j). F_i is the gross output of sector i that is consumed as final products.

We begin by considering an n industries input-output structure with no investment or inventories, where exports are recorded as final uses. For each industry $i \in \{1, 2, ..., n\}$, the value of total input, Z_i, equals

[2] For more on the empirics of agglomeration economies please see Combes and Gobillon (2014) and Rosenthal and Strange (2004) which provide nice surveys on the topic.

the sum of its use as intermediate inputs in the production of other industries and itself, which is shown in equation (1) below.

$$Z_i = \sum_{k=1}^{n} a_{ik}, \tag{1}$$

where Z_i is the industry i's output that is only used as intermediates, so that the industry i's total output equals the sum of Z_i and a final consumption in i, F_i. Note that $Y_i = Z_i + F_i$.

Now we introduce the industry-pair input structure that describes how output from one sector can be used as an input in the production of another sector. In fact, two distinct channels should be considered to fully understand the picture of the linkage between two industries, that is, the direct and the indirect channel.

Let's consider how upstream industry i's output becomes an input to downstream industry j. First, the output from industry i can directly become an intermediate input in the production of industry j (direct channel). Second, the industry i's output may become an input to industry j indirectly through other industries (including industry i itself), which are eventually used as an input to industry j (indirect channel). For example, i's output first is used to produce goods classified in industry k, then k's output is used as an intermediate input in the production of the destination industry j. Theoretically, i's output can be indirectly used as an input to j through infinitely many vertical production stages. Given the importance of the indirect channel, one should take into account both channels in order to fully capture input flows between two industries.

The input flow from industry i to j, denoted by Z_i^j, is decomposed into three main components, which can be written as

$$\begin{aligned}
Z_i^j &= Z_i = \sum_{k=1}^{n} a_{ik} \\
&= \overbrace{a_{ij}}^{\text{Direct Channel}} + \cdots + a_{ik}\overbrace{\left(\frac{a_{k1}+a_{k2}+\cdots+a_{kn}+F_k}{Z_k+F_k}\right)}^{=1} + \cdots + a_{in}\overbrace{\left(\frac{a_{n1}+a_{n2}+\cdots+a_{nn}+F_n}{Z_n+F_n}\right)}^{=1} \\
&= a_{ij} + \overbrace{\sum_{\substack{k=1\\k\neq j}}^{n}\frac{a_{ik}a_{kj}}{(Z_k+F_k)} + \sum_{\substack{l=1\\l\neq j}}^{n}\sum_{\substack{k=1\\k\neq j}}^{n}\frac{a_{ik}a_{kl}a_{lj}}{(Z_k+F_K)(Z_l+F_l)} + \sum_{\substack{m=1\\m\neq j}}^{n}\sum_{\substack{l=1\\l\neq j}}^{n}\sum_{\substack{k=1\\k\neq j}}^{n}\frac{a_{ik}a_{kl}a_{lm}a_{mj}}{(Z_k+F_K)(Z_l+F_l)(Z_m+F_m)} + \cdots}^{V_i^j=\text{Indirect Channel}} \\
&\quad + \underbrace{\sum_{\substack{k=1\\k\neq j}}^{n}\frac{a_{ik}F_k}{Z_k+F_k} + \sum_{\substack{l=1\\l\neq j}}^{n}\sum_{\substack{k=1\\k\neq j}}^{n}\frac{a_{ik}a_{kl}F_l}{(Z_k+F_K)(Z_l+F_l)} + \sum_{\substack{m=1\\m\neq j}}^{n}\sum_{\substack{l=1\\l\neq j}}^{n}\sum_{\substack{k=1\\k\neq j}}^{n}\frac{a_{ik}a_{kl}a_{lm}F_m}{(Z_k+F_K)(Z_l+F_l)(Z_m+F_m)} + \cdots}_{U_i^j=\text{Final Consumption}},
\end{aligned} \tag{2}$$

where F_k denotes the final consumption in industry k. It should be noted that Z_i^j is equal to Z_i in equation (1). The reason that we add the superscript j to Z_i is for the purpose of illuminating that equation (2) is intended to show a linkage from upstream industry i to downstream industry j.

The equation (2) consists of three components. First, the value of input, that directly flows from industry i to j, is captured by a_{ij} in equation (2). Second, the indirect channel through which industry i's output can become an input in the production of industry j is seen as V_i^j in equation (2). The indirect channel, V_i^j, can be expressed as an infinite sequence of terms, where each term reflects the number of production stages that industry i's output goes through before it ends up as an input in the production of industry j. To be more specific, the sector i's output can be used as an input to destination industry j after being used as inputs to other industries (including industry i itself) except for industry j, and this indirect procedure could occur through many stages. For example, $a_{ik}a_{kj}$ denotes the case that the industry i's output is first used as an input to industry k and then the industry k's output becomes an input in the production of industry j. The term, $\frac{a_{kj}}{(Z_k+F_K)}$, in equation (2) is the ratio of industry k's output that is used as an input to industry j to industry k's total output (that is, the remaining proportion of industry k's output after being used as both an input to other industries except for industry j and final consumption). Third, the industry i's output may be consumed as a form of final goods after being used as inputs in the production of other industries, which is captured by U_i^j term in equation (2). Now equation (2) can be rewritten as

$$\begin{aligned} Z_i^j - U_i^j =& a_{ij} + \sum_{k\neq j}^{n} \frac{a_{ik}a_{kj}}{(Z_k+F_k)} + \sum_{l\neq j}^{n}\sum_{k\neq j}^{n} \frac{a_{ik}a_{kl}a_{lj}}{(Z_k+F_K)(Z_l+F_l)} \\ &+ \sum_{m\neq j}^{n}\sum_{l\neq j}^{n}\sum_{k\neq j}^{n} \frac{a_{ik}a_{kl}a_{lm}a_{mj}}{(Z_k+F_K)(Z_l+F_l)(Z_m+F_m)} + \cdots . \end{aligned} \tag{3}$$

The right hand side of equation (3) represents the overall value of industry i' output that is used as an intermediate input in the production of j. Dividing both sides of equation (3) by $(Z_i^j - U_i^j)$ yields

$$\begin{aligned} 1 =& \frac{a_{ij}}{(Z_i^j - U_i^j)} + \sum_{k\neq j}^{n} \frac{a_{ik}a_{kj}}{(Z_i^j - U_i^j)(Z_k+F_k)} + \sum_{l\neq j}^{n}\sum_{k\neq j}^{n} \frac{a_{ik}a_{kl}a_{lj}}{(Z_i^j - U_i^j)(Z_k+F_K)(Z_l+F_l)} \\ &+ \sum_{m\neq j}^{n}\sum_{l\neq j}^{n}\sum_{k\neq j}^{n} \frac{a_{ik}a_{kl}a_{lm}a_{mj}}{(Z_i^j - U_i^j)(Z_k+F_K)(Z_l+F_l)(Z_m+F_m)} + \cdots . \end{aligned} \tag{4}$$

Note that each term of the right hand side of equation (4) reflects the percentage (or weight) of each channel that varies with the number of production stages needed for i's output to be an input to j. The first term, $a_{ij}/(Z_i^j - U_i^j)$, captures the proportion of the "direct channel" through which i's output is directly used as an input to industry j. The remaining terms reflect the proportion of the indirect channels through which the industry i's output is indirectly used as an input to the destination industry j after being used as an input to other industries. The third term, for example, indicates the proportion of the indirect channel where i's output needs to visit two other sectors before it ends up as an input in the production of industry j (i.e., $i \to k \to l \to j$).

