Non-survey Method for Estimating a Multi-regional Input-Output Model in China

Nobuhiro Okamoto, Daito Bunka University

Abstract

This paper proposes a method for estimating a multi-regional input-output (MRIO) model for China under the limitations and constraints of trade flow data over the regions. The MRIO model proposed here is the Chenery-Moses-type column model, in which an interregional trade coefficient is estimated using the Leontief-Strout Gravity (LSG) model. It is important to estimate the spatial friction parameter Q of LSG. Thus, in this paper, we tested parameter Q, estimated with random variables and distance data. We concluded that we can construct an MRIO model using the LSG model with random variables and distance data as the information on spatial friction of trade flows.

Key words: interregional trade coefficient, Leontief-Strout Gravity model, Q parameter, Transport Distribution Index

1. Introduction

Since the start of the twenty-first century, macroeconomic issues on regional disparity have become a focal point among scholars both in and outside of China. In order to clarify the spatial structure in the vast territory of China, regional economists and input-output specialists began studying the interregional input-output accounts of China . As a result, interregional input-output tables or data were developed by a few research groups, such as Ichimura and Wang (2003), the Institute of Developing Economies-JETRO (2003) in collaboration with the State Information Center (SIC 2005), Miyagawa et al. (2008), and Li (2010).

It is well known that an interregional input-output approach to analyze regional development and the interdependency of regional economies is very useful, but it is not easy to construct input-output data or tables, especially in developing countries such as China. If there was a shortcut or an easy method to estimate interregional input-output tables in China, it might have led to researchers entering the fields of spatial structure and regional development and contributing to the understanding of China’s regional economy.

Thus, I would like to propose a shortcut method to estimate the interregional input-output model using the existing data of China. The method proposed in this paper is a multi-regional input-output (MRIO) model in which interregional commodity flows are estimated using the Leontief-Strout Gravity (LSG) model. The key point of the LSG model is how to estimate the spatial friction of the Q parameter. In order to estimate the Q parameter in our model, we test the random variables data with some kind of probability distribution or distance data and conclude that this method is useful.

The paper consists of three parts. First, in section 2, we introduce the MRIO model with interregional commodity flows estimated using the LSG model. Section 3 discusses the meanings of the Q parameter, which shows spatial friction of commodity transport between regions. In the next part, section 4, we analyze the results of the Q parameter estimated using the data of random variables with two kinds of probability distribution and the Trade Distribution Index of distance. Finally, in section 5, we conclude that the MRIO model estimated using distance data could be practically useful for analyzing interregional interaction among regions.

2. Multi-regional Input-Output Model and the Q Parameter

2-1 Previous research

Interregional input-output models are of two types (Miller and Blair 2009, 76-77): the interregional input-output table or the so-called Isard type (Isard 1951), which shows the input commodities recorded differently by region of origin, and the multi-regional input-output model or the so-called Chenery-Moses type (Chenery 1953, Moses 1955), which shows the commodities by region, although we do not know which sector uses it.

Usually, it is very difficult for researchers to construct an Isard-type interregional input-output model, because it requires a lot of data related to interregional commodity flows among regions. On the other hand, there are two merits of constructing the Chenery-Moses-type MRIO model: first, it requires less data, and second, it is easy to update the model because interregional trade coefficients and technical coefficients are independently estimated. According to Polenske (1970), the column model of the MRIO table and the Leontief-Strout model (Leontief and Strout 1963) are highly reliable.

A few methodologies have been used to develop interregional input-output models in China. Akita, Kawamura, and Xie (1999) constructed the model using the location quotient; Ando and Shibata (1996) and Okuda et al. (2004) estimated the model using the RAS or entropy method; and Miyagawa et al. (2008) tried to construct the model using the LSG model. However, the problem with the location quotient approach is that cross-hauling is not allowed, and the problem with the matrix convergence approach, such as RAS and entropy, is that it lacks economic sense because it is used for balancing row and column figures mathematically. Miyagawa et al. (2008) estimated the interregional commodity flows using the LSG model. Although we also use the LSG model, it is different in terms of estimating the Q parameter. Our proposed methodology is simpler than that of

Miyagawa et al. (2008), while maintaining data accuracy.

