Explanation of the Estimation Procedure for the Interregional Input–Output Database for China
(Version 1.0, June, 2012)

Nobuhiro Okamoto, Daito Bunka University

Abstract
The database containing interregional input–output tables for China has been released so that it might contribute to analytical research on spatial structure in China. The data are compiled through nonsurvey methods to be consistent with the national input–output table. This paper presents the estimation procedure used.

Keywords: Interregional input–output table, Noncompetitive import type, Leontief–Strout gravity model.

1. Introduction
The priority of the Hu Jintao administration has been to resolve problems associated with rapid economic development—to create, in his words, a "harmonious society" ("He Xie She Hui"). Although there have been many problems, regional disparities have been especially challenging. "The western regions development" program was initiated in 1999 under the Jiang Zemin administration. Further, Hu jintao succeeded the inland region development policy and headed its regional extensions, "the promotion of the northeastern economy" and "the excavation of the central region," in 2002 and 2006, respectively. Then, the central theme of the 11th Five-Year Plan, which had started in 2006, focused on the balanced development of regional economies.

Input–output data constitute a very powerful tool in analyzing regional development and interregional interdependence. For this reason, I developed the China Interregional Input–Output Database, the data of which were captured through nonsurvey methods that had been discussed and validated in previous studies.

In this brief paper, I would like to introduce, as a reference for researchers, how to construct a database. Briefly, these interregional input–output tables are compiled from the national input–output table, using regional data and the interregional trade coefficient, the latter of which is estimated through the Leontief–Strout gravity model.

2. Estimation Procedure
2.1 Basic Framework

With regard to defining regions, 31 regions have been defined at the provincial level, such as the municipalities, provinces, and autonomous region not including Hong Kong, Macau, and Taiwan. Chongqing was once a city in Sichuan province, but was upgraded to a municipality in 1997; nonetheless, Chongqing is treated here as an independent region, like other provinces, since it has data available since the year 1987.

Sectors are classified as being of the primary industry (agricultural sector), secondary industry (manufacturing sector with mining, construction and electricity, water, or gas supply), or tertiary industry.

In China, the total sum of each provincial gross regional product (GRP) has exceeded the national gross domestic product (GDP) by around 7–8%, and regional data is not as reliable as national data. Thus, in the estimation procedure, it is essential that I divide the national input–output data at the regional level, based on regional data.

An image from the database is provided below, in Figure 1.

Figure 1. Layout of the Interregional Input–Output Database for China

			Intermediate Demand (Z)							Final Demand (Y)									
		code	Beijing 1	Tianjin 2		s		Xinjiang 31	Beijing 1	Tianjin 2		s		Xinjiang 31	Export E	Total Output X	Statistical Discrepancy SD
Intermediate Input	Beijing	1	Z^{11}	Z^{12}	$\cdots$	Z^{1s}	$\cdots$	Z^{131}	Y^{11}	Y^{12}	$\cdots$	Y^{1s}	$\cdots$	Y^{131}	E^{1}	X^{1}	SD^{1}
	Tianjing	2	Z^{21}	Z^{22}	$\cdots$	Z^{2s}	$\cdots$	Z^{231}	Y^{21}	Y^{22}	$\cdots$	Y^{21}	$\cdots$	Y^{231}	E^{2}	X^{2}	SD^{2}
			$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
		r	Z^{r1}	Z^{r2}	$\cdots$	Z^{rs}	$\cdots$	Z^{r31}	Y^{r1}	Y^{r2}	$\cdots$	Y^{rs}	$\cdots$	Y^{r31}	E^{r}	X^{r}	SD^{r}
			$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	Xinjiang	31	Z^{311}	Z^{312}	$\cdots$	Z^{31s}	$\cdots$	Z^{3131}	Y^{311}	Y^{312}	$\cdots$	Y^{31s}	$\cdots$	Y^{3131}	E^{31}	X^{31}	SD^{31}
Import		M	mi^{1}	mi^{2}	$\cdots$	mi^{s}	$\cdots$	mi^{31}	mf^{1}	mf^{2}	$\cdots$	mf^{s}	$\cdots$	mf^{31}			
Value Added		V	V^{1}	V^{2}	$\cdots$	V^{s}	$\cdots$	V^{31}									
Total Input		X	X^{1}	X^{2}	$\cdots$	X^{s}	$\cdots$	X^{31}									

2.2 Estimation Procedure

The national input–output table for China has a competitive import type of format, as shown in Table 1.

Table 1. Competitive Import Type of Input–Output Table

	Intermediate Output	Final Demand	Export	Import	Total Output
Intermediate Input	z_{ij}^N	y_i^N	e_i^N	$-m_i^N$	x_i^N
Value Added	v_j^N				
Total Input	x_j^N				

I calculate the import ratio from the equation below, which allows us to change it to a noncompetitive import type of table format (see Table 2).

$$s_i = \frac{m_i}{x_i - e_i - m_i} = \frac{m_i}{\sum_j z_{ij} + y_i}$$

Table 2. Noncompetitive Import Type of Input–Output Table

	Intermediate Output	Final Demand	Export	Total Output
Intermediate Input	$(1 - s_i^N)z_{ij}^N$	$(1 - s_i^N)y_i^N$	e_i^N	x_i^N
Import	$mi_j^N = \sum_i s_i^N z_{ij}^N$	$mf_j^N = \sum_i s_i^N y_i^N$		
Value Added	$\mathrm{v}_\mathrm{j}^\mathrm{N}$			
Total Input	$\mathrm{x}_\mathrm{j}^\mathrm{N}$			

I introduce the national technical coefficient, $a_{ij}^N = \frac{(1-s_i^N)z_{ij}^N}{x_j^N}$, and consider it the regional technical coefficient—in line with the result of Okamoto, Zhang, and Zhao (2005)—since the provincial input–output tables for 1987, 1992, 1997, and 2007 are not open.[1] Then, I break out Table 2 at the regional level, by using the interregional trade coefficient t_i^{rs} and regional data.