As suggested by Antrás and Chor (2013), the weighted average inter-industry distance, that is a measure of the extent to which two industries are related by input-output linkage, can be obtained by assigning weights to each term of the right hand side of equation (4). To be more specific, the first term on the right hand side of equation (4), $a_{ij}/(Z_i^j - U_i^j)$, is weighted by 1. The industry i's output used indirectly in the production of industry j after being used as inputs to other industries is weighted by 2, 3, 4 and so on. That is, each weight (each term of right hand side of equation (4)) is multiplied by the distance between two industries i and j. Finally, we obtain the inter-industry distance, D_i^j, which is as follow.

$$\begin{aligned} D_i^j =& 1 \cdot \frac{a_{ij}}{(Z_i^j - U_i^j)} + 2 \cdot \sum_{k \neq j}^{n} \frac{a_{ik}a_{kj}}{(Z_i^j - U_i^j)(Z_k + F_k)} + 3 \cdot \sum_{l \neq j}^{n} \sum_{k \neq j}^{n} \frac{a_{ik}a_{kl}a_{lj}}{(Z_i^j - U_i^j)(Z_k + F_K)(Z_l + F_l)} \\ &+ 4 \cdot \sum_{m \neq j}^{n} \sum_{l \neq j}^{n} \sum_{k \neq j}^{n} \frac{a_{ik}a_{kl}a_{lm}a_{mj}}{(Z_i^j - U_i^j)(Z_k + F_K)(Z_l + F_l)(Z_m + F_m)} + \cdots . \end{aligned} \quad (5)$$

It is clear that $D_i^j \geq 1$, and $D_i^j = 1$ means that all industry i's output remaining after being consumed as final products (i.e., $Y_i - F_i = Z_i$) is used as an input in the production of industry j. Note that a larger value of D_i^j is associated with a relatively lower level of the inter-industry connectedness between industry i and j.[3] Theoretically, the maximum value of inter-industry distance, D_i^j, is infinite, but in practice, the largest degree of the inter-industry distance takes a finite value. Using the 2007 U.S. input-output table, the value of the inter-industry distance is distributed between 1.00 and 9.47. For example, the value of the inter-industry distance between 'Household laundry equipment manufacturing' and 'Dry-cleaning and laundry services' is close to one, which indicates that 'Household laundry equipment manufacturing' products are mostly used as an input of 'Dry-cleaning and laundry services' through either a direct or an indirect channel. On the other hand, the inter-industry distance from 'Ophthalmic goods manufacturing' to 'Aircraft engine and

[3] To be more specific, D_i^j is the downstream distance (or downstream closeness) from i to j and D_j^i - the downstream distance from j to i - also defines the relation between i and j.

engine parts manufacturing' is about 9.47, which implies that these two industries are almost not connected by the input-output linkages.

Inter-industry Distance using the U.S. Input-Output Table

Here, we construct the measure of inter-industry distance using the U.S. Input-Output (IO) Table provided by the BEA. To do this, we employ the Use Table from the year 2007 that contains the input-output flows of 389 industries at producer prices.[4] The Use Table defines the dollar value of each commodity that is purchased by each industry. Thus, the (i, j)-th entry of the Use Table represents the dollar value of commodity i used in the production of industry j. Since some industries (4 industries) appear as industries (outputs) at the detailed levels but not as commodities (inputs), these industries were removed from the Use Table. Additionally, some industries (44 industries) are removed from the Use Table because outputs from these industries are never used as intermediates in the production of other industries. Then we are left with 341 industries. Equation (5) allows us to calculate the industry-pair distances D_i^j.[5]

Table 1 illustrates the inter-industry distances of three specific upstream industries: Metal ore mining, Motor vehicle body, and Semiconductor. For each upstream industry, the top ten downstream industries according to our measures are shown in Table 1. For example, the inter-industry distance from 'Iron, gold, silver, and other metal ore mining (2122A0)' to 'Primary smelting and refining of non-ferrous metal (331419)' is close to one, and it is ranked first in the 340 industries. This means that these two industries are closely connected by the inter-industry linkages: the output of 'Iron, gold, silver, and other metal ore mining' is considered as an important input in the production of 'Primary smelting and refining of non-ferrous metal'. Table 1 also shows the basic statistics of each upstream industry, such as the mean value of inter-industry distance. the mean value of 'Semiconductor and related device manufacturing' is 3.32, while the average inter-industry distance of 'Motor vehicle body manufacturing' is 7.03. It should be noted that 'Semiconductor and related device manufacturing' is among the most upstream industries, with almost all of their output going directly or indirectly to the production of other downstream industries as an intermediate input.

The number in parenthesis on ranking variable indicates the ranking of inter-industry distance that is measured by considering only the direct linkage from the input-output table. For instance, 'Motor vehicle body manufacturing' is never used directly as an intermediate inputs in the production of 'Iron, gold, silver,

[4]The 2007 Use Table were released on November, 2014, and it is the most recent data that provides the 389-industry level of aggregation (i.e., six-digit IO industry codes).

[5]The solution mechanism of Equation (5) is available upon request from the authors.

Table 1. Inter-industry Distance for Three Specific Industries

Downstream Industry Description (j)	Code	Ranking	D_i^j
Iron, gold, silver, and other metal ore mining (i=2122A0)			
Primary smelting and refining of nonferrous metal	331419	1 (2)	1.11
Iron and steel mills and ferroalloy manufacturing	331110	2 (1)	1.20
Alumina refining and primary aluminum production	33131A	3 (6)	1.24
Abrasive product manufacturing	327910	4 (16)	1.27
Synthetic dye and pigment manufacturing	325130	5 (7)	1.30
Other basic inorganic chemical manufacturing	325180	6 (4)	1.42
Industrial gas manufacturing	325120	7 (11)	1.45
Nonferrous metal rolling, extruding and alloying	331490	8 (3)	1.53
Fertilizer manufacturing	325310	9 (10)	1.61
Jewelry and silverware manufacturing	339910	10 (9)	1.62
Motor vehicle body manufacturing (i=336211)			
Downstream Industry Description (j)	Code	Ranking	D_i^j
Truck trailer manufacturing	336212	1 (4)	1.51
Light truck and utility vehicle manufacturing	336112	2 (1)	1.58
Automotive repair and maintenance	811100	3 (6)	2.00
Heavy duty truck manufacturing	336120	4 (7)	2.08
Construction machinery manufacturing	333120	5 (-)	2.46
Iron, gold, silver, and other metal ore mining	2122A0	6 (-)	3.76
Other nonmetallic mineral mining and quarrying	2123A0	7 (-)	3.87
Stone mining and quarrying	212310	8 (-)	3.88
Coal mining	212100	9 (-)	3.90
Other support activities for mining	21311A	10 (-)	3.91
Semiconductor and related device manufacturing (i=333295)			
Downstream Industry Description (j)	Code	Ranking	D_i^j
Printed circuit assembly (electronic) manufacturing	334418	1 (-)	1.16
Computer storage device manufacturing	334112	2 (-)	1.21
Wireless telecommunications carriers	517210	3 (-)	1.26
Photographic equipment manufacturing	333315	4 (-)	1.36
Motor vehicle electronic equipment manufacturing	336320	5 (-)	1.40
Electronic computer manufacturing	334111	6 (-)	1.46
Telephone apparatus manufacturing	334210	7 (-)	1.54
Broadcast and wireless communications equipment	334220	8 (-)	1.65
Electronic equipment repair and maintenance	811200	9 (-)	1.65
Office supplies (except paper) manufacturing	339940	10 (-)	1.70
Upstream Industry Description (i)	Min	Max	Mean
Iron, gold, silver, and other metal ore mining	1.11	7.66	4.48
Motor vehicle body manufacturing	1.51	8.77	7.03
Semiconductor and related device manufacturing	1.16	5.66	3.32

Notes: The number in parenthesis on ranking variable indicates the ranking of inter-industry closeness that is measured by a direct linkage.

and other meta ore mining', while it is used indirectly in the production of 'Iron, gold, silver, and other meta ore mining' through truck manufacturing.