2-2 The model

Okamoto and Zhang (2003) and Okamoto et al. (2005) construct the MRIO model as follows, matrix-partitioning by region:

$$\begin{bmatrix} X^1 \\ \vdots \\ X^r \end{bmatrix} = \begin{bmatrix} \hat{T}^{11} & \cdots & \hat{T}^{1r} \\ \vdots & \ddots & \vdots \\ \hat{T}^{r1} & \cdots & \hat{T}^{rs} \end{bmatrix} \begin{bmatrix} A^1 & & 0 \\ & \ddots & \\ 0 & & A^r \end{bmatrix} \begin{bmatrix} X^1 \\ \vdots \\ X^r \end{bmatrix} + \begin{bmatrix} \hat{T}^{11} & \cdots & \hat{T}^{1r} \\ \vdots & \ddots & \vdots \\ \hat{T}^{r1} & \cdots & \hat{T}^{rs} \end{bmatrix} \begin{bmatrix} F^1 \\ \vdots \\ F^r \end{bmatrix} \tag{1}$$

Here,

$X^r = [x_i^r]$: the total output of industry *i* in region *r*.

$\hat{T}^{rs} = \begin{bmatrix} t_1^{rs} & & 0 \\ & \ddots & \\ 0 & & t_i^{rs} \end{bmatrix}$: the interregional trade coefficient between regions *r* and *s*.

We assume that we apply the column model, so $\sum_r^m \sum_i^n t_i^{rs} = 1$ holds.

$A^r = \left[a_{ij}^r\right]$: the technical coefficient matrix in region *r*.

$F^r = [f_{ik}^r]$: the final-demand items in region *r*.

Further, $r, s = 1,2, \cdots, m,$; this shows the number of region, $i, j = 1,2, \cdots, n,$ in the number of sector, $k = 1,2, \cdots, l$, for the number of final-demand items.

Equation (1) will be transformed by the matrix equation:

$$X = TAX + TF \tag{2}$$

Then, we can derive its analysis form:

$$X = (I - TA)^{-1}TF \tag{3}$$

In this paper, we propose to construct not an input-output table, but a model for analyzing the economic impact of interregional interdependence. In order to satisfy our analysis requirements, we need to estimate the interregional output multiplier, or the Leontief inverse matrix $(I - TA)^{-1}$, including the regional technical coefficient and interregional trade coefficient. Hereafter, when we refer to constructing a “model,” we mean the estimation of the multiplier and other exogenous data such as final-demand items.

From the column model condition, we derive

$$t_i^{rs} = \frac{x_i^{rs}}{\sum_r^m x_i^{rs}} \tag{4}$$

Here, x_i^{rs} represents the outflow of commodity *i* from region *r* to region *s* (or inflow from region *s* to region *r*). This interregional inflow/outflow of a commodity between regions can be represented in the form of the famous LSG model:

$$x_i^{rs} = \frac{x_i^r x_i^s}{\sum_r^m x_i^r} Q_i^{rs} \qquad (5)$$

Here, x_i^r represents the total output of region *r*, and x_i^s represents the total output of region *s*; $\sum_r^m x_i^r = \sum_s^m x_i^s$ holds. Further, Q_i^{rs} gives the spatial friction of commodity flows.
We substitute the LSG model (5) for the column model (4), and get

$$t_i^{rs} = \frac{x_i^r Q_i^{rs}}{\sum_r (x_i^r Q_i^{rs})} \qquad (6)$$

From this equation, we derive the interregional coefficient matrix of model (1) (or [2], [3]) from the total output of each region and the Q parameter.

The original MRIO table of the LSG model estimates the transactions between industries among the region as follows:

$$x_{ij}^{rs} = \frac{x_i^r x_j^s}{\sum_r^m x_i^r} Q_{\mathrm{ij}}^{rs} \qquad (7)$$

As the difference between equations (5) and (7) is very clear, our model estimates the interregional trade coefficient matrix by using the LSG model, and not each transaction between industries located in each region.[1]

2-3 Discussion on the Q parameter

It is essential that our model defines or estimates the Q parameter. According to the LSG model, the Q parameter gives the transportation cost of one unit of a commodity when it is moved, or the relative economic position of region r/s to the whole nation (the outflow power of the supply side or inflow power of the demand side). Although it is very difficult to estimate the Q parameter, this is possible by substituting the alternative information on interregional flow data such as transportation data. If so, we should be able to estimate the Q parameter as the Transport Distribution Index (Ihara 1996, Miller and Blair 2009, pp. 356-366).

The Transportation Distribution Index is defined as follows:

$$Q_i^{rs} = \frac{h_i^{rs}}{\frac{h_i^{r0} h_i^{0s}}{h_i^{00}}} \qquad (8)$$

where h_i^{rs} is the amount (value) of interregional commodity flow,
$h_i^{r0} = \sum_r h_i^{rs}$ is the total amount of commodity outflow from region *r*,
$h_i^{0s} = \sum_s h_i^{rs}$ is the total amount of commodity inflow to region *s*, and

1 This is different from Miyagawa et al. (2008). The detailed process of their estimation is shown in Wang (2003).

$h_i^{00} = \sum_r \sum_s h_i^{rs}$ is the total amount of commodity transportation of all regions.