[1] However, the provincial input–output table for the year of 2002 can be obtained (Li, 2010).

Table 3. Interregional Input–Output Table

	Intermediate Output by Region	Final Demand by Region	Export	Total Output
Intermediate Input, by Region	$t_i^{rs}(1-s_i^N)a_{ij}^N x_j^s$	$t_i^{rs}(1-s_i^N)y_i^N$	$e_i^r = \frac{e^r}{e^N} e_i^N$	$x_i^r = \frac{g_i^r}{g_i^N} x_i^N$
Import	$mi_j^s = \frac{m^s}{m^N} mi_j^N$	$mf_j^s = \frac{m^s}{m^N} mf_j^N$		
Value Added	$v_j^s = \frac{g_j^s}{g_j^N} v_j^N$			
Total Input	$x_j^s = \frac{g_j^s}{g_j^N} x_j^N$			

Here, the interregional trade coefficient holds $\sum_r \sum_i t_i^{rs} = 1$.

The interregional trade coefficient matrix is estimated by using the Leontief–Strout gravity model, according to the result of Okamoto (2011), as follows.

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

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

As I understand it, from the above equation, it is sufficient to estimate the parameter Q to derive data concerning distance and the α parameter.

Next, I prepare the distance data for both railway and road transport. The distance by railway transport is calculated using data from China Info (http://www.china.co.jp/, accessed April 22, 2010). However, I 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. Road-transport distances are estimated as the distances between provincial capitals, using data from the China Highway Information Service (http://www.chinahighway.gov.cn/roadInfo/indexNew.do, accessed July 13, 2010).

As for internal distances 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 used to capture the average distance between producers and consumers located in given territories, and (3) subunit-based weighted average methods that use actual data on the spatial distribution of economic activity within countries. In this paper, I 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 distances of road and railway transport are combined as a weighted average of transport mode share to total transport; here, the figures are 84:16 (road: railway), and this ratio has been quite stable over the years.

As for the α parameter, I apply –0.232 for the year 2007 as a weighted average of road and railway transport, –0.2 for road distance, and –0.4 for railway distance, as per the empirical results of Okamoto (2012). The α parameter for the years 1987, 1992, 1997, and 2002 is estimated based on the assumption that the transport distance increases when changes to the α parameter are made. The estimated α parameter values are shown in Table 4.

Table 4. α Parameter

	Freight Transport Average Distance (km)	Improvement Ratio	Estimated α Parameter
1987	234		–0.512
1992	279	19.2%	–0.435
1997	300	7.6%	–0.404
2002	341	13.6%	–0.348
2007	446	30.7%	–0.232

As for the other estimated data used to generate the interregional input–output table, the total output of each sector, technical coefficient, final demand, export and import, and value added of each province are collected.

As shown in Table 3, in order to calculate regional data, the national value added and the total output (input) by sector are disaggregated and presented in detail by the sectors' GRP of the provincial level. The final demand of each sector is divided by the total GRP owing to the lack of data showing final demand by sector.

The export and import values according to the national input–output table are disaggregated to the provincial level by using information from each province's data.

Although the interregional input–output table uses national input–output data divided by regional information, there is no balance in terms of rows or columns. Thus, the mathematical interaction method—also called the RAS method, which is very commonly used in the input–output literature—was applied for matrix balancing. Remaining discrepancies in terms of rows were allocated as balancing items, under “Statistical Discrepancy” in the table.

3. Conclusion

The information in the database was obtained using nonsurvey methods. It would therefore be useful to capture a more complete picture of the interdependence among the regional economies of China by undertaking analysis through the use of input coefficients and the Leontief inverse. Although all macro-level data—such as final demand, export, and import—were found to be consistent with national-level data, this does not necessarily mean that the transaction values of each cell of the interregional input–output table is accurate. Thus, I can say that accuracy is guaranteed only in terms of a “big picture” analysis.

The database is available as an MS Excel file. I would very much appreciate it if you would inform me of your use of it. Although I hold the copyright for this data, I will make it available for noncommercial or academic use, as long as you acknowledge in your paper that I had compiled the data.

Acknowledgment

This study was supported by KAKENHI, a grant-in-aid for scientific research (C) 21530233 funded by the Japan Society for the Promotion of Science (JSPS).

References

Department of Comprehensive Statistics of National Bureau of Statistics (2005), China Compendium of Statistics 1949–2004. China Statistics Press.

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

Koshizuka, Y. (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. (ed.) (2010), *2002 China Regional Extended Input–Output Model (2002 Nian Zhongguo Diqu Kuozhan Touru Chanchubiao Bianzhi yu Yingyong)*. Beijing: Economic Science Press (in Chinese).

National Bureau of Statistics, *China Statistical Yearbook*, various years.

Okamoto, N., Zhang, Y. and Zhao K (2005), Non-survey Methods for Estimating Regional and Interregional Input–Output Multipliers, *Spatial Structure and Regional Development in China: An Interregional Input–Output Approach*, Okamoto, N. and Ihara, T. (Eds), Palgrave, Hampshire, United Kingdom.

Okamoto, N. (2012), Multi-regional Input–Output Modeling in China: Revisited (Chugoku no Chikikikan Sangyo Renkan Hyo no Kouchiku: Saikentou), *Journal of Applied Regional Science*

(Ouyo Chikigaku Kenkyu), 17, 1–14 (in Japanese).