3. The Empirical Analysis

In this section, using our measure, we construct aggregate measures of 'industry connectedness' for each region in the U.S. Regions classified according to the Combined Statistical Area (CSA). 'Industry connectedness' measures the degree of industrial linkages of a region. In essence if there is a high concentration of industries that are closely linked in terms of input-output linkages, then the region has a high degree of 'industry connectedness'.[6] Using this measure we then empirically test the theory of Marshall (1920). He argued that there are three factors that cause agglomeration of economic activities within a region: knowledge spillovers, labor market pooling and input-output linkages among industries. Our measure is related to the third factor, the I-O linkage. It is also well established that the agglomeration of economic activities increases productivity (Ciccone and Hall (1996); Ciccone (2002)). We regress our measure of 'industry connectedness' on manufacturing labor productivity of each region in the U.S. Rigby and Essletzbichler (2002), Lopez and Sudekum (2009) and Ellison et al. (2010) employ similar approaches, using I-O tables, to test the Marshallian theory of agglomeration.

Since the industry-level variables including labor productivity from the U.S. regional data are classified by the 3-digit NAICS code, we need to collapse the 6-digit IO industries into the 3-digit NAICS industries in order to construct the input-output table corresponding to the 3-digit NAICS industries. To do this, we first match the IO industries to the NAICS industries using the corresponding tables between the 2007 6-digit IO industry codes and the 2007 3-digit NAICS codes provided by the BEA.[7] Second, we construct a new Input-Output Table corresponding to the 3-digit NAICS industries by collapsing the 6-digit IO industries into the 3-digit NAICS industries. Thus, we have the new Input-Output table corresponding to 78 NAICS industries, so that we have total 78×77 measures of industry-pair distance. Table 4 in Appendix lists the inter-industry distance of 21 NAICS manufacturing industries.

[6] Throughout this paper, we use the terms 'industry connectedness', 'downstream closeness' and 'inter-industry connectedness (closeness)', analogously.

[7] In some cases, a single 6-digit IO industry is comprised of multiple 3-digit NAICS codes. For instance, the 'Wholesale Trade', classified by 420000 IO industry code, corresponds to three different 3-digit NAICS codes, 'Merchant Wholesalers (durable goods)', 'Merchant Wholesalers (nondurable goods)', 'Wholesale Electronic Markets and Agents and Brokers', which are classified by 423, 424, 425 NAICS codes, respectively. To deal with this issue, we use the 2-digit NAICS industry classification only for these cases, so that a single 6-digit IO industry corresponds to one 2-digit or 3-digit NAICS industry. To be more specific, we have the only 4 IO industries that correspond the 2-digit NAICS, and the other IO industries exactly match the 3-digit NAICS industries.

3.1. *Model Specification*

In this section we describe our empirical methods for testing Marshallian theory of agglomeration. Essentially we are interested in regional differences in labor productivity (i.e., value added per hour worked) which are potentially related to the degree of industry agglomeration. Thus, we first construct a variable that measures how industries in each region are inter-connected in terms of downstreamness. We then use them as explanatory variables to see the impact of 'inter-industry connectedness' on labor productivity of regions.

'Inter-industry connectedness' of industry i in region r is measured as

$$DM_{i,r} = \sum_{s \neq i}^{n} (D_i^s)^{-1} \left(\frac{L_{s,r}}{L_r - L_{i,r}} \right). \tag{6}$$

$L_{s,r}$ is labor in region r employed in industry s and L_r is the total employment in region r. Equation (6) basically says that 'downstream closeness' of industry i in region r is the proportion of all other industries weighted by our measure of inter-industry distance. A large $DM_{i,r}$ implies that from industry i's point of view there is a large presence of downstream industries in region r that are *close* to industry i.

It should be noted that since the data including labor productivity is available only at the regional level, but not at the industry-regional level, we additionally construct the regional level of downstream closeness.[8] Now we construct aggregate measures of industry connectedness of region r from the equation (6), which is as follows.

$$DM_r = \sum_{s}^{n} DM_{s,r} \left(\frac{L_{s,r}}{L_r} \right). \tag{7}$$

Again a large $DM_{s,r}$ means that for industry s there are relatively large proportions of downstream industries that are closely linked to s. In what follows a large DM_r means that *on average* for an industry in region r there are relatively large proportions of downstream industries that are close to the industry within the same region. Equation (7) is the measures of what we termed 'inter-industry connectedness' or 'downstream closeness' of each region in the introduction.

[8] In fact, the analysis using the industry-regional level data may give more implications about the degree to which the downstream closeness affects labor productivity.

The degree to which the inter-industry linkages affects labor productivity is estimated using a translog regression, where the natural log of labor productivity is regressed against the natural log of 'downstream closeness' and control variables. Equation that we estimate is

$$ln(LP_r) = \alpha + \beta ln(DM_r) + X_r^{'}\gamma + \varepsilon_r, \tag{8}$$

where $ln(LP_r)$ is the logarithm of the labor productivity for region r and X_r is a vector of region controls, such as capital intensity and education. ε_r is a random disturbance. We employ cross sectional data for year 2012 when the most recent Census data are available.

3.2. Data Description

The U.S. regions are divided according to *Combined Statistical Area* (CSA).[9] We chose CSA as our regional unit because it is composed of smaller geographic units with high degree of economic integration. As far as we know it is the most ideal geographical unit available for our purpose as it combines regions with strong economic ties. In 2012 there is a total of 166 CSA regions in the U.S.

The dependent variable, labor productivity, is measured as value added per production hours worked. We include the capital intensity and the share of high school and college graduates in working age population for each region as other control variables. The capital intensity is measured as the ratio of capital expenditures to labor compensations. We obtain the labor productivity and the capital intensity from *2012 Economic Census of the United States* and the educational data from *American Community Survey* of *the Census*. With regards to the region-industry employment ($L_{s,r}$), because of limited data availability we had to resort to different means. *County Business Patterns* of *the Census* reports the number of establishments by the size of employees and industry for each core based statistical area. For example in 2012, there are 39 establishments in 'construction and buildings' sector (NAICS 236) in Aberdeen (SD Micro Area) that employ 1 to 4 employees. Because the number of establishments are reported by the range of employment size we take the median value of each range and multiply this by the number of establishments that falls into that category. We then add them across all the ranges to obtain $L_{s,r}$[10]

[9]Two or more adjacent *Core Based Statistical Areas* (CBSA) form a CSA if there is a substantial degree of employment interchanges between the *Core Based Statistical Areas.* According to U.S. Office of Management and Budget CBSA is defined as *a statistical geographic entity consisting of the county or counties associated with at elast one core (urbanized area or urban cluster) of at least 10,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties with the counties containing the core.*

[10]It classifies them by 9 employment size groups: establishments with 1 to 4, 5 to 9, 10 to 19, 20 to 49, 50 to 99, 100 to 249, 250