The Transport Distribution Index shows the ratio of the observed interregional transportation data h_i^{rs} to the transportation data proportionally distributed from total transportation. The transportation data are in proportion to both the supply and demand pools, and the distributed transportation data become larger when the amount of goods transacted in the region increases. When the observed transportation data is larger than the distributed transportation data, other factors besides the amount of supply and demand pools are involved. In other words, this indicates that the spatial friction in commodity transportation between regions, for example, physical distance, time, and cost, would be small. In this case, Q_i^{rs} will be bigger than a unit. On the other hand, when the observed transportation data is smaller than the distributed transportation data, it indicates some sort of difficulty in commodity transportation and larger spatial friction. Thus, the value of Q_i^{rs} shows the following:

$Q_i^{rs} > 1$: less spatial friction and more interregional transactions
$Q_i^{rs} = 1$: interregional transactions are independent of spatial friction
$Q_i^{rs} < 1$: more spatial friction and less interregional transactions

From the above, we conclude that Q_i^{rs} is the friction resisting the smooth interregional transportation of commodities based on distance, transportation costs, and so on, other than the supply side's push factor and demand side's pull factor.

3. How to Estimate Q Parameter

Since it is very difficult to obtain data on the interregional flow of goods and services, we are unable to estimate the Q parameter using equation (8). However, in the case of China, we have data on the transportation of goods by railway transport (hereafter, the railway origin-destination [OD] table), published by the Ministry of Transportation, but this is published only for the total goods and coal transportation, and not detailed by commodities. Further, there are various modes of transportation, such as water, road, and air transport, besides railway transport. In China, road transport is the main mode, covering 72% of the total transportation, with railway transport covering only 13.8% (2007 figures).

In the following section, we discuss the estimation of the Q parameter based on the above data.

3-1 Random variable

As mentioned above, the Q parameter is the key point of estimation of interregional commodity flows. In case there is data limitation, this would serve as one method to estimate the preliminary figures for spatial friction by engineering methods.

First, we assume that interregional commodity flows among regions occur uniformly in a geographical range and that uniform distribution is the random variable. For example, imagine that a kind of fish has its own territory and is uniformly distributed in the river. We regard interregional commodity flows as the above fish. Trade between regions spreads uniformly at random in a certain geographic area. This assumption is slightly different from the reality of interregional commodity flows, but it is set to contrast with the following random variables.

Second, we assume that interregional commodity flows concentrate in some geographical areas while random variables mainly concentrate in some ranges. This is somewhat close to the reality of interregional commodity flows. The commodity flows between neighboring provinces might become larger than those between provinces located at very far distances. In this case, in what form does a random variable distribute? We check the interregional flow data of railway transportation (railway OD table) in China and the interregional transactions in the interregional input-output table of Japan, which is in the frequency distribution, shown in Figures 1 and 2, respectively. From these figures, we find that (1) the form of distribution is very similar regardless of volume or amount; China or Japan, (2) transactions of small amounts (volume) are mostly in interregional trade, whereas those of large amounts (volume) are relatively few; and (3) the places of transactions become fewer or decrease in an exponential manner as the transaction amounts (volume) increase. Here, the random variable is assumed to be exponential distribution.

Insert Figure 1 and 2

Figure 1. Distribution of Trade between Provinces in China (2007)

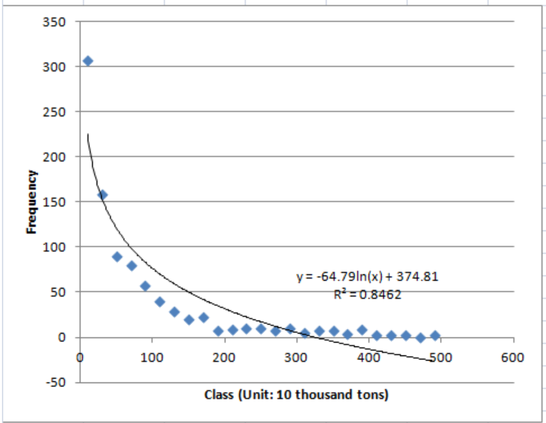
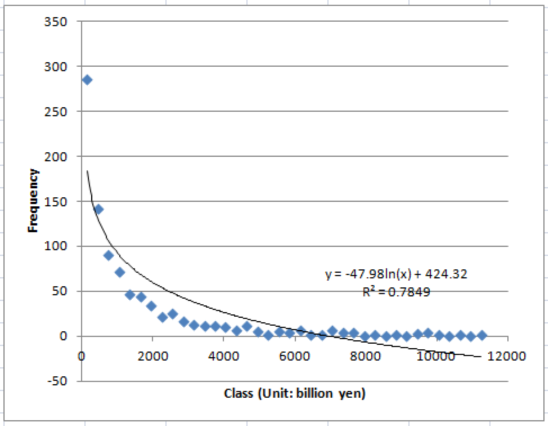


Figure 2. Distribution of Trade between Prefectures in Japan (2005)



3-2 Distance

In China, the outflow of goods data by region can be obtained from statistical yearbooks, but we would not know where the goods go. In fact, Okamoto and Zhang (2003) and Okamoto et al. (2005) used these data to estimate the interregional commodity flow matrix with a supply-constrained gravity model. This estimated commodity flow matrix was used as the Transport Distribution Index to estimate the Q parameter.

Here, we would like to clarify and develop the estimation methods of Okamoto and Zhang (2003) and Okamoto et al. (2005). In these studies, the outflow of commodities in region *r* is assumed to be distributed to each region on the basis of the demand pool of region *s* and the distance between regions *r* and *s*. Therefore, a supply-constrained gravity model was used as follows:

$$h_i^{rs} = AY_i^r Y_i^s (d^{rs})^{\alpha} \tag{9}$$

$$\text{constraint} \;\; h_i^{r0} = \sum_s h_i^{rs}$$

$$ad_i = \frac{d_i^{rs} h_i^{rs}}{h_i^{r0}}$$

where Y_r is the total outflow of region *r;* Y_s is the demand pool of region *s*; ad_i gives the average distance of good *i*, which is obtained from the statistics data; and α is determined by holding a constraint condition. Using the interregional commodity data estimated from this supply-constrained gravity model, Okamoto and Zhang (2003) and Okamoto et al. (2005) estimated the Q parameter by commodity.

However, this estimation is the same as the Transport Distribution Index of distance. If we substitute (8) with (9), we get the following relation:

$$Q_i^{rs} = \frac{AY_i^r Y_i^s (d^{rs})^{\alpha}}{\frac{\sum_r AY_i^r Y_i^s (d^{rs})^{\alpha} \sum_s AY_i^r Y_i^s (d^{rs})^{\alpha}}{\sum_r \sum_s AY_i^r Y_i^s (d^{rs})^{\alpha}}}$$

$$= \frac{AY_i^r Y_i^s (d^{rs})^{\alpha}}{\frac{AY_i^r Y_i^s \sum_r (d^{rs})^{\alpha} \, AY_i^r Y_i^s \sum_s (d^{rs})^{\alpha}}{AY_i^r Y_i^s \sum_r \sum_s (d^{rs})^{\alpha}}}$$

$$= \frac{(d^{rs})^{\alpha}}{\frac{\sum_r (d^{rs})^{\alpha} \sum_s (d^{rs})^{\alpha}}{\sum_r \sum_s (d^{rs})^{\alpha}}} \tag{10}$$

As we know, the Q parameter is determined on the distance *d* and *α*, independent of the outflow of commodity and the demand pool of each region. Based on this, we consider the Q parameter of the LSG model as the Transportation Distribution Index, and it is not necessary to use the commodity outflow or demand pool data of each province published by the Chinese authorities.

4. Empirical Test

4-1 Data and comparison method

In case of lack of data on the “real” or “true” interregional commodity flow, we can consider the effectiveness of the result estimated using the alternative data proposed above. In order to estimate the Q parameter, we prepare the following alternative data in line with the previous section:

(1) Two types of random variables (with uniform distribution and exponential distribution)

(2) Two types of physical distance (based on railway and road transport)

To examine the effectiveness of the estimated data, we construct a simple MRIO model and compare it with the survey-based input-output table. However, we have to overcome two problems.

First, we have to construct the MRIO model. As seen in equation (3), we can make an impact analysis if we get the technical coefficient A of each province and the interregional trade coefficient matrix T. So, in order to examine the model effectiveness, it would be enough to estimate the Leontief inverse (multiplier), with no need to estimate the full MRIO data set. Although it would be better if we could get the technical coefficient matrix A of each province, we can use the technical coefficient A of the 2007 national input-output table for each province and construct a one-sector model in order to evaluate the interregional coefficient matrix estimated by our method. We can then calculate the total output of these two models and check the overall percentage error (OPE) established by Miller and Blair (2009).

Second, we do not have a survey-based interregional input-output table with regard to China. Further, we are not able to compare a non-survey model with a survey-based one. Therefore, we can evaluate the accuracy of the estimated model using the following two methods. First, the MRIO model estimated from the 2007 railway OD table across the provinces published by the Chinese authorities is assumed to be a survey-based input-output table, so we can compare it with our MRIO model estimated from random variables or distance. Second, we calculate the total output induced by one unit final demand of each province (the row sum of the Leontief inverse) and compare it with the Gross Regional Product (GRP) of each province. We check the correlation coefficient of the two data sets. Based on the interregional input-output table in Japan, the total output of each prefecture induced by one unit final demand $(I - TA)^{-1}i$ and the GRP data of each province are highly correlated[2].