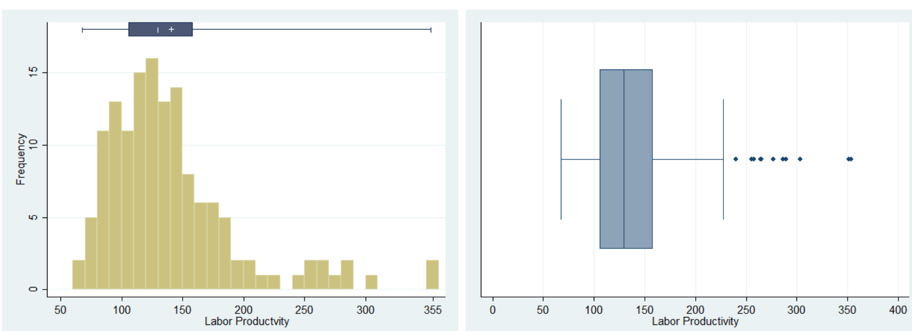
Table 2. Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
LP	142	140.742	54.173	67.600	353.300
DM	160	0.404	0.013	0.374	0.428
CI	154	0.289	0.246	0.040	2.770
EDU	158	26.376	6.811	13.200	50.800

Notes: LP stands for the labor productivity. DM is a measure of downstream closeness. CI is the level of the capital intensity, that is calculated as capital expenditure divided by total labor compensation. EDU represents the percentage of the population 25 years and older who had a bachelor's degree or higher.

Table 2 provides the descriptive statistics of the data. Note that the range (i.e., maximum minus minimum) of labor productivity is large (286) due to the outliers. In fact, labor productivity is distributed with a mean of 127 and a standard deviation of 32 when excluding the outliers, and its range between the first and third quartile (that is, interquartile range) is relative small (140), which as illustrated in Figure 2.

Figure 2. Histogram and Box Plot for Labor Productivity



(a) Histgram with Box Plot (b) Box Plot

to 499, 500 to 999 employees and 1000 employees or more.

3.3. Empirical Results

The estimation results, summarized in Table 3, indicate that downstream closeness at the regional level are important determinants of regional labor productivity. Each column in Table 3 reports the estimation results obtained under the ordinary least squares (OLS) regression with log transformed variables.[11] The first three columns (i.e., column (1)-(3)) use the full sample, while the rest of the columns (i.e., column (3)-(6)) report the estimation results obtained by using the restricted sample that excludes the outliers of labor productivity variable. Note that the outliers of labor productivity, that are unusual observations numerically far from the mass of data, is defined as a data point that is located outside the whiskers of the boxplot in Figure 2 (e.g., outside 1.5 times the interquartile range above the upper (third) quartile and below the lower (first) quartile).[12]

The effect of downstream closeness (DM) is positive and significant at the conventional level. That is, closer the industries locate with each other in terms of the downstreamness linkages, higher the regional productivity will be. As indicated in Table 3, the elasticity of labor productivity with respect to downstream closeness is above 2, which implies that a 10% increase in downstream closeness would lead to a 20% increase in labor productivity. Assuming that the elasticity of labor productivity with respect to downstream closeness is 2.5 and both labor productivity and downstream closeness index are located near the mean values (that is, the mean values of labor productivity and downstream closeness are 140, 0.404, respectively), a 1% increase in downstream closeness leads to a 3.5 U.S. dollars increase in value added per hour worked. This increase in value added per hour worked corresponding to a change in downstream closeness index can be sufficiently large from an economic point of view in that the downstream closeness index ranges from .374 to .428.[13] Note that the coefficients obtained by using the restricted sample (i.e., column (3)-(6)) is slightly smaller than those obtained from the full sample, but there is an increase in the adjusted *R*-squared when excluding the outliers of labor productivity.

Capital intensity (CL) is positively associated with regional labor productivity as expected, while education variable (EDU, the percentage of the population 25 years and older who had a bachelor's degree or higher) does not influence regional labor productivity. It is probably because we use the percentage of the labor forces who had a bachelor's degree or higher as a proxy for a education level. The average year of

[11] The estimation results from log-level regression is shown in Appendix Table 6.
[12] For more details about boxplot and the outliers of labor productivity, see Figure 2.
[13] For more details about the descriptive statistics including the mean values, see Table 2 and Figure 2.

Table 3. Estimation Results: Log-Log Regression Model

	Full Sample			Sample Excluding Outliers of LProd		
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	ln(LProd)	ln(LProd)	ln(LProd)	ln(LProd)	ln(LProd)	ln(LProd)
$ln(DM)$	3.078***	3.088***	2.607**	2.888***	2.845***	2.026*
	(0.863)	(0.870)	(1.286)	(0.671)	(0.688)	(1.053)
$ln(CI)$		0.137*	0.131*		0.120**	0.114**
		(0.070)	(0.078)		(0.049)	(0.057)
$ln(EDU)$			0.119			0.184
			(0.181)			(0.134)
Observations	142	140	138	130	128	126
Adj. R^2	0.074	0.100	0.097	0.112	0.141	0.156

Notes: The ln(LProd) is the logarithm of labor productivity at the regional level. Robust standard errors are in parentheses. ***,**, and * refer to statistical significance at the 1%, 5%, and 10% level, respectively.

education per capita as a measure of the human capital stock can be an alternative to link between regional labor productivity and the level of eduction, but it is not attainable.

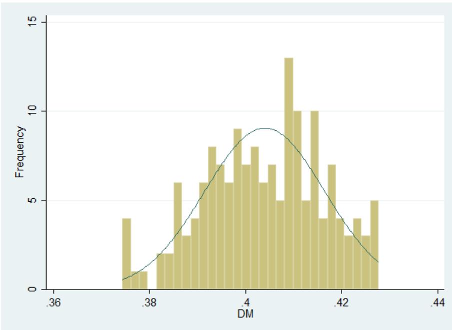
Instead of using a log transformed downstream closeness, we regress the logarithm of regional labor productivity on downstream closeness (the original, non-transformed variable) and the logarithm of regional controls. Table 6 in Appendix reports the estimation results using log-level regression. As shown in the previous results in Table 3, the downstream closeness has a positive effect on regional labor productivity at the conventional significance level. The coefficients for the downstream closeness in this regression are much greater than those with log-log regression reported in Table 3 because the logarithmic transformations of downstream closeness make distribution more spread out.[14]

The coefficients and standard errors for the other control variables are nearly identical regardless of whether taking a logarithm of downstream closeness variable or not. The capital intensity variable, CI, performs as expected in terms of sign and significance level, which is reported in Appendix Table 6. The effect of education on regional labor productivity is positive, but not statistically significant, which is the same as in Table 3.

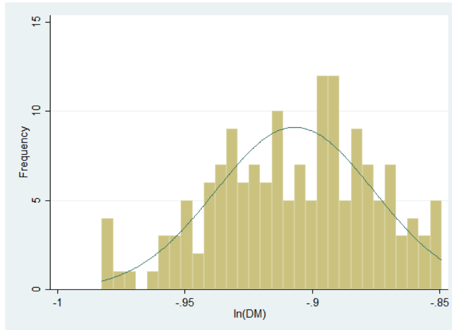
Lastly, we use the alternative measure of downstream closeness to see whether the main results are

[14] In general, log transformations are commonly used as a means of transforming a highly skewed variable into one that is more approximately normal. However, both forms of the variable of interest (that is, DM and $ln(DM)$) seem to be normally distributed, which is illustrated in Figure 3.

Figure 3. Histogram of Downstream Closeness: Original v.s. Logarithm



(a) DM



(b) Log(DM)

robust to the different measure of downstream closeness. An alternative measure of downstream closeness can be calculated by removing the inverse term of D_i^s in equation (6), so that a small $DM_{i,r}$ implies that form industry i's point of view there is a large presence of downstream industries in region r that are close to industry i. Thus, the estimation results, which is reported in Table 7, obtained using the alternative measure of downstream closeness are now expected to be negative. As shown in Table 7, the coefficient of downstream closeness is statistically significant and negative as expected. That is, the closer the industries locate with each other in terms of the downstreamness linkages, higher the regional productivity will be. The coefficients and standard errors for the other control variables (i.e., capital intensity and education) are nearly identical to the results shown in Table 3.

4. *Concluding Remarks*

In this paper, we introduce a new measure of inter-industry distance that emphasizes the role of input-output linkages in determining the industry agglomeration. Previous studies only consider a direct channel of input-output flows between industries (that is, upstream industry i's output can be directly used as an intermediate input in the production of downstream industry j), so that they overlook an indirect channel that includes crucial information on input-output linkages between industries (that is, upstream industry i's output can be indirectly used as an intermediate input in the production of downstream industry j after being used as inputs to other industries). In this context, we take into account both direct and indirect channels to accurately measure the degree of inter-industry distance between two industries.

Using the 2007 U.S. I-O table, we first construct the downstream closeness for 160 U.S regions. In addition to the regional-level downstream closeness, we also explore the *Economics Census of the United States* data to examine a link between the downstream closeness and regional labor productivity. We then empirically establish that the regional labor productivity is positively associated with the downstream closeness, which implies that a region with a higher downstream closeness is more likely to be productive.

This paper contributes to two streams of literature. One is the literature on the measure of downstream closeness using input-output linkages, and the other is the literature on regional agglomeration and productivity. Our new measure can be used in the future studies that stress the importance of the indirect channels in measuring downstream closeness using the input-output structure, so as to fully captures the relationship between the industrial linkages in terms of downstream closeness and productivity.

Throughout the paper, we introduce a new measure of downstream closeness, that focuses on the input flows from upstream to downstream industry, and quantify a relationship between downstream closeness and labor productivity. However, upstream closeness is also an important measure of the interindustry connectedness. In contrast to downstream closeness, upstream closeness emphasizes the importance of the input-output linkages from a downstream industry point of view. That is, upstream closeness can be constructed by using the relative importance of upstream industries whose output is used as an input to the production of downstream industry. Since the labor productivity at the regional level may be associated with upstream closeness as well as downstream closeness, it is desirable to additionally construct upstream closeness and decompose the impact of interindustry closeness on labor productivity into both downstream and upstream closeness, which will be addressed in subsequent research.

References

Alfaro, Laura and Andrew Charlton. 2009. "Intra-Industry Foreign Direct Investment." *AmericanEconomic Review*, 99, 2096−2119.

Antrás, Pol and Davin Chor. 2013. "Organizing the Global Value Chain." *Econometrica*, 81 (6), 2127−2204.

Antrás, Pol, Davin Chor, Thibault Fally, and Russell Hillberry. 2012. "Measuring the Upstreamnessof Production and Trade Flows." *American Economic Review Papers & Proceedings*, 102, 412−416.

Ciccone, Antonio. 2002. "Agglomeration Effects in Europe." *European Economic Review*, 46, 213−227.

__and Robert Hall. 1996. "Productivity and the Density of Economic Activity." *American EconomicReview*, 86, 54−70.

Combes, Pierre-Philippe and Laurent Gobillon. 2014. "The Empirics of Agglomeration Economics." IZA Discussion Paper.

di Giovanni, Julian and Andrei Levchenko. 2010. "Putting the Parts Together: Trade, Vertical Linkagesand Business Cycle Comovement." *American Economic Journal: Macroeconomics*, 2, 95−124.

Ellison, Glenn, Edward Glaeser, and William Kerr. 2010. "What Causes Industry Agglomertaion? Evidence from Coagglomeration Patterns." *American Economic Review*, 100, 1195−1213.

Lopez, Ricardo and Jens Sudekum. 2009. "Vertical Industry Relations, Spillovers and Productivity: Evidence from Chilean Plants." *Journal of Regional Science*, 49, 721−747.

Marshall, Alfred. 1920. *Principles of Economics*. London: MacMillan.

Rasmussen, P.N. 1956. *Studies in Inter-sectoral Relations*. Amsterdam: North-Holland.

Rigby, David and Essletzbichler. 2002. "Agglomertaion Economics and Productivity Differences inUS Cities." *Journal of Economic Geography*, 2, 407−432.

Rosenthal, Stuart and William Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In J.V. Henderson and J.-F. Thisse eds. *Handbook of Urban and Regional Economics*, 4, 2119−2172.

Table 4. Inter-industry Closeness for NAICS 3-digit Manufacturing Industries

NAICS	311	312	313	314	315	316	321	322	323	324	325	326	327	331	332	333	334	335	336	337	339
23	3.0	3.3	3.0	3.4	3.4	3.3	2.7	2.4	2.7	2.0	2.6	3.0	2.5	2.5	2.8	3.1	2.7	3.2	3.3	3.2	3.2
42	2.0	2.2	2.2	2.2	1.8	1.8	1.8	1.9	2.5	2.6	1.7	2.2	2.2	1.7	2.1	1.9	1.7	1.9	1.9	2.1	2.0
47	3.6	3.1	3.6	3.2	4.1	3.9	3.4	3.2	1.6	3.0	2.2	3.3	3.4	3.4	3.6	3.5	3.5	3.7	3.7	2.5	3.0
48	3.1	3.6	3.4	3.5	3.5	3.5	2.5	2.9	3.1	3.8	3.1	3.3	2.6	3.0	3.4	3.5	3.3	3.6	3.5	3.1	3.3
111	1.3	1.6	1.2	2.5	2.5	2.7	3.5	3.0	3.2	4.0	1.4	3.1	3.6	4.4	4.2	3.9	4.1	4.3	4.1	2.9	2.1
112	1.2	2.7	1.7	2.3	3.1	2.6	4.0	3.3	3.2	4.8	2.9	4.2	4.1	5.3	5.1	5.0	5.1	5.4	5.1	4.0	4.1
113	3.4	3.3	4.2	4.2	2.2	2.6	1.2	1.4	3.0	4.7	3.0	1.7	3.6	4.0	4.2	3.8	3.7	3.8	3.5	2.6	2.6
114	1.0	2.5	3.9	4.2	4.1	2.5	3.6	3.4	3.1	4.7	2.5	3.7	3.6	4.1	3.8	3.9	4.0	4.5	4.4	4.0	3.8
115	2.3	2.8	2.3	3.6	3.4	3.6	2.2	2.6	4.1	5.5	2.8	3.0	4.6	5.1	5.2	4.8	4.8	5.0	4.7	3.6	3.4
211	3.8	4.2	3.7	4.1	4.4	4.4	3.6	3.3	3.4	1.1	2.4	3.7	3.4	3.4	4.1	4.1	4.1	3.9	4.4	4.2	4.1
212	3.1	3.2	3.0	3.7	4.3	4.3	3.7	2.1	3.9	3.2	2.1	3.4	1.3	1.2	2.6	3.1	3.3	2.8	3.2	3.4	3.1
213	4.8	5.0	4.7	5.2	5.5	5.5	4.7	4.0	4.7	2.3	3.6	4.8	3.1	2.9	4.4	4.7	4.9	4.5	5.0	5.0	4.8
221	2.2	2.6	1.9	2.6	2.6	2.7	2.0	1.5	2.3	2.6	1.5	2.2	1.6	1.6	2.2	2.6	2.3	2.6	2.7	2.6	2.6
311	-	1.4	3.2	3.5	4.0	1.4	3.4	2.1	2.0	3.6	1.7	3.1	2.9	4.2	4.0	4.1	4.0	4.3	4.3	4.0	3.8
312	1.5	-	3.8	4.2	3.8	3.2	3.4	3.6	3.2	4.8	3.6	3.7	3.5	3.9	3.6	3.7	3.8	4.3	4.1	3.7	3.5
313	3.5	3.3	-	1.2	1.2	1.3	1.6	1.3	1.9	4.4	3.6	1.8	2.1	4.0	4.0	1.9	3.5	3.6	2.1	1.3	1.3
314	2.9	2.8	2.1	-	1.1	3.3	1.6	3.3	3.6	3.4	3.2	1.1	1.9	3.6	3.7	3.2	3.2	3.6	1.3	1.5	2.7
315	3.5	3.7	1.1	1.8	-	1.9	3.1	3.0	1.1	4.3	3.2	3.4	3.4	3.3	3.5	3.4	3.2	3.5	3.3	2.9	2.9
316	3.2	4.1	4.4	1.1	1.1	-	3.7	4.3	1.1	3.9	3.9	3.7	4.1	4.3	4.4	3.2	3.9	4.3	1.1	3.9	1.4
321	3.7	3.3	4.0	3.5	4.4	3.8	-	1.7	3.4	3.5	3.3	2.4	2.8	3.1	3.5	2.9	2.7	3.1	2.6	1.4	2.3
322	1.7	1.9	2.3	2.7	3.2	2.7	2.4	-	1.5	4.1	2.0	1.9	2.0	2.4	2.4	2.5	2.3	2.2	2.7	2.1	2.0
323	3.5	3.0	3.7	4.0	1.9	3.5	3.3	3.7	-	4.4	2.5	3.7	3.8	3.6	3.6	3.6	3.4	3.7	3.7	3.7	3.5
324	2.8	3.2	2.9	3.3	3.5	3.5	2.6	2.4	2.4	-	1.7	2.9	2.5	2.5	3.2	3.1	3.2	2.8	3.5	3.2	3.1
325	3.0	3.0	1.7	2.0	2.9	2.3	2.5	2.0	2.2	2.6	-	1.6	2.3	3.1	2.4	2.8	2.3	2.5	2.8	2.6	2.2
326	2.0	1.6	2.7	2.6	3.6	1.8	2.4	2.1	2.7	3.6	1.7	-	2.4	2.8	2.4	1.8	1.9	2.1	1.6	1.5	1.7