For comparison of the estimated and original data sets, we will study various aspects of evaluation such as the correlation coefficient and the OPE established by Miller and Blair (2009), as mentioned above. The standard total percentage error (STPE) and mean absolute difference (MAD) developed by Lahr (2000) will be used in the modified version as follows:

$$STPE = 100\frac{\sum_i \left| X_i - \hat{X}_i \right|}{\sum_i X_i}$$

[2] This idea came from Dr. Pongsun Bunditsakulchai (Central Research Institute of Electric Power Industry) at the 24th Annual Meeting of the Applied Regional Science Conference held in Nagoya University on Dec 4–5, 2010.

$$MAD = 100\frac{\sum_i \left|X_i - \hat{X}_i\right|}{n^2}$$

STPE gives the output error that will be induced by final demand and MAD, the output error per region.

4-2 Empirical test: random variable

In case the interregional commodity flows are regarded as random variables, we have to generate two types of random variables, one with uniform distribution and the other with exponential distribution, and these should be regarded as the origin-destination table; the Q parameter was estimated as the Transport Distribution Index in equation (8), and then compiled as the MRIO model.

Random variables generate different values each time. So, we generate 30 different values and calculate the total output of each MRIO model, using 30 different random variables. We then compare their mean and standard deviations with the original values (total output induced by one unit final demand of each province in the MRIO model with the railway OD table).

Insert Table 1

When we consider the Q parameter as the Transport Distribution Index of a random variable, the uniform distribution is 3.186 and exponential distribution 3.078 in terms of the average of total output induced by final demand. The correlation coefficient of our data estimated by uniform distribution with both GRP and the railway transport data is higher than that estimated by exponential distribution. Our data estimated by uniform distribution fits better in terms of both STPE and MAD.

However, interregional commodity flows might be overestimated in a uniform distribution, and the multiplier estimated by uniform distribution might become larger. The multiplier estimated by exponential distribution is 3.078, and it is the same for the data estimated by railway transport data. In view of the small difference between the results of the two random variables, we can conclude that the Q parameter of the LSG model has a small impact on estimating the MRIO model. It would be proper to use both the random variables, but it would be better to use the random variables with exponential distribution, because it would give a good estimation result of the total multiplier, and the exponential distribution might be regarded as nearer the real situation of interregional commodity flows.

Table 1 Total output induced by a unit of final demand for each sector (random variables)

	GDP (2007)	Railway OD Table (a)	Uniformed Distribution			Exponential Distribution		
			Average	Difference from (a)	Standard Deviation	Average	Difference from (a)	Standard Deviation
Beijing	9353	3.413	3.261	-0.153	0.148	3.257	-0.157	0.245
Tianjin	5050	2.339	2.228	-0.111	0.090	2.190	-0.150	0.147
Hebei	13710	4.480	4.396	-0.085	0.205	4.228	-0.253	0.300
Shanxi	5733	2.542	2.417	-0.125	0.086	2.389	-0.153	0.158
Inner Mongoria	6091	2.621	2.512	-0.109	0.081	2.422	-0.199	0.160
Liaoning	11023	3.179	3.748	0.569	0.236	3.604	0.426	0.238
Jilin	5285	2.252	2.282	0.029	0.083	2.248	-0.004	0.125
Heilongjiang	7065	2.614	2.735	0.121	0.104	2.647	0.032	0.192
Shanghai	12189	3.983	4.040	0.057	0.204	3.968	-0.015	0.295
Jiangsu	25741	7.590	7.169	-0.421	0.544	6.650	-0.940	0.569
Zhejiang	18780	4.896	5.596	0.700	0.279	5.403	0.507	0.319
Anhui	7364	3.031	2.843	-0.189	0.141	2.811	-0.220	0.200
Fujian	9249	2.892	3.278	0.386	0.164	3.201	0.309	0.263
Jiangxi	5500	2.070	2.345	0.274	0.076	2.304	0.233	0.116
Shandong	25966	8.041	7.368	-0.674	0.393	6.846	-1.195	0.521
Henan	15012	5.074	4.771	-0.302	0.217	4.468	-0.605	0.289
Hubei	9231	3.164	3.303	0.139	0.160	3.237	0.073	0.274
Hunan	9200	3.008	3.258	0.249	0.143	3.120	0.112	0.193
Guangdong	31084	6.761	8.605	1.844	0.547	7.987	1.226	0.583
Guangxi	5956	2.119	2.461	0.343	0.091	2.481	0.362	0.188
Hainan	1223	1.000	1.296	0.296	0.020	1.295	0.295	0.031
Chongqing	4123	1.909	1.997	0.088	0.081	1.992	0.083	0.115
Sichuan	10505	3.207	3.665	0.458	0.203	3.529	0.322	0.267
Guizhou	2742	1.599	1.677	0.078	0.042	1.674	0.075	0.062
Yunnan	4741	1.955	2.161	0.206	0.075	2.128	0.173	0.092
Tibet	342	1.065	1.084	0.019	0.006	1.088	0.023	0.009
Shaanxi	5466	2.365	2.341	-0.024	0.088	2.305	-0.060	0.110
Gansu	2702	1.759	1.652	-0.107	0.050	1.686	-0.073	0.072
Qinghai	784	1.327	1.191	-0.137	0.012	1.190	-0.137	0.017
Ningxia	889	1.234	1.215	-0.018	0.013	1.218	-0.015	0.024
Xinjiang	3523	1.924	1.859	-0.065	0.049	1.850	-0.074	0.103
Average		3.078	3.186	0.108	0.149	3.078	0.000	0.202
STPE				8.48%			8.91%	
MAD				27.01%			27.42%	
Correlation Coefficient (1)			0.9731			0.9716		
Correlation Coefficient (2)		0.9736	0.9998			0.9991		