Continued on next page

Table 4 – *Continued from previous page*

NAICS	311	312	313	314	315	316	321	322	323	324	325	326	327	331	332	333	334	335	336	337	339
327	3.5	1.7	2.7	3.9	4.0	4.1	1.9	3.2	3.8	3.0	2.8	2.4	-	2.1	2.7	2.5	2.7	2.1	2.3	2.8	2.7
331	4.4	2.4	4.6	3.5	4.8	4.3	3.3	3.5	3.9	4.5	3.7	3.2	3.2	-	1.5	1.8	2.1	1.7	2.0	2.2	2.0
332	2.8	2.1	3.0	2.9	3.4	2.6	2.0	2.2	2.6	3.6	2.3	2.2	2.3	2.3	-	1.6	1.7	1.7	1.8	2.0	2.1
333	3.4	2.6	3.8	4.2	4.3	4.4	2.8	2.6	2.3	3.4	2.7	2.8	3.1	2.6	2.4	-	2.9	2.3	1.7	3.7	2.7
334	4.1	2.7	3.0	2.9	3.5	3.2	2.8	2.5	2.5	4.4	2.6	2.8	2.8	2.9	2.5	2.2	-	2.0	1.9	2.6	2.4
335	3.6	3.2	3.7	3.8	4.0	4.1	2.1	3.0	3.3	3.6	3.1	2.8	3.4	2.1	2.2	1.5	1.5	-	2.1	3.4	2.5
336	3.6	3.3	4.0	4.0	4.3	4.5	2.7	3.2	3.6	4.4	3.5	3.7	3.2	3.3	3.3	2.0	2.8	3.2	-	3.3	3.5
337	4.0	4.1	4.1	4.4	4.3	4.1	1.3	3.3	3.7	3.2	3.6	2.3	3.6	3.6	3.9	2.4	1.9	3.9	2.9	-	2.1
339	3.6	3.7	3.0	2.3	1.4	1.4	3.2	3.6	3.3	4.2	2.0	3.0	2.1	3.4	3.3	1.4	3.4	1.7	2.0	3.1	-
441	2.9	2.4	3.1	3.0	3.7	3.8	2.3	2.4	2.8	3.7	2.6	3.0	2.3	2.8	2.7	1.5	2.4	3.5	1.8	2.5	2.5
445	2.9	1.9	3.4	3.7	3.7	3.7	3.5	2.8	1.2	3.3	1.6	2.7	2.6	3.3	3.1	2.7	3.1	2.7	2.8	1.5	1.9
452	3.1	2.2	3.6	2.6	3.6	3.8	3.1	2.9	1.2	3.3	2.1	2.8	3.0	3.4	3.3	2.5	3.1	3.1	2.7	1.8	2.2
481	2.6	2.9	2.7	3.0	2.8	2.5	2.1	2.4	2.0	3.8	2.3	2.3	2.2	2.6	2.3	2.4	2.3	2.9	2.7	2.3	2.4
482	1.9	2.7	2.6	3.2	3.8	3.1	1.7	1.5	2.7	3.0	1.5	2.1	1.5	1.3	2.4	2.8	2.9	2.6	2.7	2.4	2.7
483	1.5	2.2	2.3	3.4	3.5	3.0	3.0	2.2	3.3	2.7	1.7	2.9	1.5	1.4	2.6	3.0	2.7	2.4	3.1	3.0	2.6
484	1.6	2.2	1.9	2.1	2.0	2.2	1.7	1.8	2.3	2.2	1.9	2.1	1.7	1.6	2.1	2.1	2.1	2.2	2.1	2.1	2.2
485	3.1	3.2	3.2	3.4	3.0	2.8	2.3	2.6	2.1	4.2	3.0	2.7	2.4	2.9	2.5	2.6	2.7	3.4	3.2	2.5	2.4
486	3.6	3.9	3.2	3.9	4.1	4.2	3.4	3.0	3.4	1.5	2.5	3.5	3.1	3.1	3.7	3.9	3.8	3.8	4.2	4.0	3.9
491	3.2	3.6	3.4	3.6	3.3	3.4	3.1	3.3	2.6	3.8	3.1	3.5	3.2	3.0	3.3	2.5	3.1	2.5	3.4	3.4	3.2
492	3.1	3.6	3.4	3.6	3.3	3.3	3.0	3.2	3.2	3.8	3.2	3.5	3.2	3.0	3.4	3.4	3.2	3.4	3.4	3.4	3.4
493	2.9	2.9	2.4	2.5	2.4	2.5	2.1	2.5	2.0	3.9	2.5	2.6	2.5	2.7	2.3	2.6	2.3	2.8	2.6	2.3	1.9
511	3.7	3.6	3.5	3.9	3.4	3.4	3.4	3.6	2.8	4.4	3.0	3.6	3.5	3.6	3.3	3.3	2.7	3.4	3.3	3.5	2.9
512	5.0	4.9	4.8	5.0	4.5	4.6	4.3	4.6	4.2	5.6	4.5	4.7	4.5	4.9	4.5	4.5	4.1	4.9	4.8	4.2	4.3
515	4.9	4.8	4.7	4.9	4.4	4.5	4.2	4.5	4.1	5.2	4.4	4.6	4.4	4.8	4.3	4.3	3.9	4.7	4.7	4.0	4.2
517	3.3	3.3	3.1	3.3	2.8	3.0	2.7	2.9	2.6	4.0	2.9	3.0	2.8	3.2	2.8	2.8	2.4	3.2	3.2	2.4	2.6
518	3.0	3.1	2.7	2.9	2.6	2.3	2.0	2.3	2.4	4.2	2.7	2.6	2.2	2.8	2.3	2.3	2.4	2.9	2.6	2.3	2.2
519	4.1	3.6	3.8	4.4	3.3	3.6	3.7	3.6	3.5	4.6	3.2	3.9	3.7	3.8	3.5	3.7	3.2	3.9	3.8	3.9	3.4
521	2.9	3.1	2.9	3.4	2.9	2.9	3.0	3.2	3.1	4.0	3.0	2.9	2.6	3.0	2.8	3.0	3.0	3.4	3.4	3.0	2.8

Continued on next page

Table 4 – *Continued from previous page*

NAICS	311	312	313	314	315	316	321	322	323	324	325	326	327	331	332	333	334	335	336	337	339
522	3.2	3.2	3.0	3.2	2.6	2.5	2.7	2.9	2.8	4.1	2.9	3.1	2.8	2.9	3.0	3.0	2.9	3.1	3.3	2.9	2.9
523	3.7	3.8	3.7	4.2	3.6	3.6	3.5	3.7	3.4	4.2	3.5	3.6	3.3	2.4	2.4	2.4	3.0	3.0	3.0	2.5	2.4
524	3.3	3.4	3.2	3.5	2.3	2.2	3.3	3.5	3.5	4.4	3.2	3.7	3.2	3.5	3.5	3.4	3.1	3.3	3.7	3.4	3.4
622	4.5	4.6	4.4	4.8	3.6	3.5	4.4	4.6	4.5	5.3	4.4	4.7	4.3	3.8	3.8	3.8	4.1	4.2	4.3	3.9	3.8
531	3.6	3.7	3.6	3.7	3.3	3.3	3.2	3.5	3.3	4.3	3.3	3.7	3.5	3.5	3.4	3.6	3.3	3.7	3.7	3.4	3.4
532	2.9	3.3	3.1	3.5	3.2	3.4	2.7	2.8	2.5	3.0	2.8	3.1	2.6	2.8	2.9	2.9	2.8	3.3	3.2	3.1	3.1
533	2.9	1.9	2.7	3.5	2.0	2.4	2.9	2.3	2.5	3.3	1.8	2.7	2.5	2.8	2.2	2.3	1.7	2.5	2.5	3.1	2.0
541	3.2	3.2	3.0	3.3	2.6	2.6	2.6	2.9	2.7	3.8	2.7	3.0	2.8	3.0	2.8	2.9	2.5	3.2	3.1	2.8	2.5
551	2.2	1.9	2.0	2.7	2.1	2.4	2.0	2.0	2.2	2.8	1.4	2.4	2.1	2.5	2.3	2.3	1.8	2.4	1.9	2.4	1.8
561	3.3	3.3	3.0	3.5	2.0	2.8	2.7	2.7	2.1	4.0	2.8	3.0	2.7	2.7	2.4	2.9	2.3	3.1	2.9	3.1	2.9
562	2.5	2.5	2.3	2.8	2.8	2.7	2.4	2.2	2.5	3.1	1.7	2.6	2.3	2.5	2.5	2.8	2.5	2.9	2.9	2.6	2.6
611	2.9	3.6	3.3	3.9	3.5	3.6	3.4	3.5	1.7	4.0	3.3	3.7	3.7	3.4	3.7	3.7	3.5	3.8	3.8	3.7	3.6
621	4.2	4.3	4.0	4.4	3.5	3.7	3.7	3.9	3.5	4.9	3.8	4.0	3.8	4.0	3.7	3.9	3.5	4.2	4.1	3.9	3.6
711	3.7	3.9	3.6	3.9	3.3	3.6	3.1	3.3	3.2	4.6	3.4	3.5	3.3	3.5	3.5	3.5	3.6	4.1	3.8	3.6	3.4
713	3.0	3.1	3.1	3.5	3.0	3.0	2.4	2.7	2.2	4.1	2.8	2.8	2.5	3.0	2.6	2.7	2.7	3.5	3.1	2.6	2.5
721	3.2	3.2	3.2	3.5	3.0	2.9	2.4	2.7	1.7	4.1	3.0	2.5	2.5	3.0	2.5	2.6	2.7	3.4	3.2	2.6	2.5
722	3.0	3.2	2.9	3.2	2.7	2.8	2.3	2.6	2.2	3.9	3.0	2.7	2.5	2.9	2.6	2.7	2.8	3.4	3.2	2.7	2.5
811	2.8	3.0	2.7	3.0	2.8	2.8	2.4	2.2	2.4	3.5	2.3	2.7	2.3	2.3	2.5	2.7	2.4	2.8	2.9	2.8	2.7
812	3.6	3.8	2.5	3.0	3.1	3.4	3.2	3.4	3.5	4.4	3.4	3.6	2.6	2.8	2.5	3.1	3.3	3.5	3.1	3.5	3.0
813	3.0	3.2	3.0	3.4	3.0	3.1	2.5	2.7	2.7	3.7	3.0	2.9	2.7	3.0	2.9	3.1	3.1	3.5	3.4	2.9	2.8

Table 5. NAICS Description

NAICS	NAICS Description	NAICS	NAICS Description
23	Construction	42	Wholesale Trade (423, 424, 425)
47	Arbitrary Sector (442, 443, 444, 446, 448, 451, 453, 454)	48	Arbitrary Sector (487, 488)
111	Crop Production	112	Animal Production
113	Forestry and Logging	114	Fishing, Hunting and Trapping
115	Support Activities for Agriculture and Forestry	211	Oil and Gas Extraction
212	Mining (except Oil and Gas)	213	Support Activities for Mining
221	Utilities	311	Food Manufacturing
312	Beverage and Tobacco Product Manufacturing	313	Textile Mills
314	Textile Product Mills	315	Apparel Manufacturing
316	Leather and Allied Product Manufacturing	321	Wood Product Manufacturing
322	Paper Manufacturing	323	Printing and Related Support Activities
324	Petroleum and Coal Products Manufacturing	325	Chemical Manufacturing
326	Plastics and Rubber Products Manufacturing	327	Nonmetallic Mineral Product Manufacturing
331	Primary Metal Manufacturing	332	Fabricated Metal Product Manufacturing
333	Machinery Manufacturing	334	Computer and Electronic Product Manufacturing
335	Electrical Equipment, Appliance, Component Manufacturing	336	Transportation Equipment Manufacturing
337	Furniture and Related Product Manufacturing	339	Miscellaneous Manufacturing
441	Motor Vehicle and Parts Dealers	445	Food and Beverage Stores
452	General Merchandise Stores	481	Air Transportation
482	Rail Transportation	483	Water Transportation
484	Truck Transportation	485	Transit and Ground Passenger Transportation
486	Pipeline Transportation	491	Postal Service
492	Couriers and Messengers	493	Warehousing and Storage
511	Publishing Industries (except Internet)	512	Motion Picture and Sound Recording Industries
515	Broadcasting (except Internet)	517	Telecommunications
518	Data Processing, Hosting and Related Services	519	Other Information Services
521	Monetary Authorities-Central Bank	522	Credit Intermediation and Related Activities
523	Securities, Commodity Contracts	524	Insurance Carriers and Related Activities
525	Funds, Trusts, and Other Financial Vehicles	531	Real Estate
532	Rental and Leasing Services	533	Lessors of Nonfinancial Intangible Assets
541	Professional, Scientific, and Technical Services	551	Management of Companies and Enterprises
561	Administrative and Support Services	562	Waste Management and Remediation Services
611	Educational Services	621	Ambulatory Health Care Services
711	Performing Arts, Spectator Sports	713	Amusement, Gambling, and Recreation Industries
721	Accommodation	722	Food Services and Drinking Places
811	Repair and Maintenance	812	Personal and Laundry Services
813	Religious, Grantmaking, Civic	-	-