Note: Correlation Coefficient (1) is correlation with (a) and (2) with GDP.

4-3 Empirical test: distance

We have to prepare the distance data for both railway and road transport for estimating the Q parameter of equation (10). The distance by railway transport is calculated using the data of China

Info (http://www.china.co.jp/, author accessed 2010/4/22). However, we estimate the distance from each provincial capital to Lhasa (Tibet) and Haikou (Hainan) on the assumption that railway transportation would be made to Lhasa through Xining and to Haikou through Guangzhou. The distance of road transport is estimated as the distance between the provincial capitals from the data of China Highway Information Service (http://www.chinahighway.gov.cn/roadInfo/indexNew.do, author accessed 2010/7/13).

As for the internal distance within provinces, Head and Mayer (2002) discuss several methods of estimation, like (1) fractions of distances to the centers of neighboring regions, (2) area-based measures trying to capture the average distance between producers and consumers located in given territories, and (3) sub-unit-based weighted average methods using actual data on the spatial distribution of economic activity within countries. In this paper, we apply the area-based measures method (2), $d_{ii} = \frac{2}{3}\sqrt{\frac{area}{\pi}}$, to estimate the internal distances within provinces, based on Koshizuka (1978).

The other point with regard to estimating the Transport Distribution Index of distance is how to estimate parameter α, which shows the decayed function with distance: when the distance is more, transportation would become less. We estimate the gravity model by using the railway OD table between regions; the result is shown in Table 2.

Insert Table 2

Table 2 Result of the Gravity model of railway OD table

	Per capita GDP	GDP	Freight Volume
Constant (A)	6.879***	-5.229***	-2.660***
	(0.0001)	(0.0000)	(0.0000)
Distance	-0.784***	-0.458***	-0.524***
	(0.0000)	(0.0000)	(0.0000)
Outflow of region (push factor)	0.011	1.521***	1.214***
	(0.9204)	(0.0000)	(0.0000)
Inflow of region (pull factor)	0.245**	1.713***	1.718***
	(0.0223)	(0.0000)	(0.0000)
Coefficient of determination (R2)	0.125	0.413	0.618
Number of samples (N)	961	961	961

Note: The figures in parentheses are P values. "**" indicates 5% significance and "***" indicates 1% significance.

In the gravity model estimation, we use three types of supply (push factor) and demand (pull factor) pools: per capita GRP, GRP, and cargo volume. The results show that the cargo volume type fits best because the railway OD table of China is shown in cargo volume, and parameter α is between -0458 and -0.784.

Insert table 3

Table 3 Total output induced by a unit final demand for each sector (the transport distribution coefficient of distance)