Table 6. Estimation Results: Log-Level Regression Model

	Full Sample			Sample Excluding Outliers of LProd		
Variables	(1) ln(LProd)	(2) ln(LProd)	(3) ln(LProd)	(4) ln(LProd)	(5) ln(LProd)	(6) ln(LProd)
DM	7.657***	7.691***	6.536**	7.182***	7.081***	5.070*
	(2.159)	(2.171)	(3.192)	(1.679)	(1.714)	(2.610)
$ln(CI)$		0.138*	0.131*		0.120**	0.114**
		(0.070)	(0.078)		(0.050)	(0.057)
$ln(EDU)$			0.117			0.183
			(0.180)			(0.133)
Observations	142	140	138	130	128	126
Adj. R^2	0.074	0.100	0.097	0.111	0.141	0.157

Notes: The ln(LProd) is the logarithm of labor productivity at the regional level. Robust standard errors are in parentheses. ***,**, and * refer to statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7. Estimation Results: Log-Log Regression Model

	Full Sample			Sample Excluding Outliers of LProd		
Variables	(1) ln(LProd)	(2) ln(LProd)	(3) ln(LProd)	(4) ln(LProd)	(5) ln(LProd)	(6) ln(LProd)
$ln(DM)^{alternative}$	-2.672***	-2.664***	-1.987	-2.634***	-2.577***	-1.731*
	(0.813)	(0.820)	(1.212)	(0.622)	(0.637)	(0.983)
$ln(CI)$		0.136*	0.131*		0.118**	0.114**
		(0.070)	(0.079)		(0.049)	(0.056)
$ln(EDU)$			0.154			0.194
			(0.182)			(0.135)
Observations	142	140	138	130	128	126
Adj. R^2	0.066	0.092	0.088	0.110	0.139	0.152

Notes: The ln(LProd) is the logarithm of labor productivity at the regional level. Robust standard errors are in parentheses. ***,**, and * refer to statistical significance at the 1%, 5%, and 10% level, respectively.

List of KIEP Working Papers (2014-2016. 12)

A List of all KIEP publications is available at: http://www.kiep.go.kr

16-01 Access to Credit and Quality of Education in Vietnam
HUR Yoon Sun

■ 2015

15-03 Estimating Regional Matching Efficiencies in the Indian Labor Market: State-level Panel Data for 1999-2011
LEE Woong

15-02 The Distribution of Optimal Liquidity for Economic Growth and Stability
PYO Hak K. and SONG Saerang

15-01 Income Distribution and Growth under A Synthesis Model of Endogenous and Neoclassical Growth
KIM Se-Jik

■ 2014

14-05 Regional Financial Arrangement in East Asia: Policy Proposal for Strengthening the Chiang Mai Initiative Multilateralization
Pravin Krishna, Jiyoung Choi, and Tae-Hoon Lim

14-04 Labor Market Flexibility and Different Job-Matching Technologies across Regions in India: An Analysis of State-Level Dis-aggregate Matching Functions
Woong Lee

14-03 Rising Income Inequality and Competition: Evidence
Minsoo Han

14-02 Inequality and Fiscal Policy Effectiveness
Ju Hyun Pyun and Dong-Eun Rhee

14-01 Inequality and Growth: Nonlinear Evidence from Heterogeneous Panel Data
Dooyeon Cho, Bo Min Kim, and Dong-Eun Rhee

국문요약

본 연구에서는 산업간 연계성을 측정하는 새로운 지수(index)를 제시하였다. 이 지수는 기존 지수와 달리 한 산업의 생산물이 다른 산업의 중간재로 이용되기까지 거치는 수직적 생산 단계(vertical production stage)를 구분하고 각 단계에 가중치를 부여함으로써 산업간 '거리' 개념을 포함하였다. 이에 따라 본 지수는 산업간 연계성에 있어 더욱 포괄적인 정보를 포함하고 있다고 볼 수 있다. 한편 본 연구는 개발된 지수를 이용하여 미국 각 지역의 '산업간 밀접도'를 측정하는 지수를 구축, '산업간 밀접도'가 해당 지역의 노동생산성과 양의 관계를 갖고 있음을 실증적으로 입증하였다. 이는 지역 내 산업간 연계성이 집적효과(agglomeration effect)에 기여한다는 마샬의 이론을 뒷받침하는 결과로, 해당 지수가 산업간 연계성을 측정하는 대안적인 지수로 활용될 수 있음을 나타낸다.

핵심용어: 산업간 연계성, 직접효과, 생산성

윤여준(尹汝峻)

연세대학교 경제학 학사 및 석사

University of Warwick 경제학 박사

대외경제정책연구원 구미유라시아본부 미주팀 부연구위원

(現, E-mail: yoonyj@kiep.go.kr)

저서 및 논문

"A Quantitative Analysis of U.S. Economic Development, 1870-1913" (2014)

『미국 통화정책 정상화에 따른 출구전략 효과 및 시사점』(공저, 2015) 외

황운중(黃云重)

미국 콜로라도대학교 경제학 박사

대외경제정책연구원 무역통상본부 지역무역협정팀 부연구위원

(現, E-mail: ujwhang@kiep.go.kr)

저서 및 논문

"Structural Transformation and Comparative Advantage: Implications for Small Open Economies" (*The World Economy*, 2016)

"To Whom does Outward FDI Give Jobs?" (공저, 2016) 외

KIEP Working Paper 16-11

A New Measure of Inter-industry Distance and Its Application to the U.S. Regional Growth

YOON Yeo Joon and WHANG Un Jung

We propose a new measure of inter-industry 'distance'. This is constructed à la Antras *et al*. (2012). While they measure the distance of an industry from its final use – what they call 'downstreamness' of an industry – we measure the distance between a pair of industries. Our proposed index is a measure of input-output linkages between industries that incorporates a 'distance' flavor. Our measure distinguishes the number of vertical production stages that an industry's product goes through until it is finally used by another industry by assigning larger weights to the value of input use with longer production chains. Hence our measure contains more information on the relation between two industries along the vertical production chain. We use this index to construct an aggregate measure of 'industry connectedness' of regions in the U.S. It measures the degree of industrial linkages of a region. We then empirically establish that each region's labor productivity is positively associated with the 'industry connectedness'. The result contributes to the large literature of agglomeration economies that the industrial linkage is one of the main sources of agglomeration economies and productivity growth, as emphasized by Marshall (1920). It also suggests that our index can serve as an alternative measure of the industrial linkages.

9 788932 242620 94320

ISBN 978–89–322–4262–0
978–89–322–4026–8(set)

Price USD 3

KIEP Korea Institute for International Economic Policy

Building C, Sejong National Research Complex, 370, Sicheong-daero, Sejong-si, Korea
Tel. 82-44-414-1251 Fax. 82-44-414-1144
www.kiep.go.kr