	GDP (2007)	Railway OD Table (a)	Raiway Distance			Road Distance		
			α=-0.3	α=-0.4	α=-0.5	α=-0.1	α=-0.2	α=-0.3
Beijing	9353	3.413	3.114	3.078	3.033	3.178	3.166	3.147
Tianjin	5050	2.339	2.136	2.111	2.080	2.176	2.169	2.157
Hebei	13710	4.480	4.156	4.151	4.155	4.203	4.197	4.186
Shanxi	5733	2.542	2.309	2.303	2.300	2.333	2.326	2.317
Inner Mongoria	6091	2.621	2.388	2.380	2.375	2.408	2.393	2.378
Liaoning	11023	3.179	3.530	3.504	3.473	3.561	3.548	3.538
Jilin	5285	2.252	2.233	2.225	2.213	2.225	2.216	2.208
Heilongjiang	7065	2.614	2.645	2.633	2.616	2.636	2.623	2.613
Shanghai	12189	3.983	3.912	3.914	3.892	3.881	3.910	3.930
Jiangsu	25741	7.590	7.199	7.253	7.295	7.087	7.149	7.200
Zhejiang	18780	4.896	5.512	5.545	5.574	5.435	5.476	5.509
Anhui	7364	3.031	2.729	2.729	2.729	2.735	2.746	2.752
Fujian	9249	2.892	3.112	3.086	3.055	3.165	3.168	3.171
Jiangxi	5500	2.070	2.280	2.274	2.265	2.287	2.286	2.283
Shandong	25966	8.041	7.241	7.303	7.368	7.102	7.129	7.147
Henan	15012	5.074	4.585	4.612	4.639	4.507	4.502	4.494
Hubei	9231	3.164	3.151	3.154	3.161	3.151	3.142	3.131
Hunan	9200	3.008	3.124	3.123	3.128	3.145	3.138	3.132
Guangdong	31084	6.761	8.328	8.386	8.469	8.297	8.356	8.448
Guangxi	5956	2.119	2.333	2.317	2.304	2.380	2.370	2.362
Hainan	1223	1.000	1.274	1.271	1.268	1.284	1.283	1.282
Chongqing	4123	1.909	1.930	1.925	1.926	1.949	1.937	1.927
Sichuan	10505	3.207	3.393	3.385	3.385	3.413	3.378	3.351
Guizhou	2742	1.599	1.620	1.615	1.610	1.631	1.622	1.614
Yunnan	4741	1.955	2.079	2.071	2.066	2.097	2.089	2.084
Tibet	342	1.065	1.077	1.076	1.075	1.078	1.076	1.075
Shaanxi	5466	2.365	2.253	2.242	2.231	2.264	2.250	2.238
Gansu	2702	1.759	1.599	1.587	1.576	1.618	1.605	1.593
Qinghai	784	1.327	1.176	1.174	1.172	1.180	1.176	1.174
Ningxia	889	1.234	1.195	1.191	1.188	1.203	1.199	1.194
Xinjiang	3523	1.924	1.803	1.797	1.793	1.806	1.792	1.781
Average		3.078	3.078	3.078	3.078	3.078	3.078	3.078
STPE			8.31%	8.25%	8.26%	8.56%	8.54%	8.62%
MAD			25.56%	25.39%	25.41%	26.33%	26.28%	26.54%
Correlation Coefficient (1)			0.9757	0.9759	0.9755	0.9739	0.9736	0.9726
Correlation Coefficient (2)		0.9736	0.9997	0.9995	0.9993	1.0000	0.9999	0.9997

Note: Correlation Coefficient (1) is correlation with (a) and (2) with GDP.

Based on this result, we move parameter α from -0.1 to -0.7 by 0.1 point and estimate the transport distribution coefficient of distance, and construct the MRIO model. As there is no "real" OD table for road transport, we compare it with the estimated model by using the railway OD table.

The results are shown in Table 3.

The Transport Distribution Index of distance depends on parameter α—how much volume of the commodity would be transported, and how far. The larger parameter α is, the more distance the commodity is transported. If parameter α is small, it means that the commodity would be transported to a nearby region, not to a far away region. So, it is very common that parameter α for road transport is smaller than that for railway transport, because the average transport distance by road is shorter, according to the transportation statistics of China.

From Table 3, the average total output induced by one unit of final demand of each province (the row sum of the Leontief inverse) is 3.078, which is the same figure as for the model estimated using the railway OD table data. Judging from the STPE and MAD indexes, -0.4 for parameter α of railway distance and -0.2 for parameter α of road transport distance fit better than other α parameter values.

As for the correlation coefficient, it fits better when parameter α is increased. It is quite natural that the correlation coefficient (1) of the distance by railway compared with the railway OD table model is higher than the correlation coefficient of the distance by road. As a result, we can conclude that the estimated result would be acceptable for constructing the MRIO model.

5. Conclusion

In this paper, we discussed how to estimate an interregional input-output model with limited data, and we proposed a model that is both easy and practical: the Chenery-Moses-type MRIO model with the interregional trade coefficient stipulated by the LSG model, in which the Q parameter is estimated with random variables and distance data.

We evaluated the model using random variables with uniform/exponential distribution and railway/road distance. The results show that we can estimate a reliable interregional input-output model with the transport distribution coefficient of distance under the situation of no data of interregional transactions. However, this does not mean that we can obtain accurate figures in each interregional transaction. It only means that it is possible to construct an interregional input-output model with so-called holistic accuracy (Jensen 1980).

Even in the absence of data on the distance among the regions, the results suggest that we can construct a relatively reliable interregional input-output model with random variables. This means that the total output plays an important role in estimating interregional commodity flows in an LSG model.

However, there is no economic sense in using random variables. We have to consider the distance and its exponent reflected in the spatial character of the region in order to construct a model with higher accuracy and economic sense.

Acknowledgement

I would like to thank Dr. Pongsun Bunditsakulchai (Central Institute of Electric Power) for his useful comments on the preliminary version of this paper at the 24th Applied Regional Science Conference. This study was supported by KAKENHI, Grant-in-Aid for Scientific Research (C) 21530233 and 22560538 of the Japan Society for the Promotion of Science (JSPS).

References

Akita, T., K. Kawamura, and B. Xie (1999), The Regional Economic Development of Northeast China: An Interregional Input-Output Analysis, *The Journal of Econometric Study of Northeast Asia*, 1(1), pp. 53-78.

Ando, A. and T. Shibata (1996), A Multi-Regional Model for China Based on Price and Quantity Equilibrium, in Chatterji, M. (ed.), *Regional Science Perspectives for the Future*, Macmillan, London, Chap.23.

Chenery, H.B. (1953), Regional Analysis, in Chenery, H.B., Clark, P.G., and Pinna, V.C. (eds.), *The Structure and Growth of the Italian Economy*, US Mutual Security Agency, Rome, Italy.

Head, K. and T. Mayer (2002), Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade, CEPII, Working Paper No 2002-01.

Ihara, T. (1996), Regional Economic Analysis (Chiki no Keizai Bunseki), *Chuo Keizai Shuppansha*. (in Japanese).

Ichimura, S. and H. Wang (eds.) (2003), *Interregional Input-Output Analysis of the Chinese Economy*, Singapore: World Scientific Publishing Co.

Institute of Developing Economies-JETRO (2003), *Multi-regional Input-Output Model for China 2000*, Statistical Data Series No.86, Chiba: Institute of Developing Economies-JETRO.

Isard, W. (1951), Interregional and Regional Input-Output Analysis: A model of a Space Economy, *Review of Economics and Statistics*, 33, pp. 318-28.

Jensen, R.C. (1980), The Concept of Accuracy in Regional Input-Output Models, *International Regional Science Review*, 5, pp. 139-54.

Koshizuka, T. (1978), On the Random Distance Within an Area (Chiki nai Kyori), *Journal of Operations Research Society of Japan*, 21, 302-309 (in Japanese).

Li, S. (eds.) (2010), *2002 China Regional Extended Input-Output Model (2002 Nian Zhongguo Diqu Kuozhan Touru Chanchubiao Bianzhi yu Yingyong)*, Beijing: Economic Science Press (in Chinese).

Leontief, W. and A. Strout (1963), Multi-regional Input-Output Analysis, in Barna, T. (eds.), *Structural Interdependence and Economic Development*, St. Martin's Press, New York.

Miller, R.E. and P.D. Blair (2009), *Input-Output Analysis - Foundations and Extensions Second Edition*. Cambridge University Press. New York.

Miyagawa, K., Wang, Z., Hu, Z., Shimizu, M., Arai, M., Ishida, K., et al. (2008), *Analysis of Regional Industrial Structure in China*, Tokyo, Keio University Press (in Japanese).

Moses, L.N. (1955), The Stability of Interregional Trading Patterns and Input-Output Analysis, *American Economic Review* 45(5), pp. 803-832.

Okamoto, N. and Y. Zhang (2003), Compilation Procedure of Multi-regional Input-Output Model for China 2000, *Multi-regional Input-Output Model for China 2000*, Institute of Developing Economies-JETRO (eds.), Institute of Developing Economies-JETRO.

Okamoto, N., Zhang, Y., Hioki, S., Kanazawa, T., and K. Zhao (2005), A Method for Constructing an Interregional Input-Output Model of China for 2000, *The Journal of Econometric Study of Northeast Asia*, 5(2), pp. 23-36.

Okuda, T., Tanekura, F., Hatano, T., and S. Qi (2004), An Estimation of MRIO in China and the Analysis, *Proceedings of Infrastructure Planning*, No.21, pp. 247-254 (in Japanese).

Polenske, K.R. (1970), An Empirical Test of Interregional Input-Output Models: Estimation of 1963 Japanese Production, *American Economic Review* 60(1), pp. 76-82.

State Information Center (2005), *Multi-regional Input-Output Model for China (Zhongguo Quyujian Touruchanchubiao)*, Beijing: Social Sciences Academic Press (in Chinese).

Wang, Z. (2003), The Estimation of Interregional Trade Coefficient in China—Towards Compilation of 1997 Interregional Input-Output Table (Chugoku no Chikikan Koueki Keisu no Suikei), *Quarterly Report of Economics (Keizai Kihou, Economic Society of Rissho University)*, Vol. 52, No. 3/4, pp. 45-62 (in Japanese).