Assessing the estimation accuracy of LQ methods for regionalization of input coefficients: a case study in Japan

Kazuki Tamesue

Graduate School of Systems and Information Engineering, University of Tsukuba

1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573, Japan

+81-29-853-5572

tamesue.kazuki@sk.tsukuba.ac.jp

Morito Tsutsumi

Faculty of Engineering, Information and Systems, University of Tsukuba

1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573, Japan

+81-29-853-5007

tsutsumi@sk.tsukuba.ac.jp

Abstract: A survey-based technique is regarded as a reliable method for constructing input-output tables; however, it requires huge amounts of time and money. On the other hand, a non-survey technique only requires existing statistics data, and its recent advances and developments are remarkable, especially in location quotient method. The LQ method is found to be a useful and strong tool through some empirical studies, but more empirical evidences are needed to prove its efficiency. Furthermore, different behaviors of parameter δ of the FLQ and AFLQ in existing researches also suggest that more applications and examinations should be carried out. The objective of the present study is to compare accuracies of estimated regional input-output tables with various LQ techniques using nine Japanese regions data.

Keywords: location quotients, regionalization, accuracy assessment, input-output table

1. Introduction

Input-output tables summarizing transactions of goods and services between industries into a single matrix are compiled for various countries around the world owing to their usefulness as a tool for analyzing economic ripple effects and understanding economic structure. In the case of Japan, in addition to the national input-output table that is created every five years, other input-output tables, such as the regional input-output and prefecture input-output tables, target different spatial scales.

When measuring policy effects or forecasting the economy for a specific region, it is desirable to use input-output tables for smaller geographical scales (rather than using national input-output tables) to take into account the actual situation of the target region.

Since 1990, input-output tables have been created for all Japanese prefectures every five years. However, only a few major cities have managed to create their own input-output tables at the municipality level, which is a smaller administrative unit than a prefecture. The main reason for this inability is the huge cost expended in creating an input-output table. Generally, the survey method is used to construct an input-output table for a nation or region. As the name suggests, it uses surveys and primary statistics data. However, due to issues with data availability and lack of information about the method for constructing input-output tables, it is difficult for small local governments to construct their own input-output tables.

Non-survey methods, which use existing statistics to estimate an input-output table by numerical calculations, have attracted the attention of researchers and practitioners for their simplicity and economical aspects. Above all, the location quotient (LQ) method is recognized as a useful and strong tool, and various LQ indicators (introduced in the next section) have been developed to improve the estimation accuracy of regional input coefficients and/or multipliers.

Several empirical studies have assessed and compared the accuracy of LQs, such as those developed for Scotland (Flegg and Webber, 2000), Finland (Flegg and Tohmo, 2013), and Germany (Kowalewksi, 2013). In this regard, the Monte Carlo simulation (Bonfiglio and Chelli, 2008) is also notable. However, no such research currently exists for Japanese input-output tables. The objective of this study is to compare and assess the accuracy of the estimated regional input-output tables using various LQ techniques and to provide empirical evidence using data from the Japanese input-output table.

2. Location quotient method

2.1. Basic LQs

The LQ method assumes that the technological capacities (technological coefficients) of the region and nation are equal; in other words, the amount of an input necessary for producing one unit of a gross product, regardless of the source (Flegg et al., 1995), is the same for the region and the nation. To estimate the regional input coefficients r_{ij} from the national input coefficient a_{ij}, we multiply some adjustment coefficient q_{ij} to the national input coefficients: $r_{ij} = q_{ij} \times a_{ij}$. In the LQ method, we replace the adjustment coefficient by the various LQs that have been developed to date.

The simple LQ (SLQ) is defined as follows:

$$SLQ_i = \frac{RE_i}{NE_i} \times \frac{TNE}{TRE}, \tag{1}$$

where RE_i is the number of regional employees in the supplying sector i, and NE_i is the number of national employees in sector i. TRE and TNE are the regional and national total employment respectively. The SLQ can also be expressed as

$$SLQ_i = \frac{RE_i / TRE}{NE_i / TNE}, \tag{2}$$

which is interpreted as the share of sector i in the region standardized by the national share of sector i. When $SLQ_i < 1$, regional sector i is assumed to have lower productivity than the nation. Therefore, the region would import from other regions in order to satisfy the regional demand. On the other hand, the region is assumed to satisfy the regional demand if $SLQ_i \geq 1$. In this case, the regional coefficients are considered to be equal to the national coefficients, and thus, the SLQ is truncated as $SLQ_i = 1$ when it exceeds 1.

The cross industry LQ (CILQ) is defined as follows:

$$CILQ_{ij} = \frac{RE_i / NE_i}{RE_j / NE_j}, \tag{3}$$

which can also be expressed as $CILQ_{ij} = SLQ_i / SLQ_j$. When $CILQ_{ij} < 1$, the region imports from other regions, since the productivity of the supplying sector i is relatively low compared to that of the purchasing sector j. The CILQ will be truncated as $CILQ_{ij} = 1$ when $CILQ_{ij} \geq 1$. Unlike the SLQ, the CILQ considers both supplying and purchasing sectors. However, because the CILQ always equals 1 when $i = j$, Smith and Morrison (1974) suggested that the two LQs be combined.

$$q_{ij} = \begin{cases} SLQ_i & i = j \\ CILQ_{ij} & i \neq j \end{cases}. \tag{4}$$

2.2. Developments in LQs

Round (1978) noted that the adjustment coefficient q_{ij} should consider the relative sizes of the (I) supplying sector, (II) purchasing sector, and (III) region. Since (RE_i / NE_i) indicates the relative size of the supplying sector i, and (TNE/TRE) indicates the relative size of the region, the SLQ considers (I) and (III) but does not take (II) into account. The CILQ satisfies (I) and (II) but not (III). To take all these variables into account, Round (1978) devised a semilogarithmic quotient called the RLQ.

$$RLQ_{ij} = \frac{SLQ_i}{\log_2(1+SLQ_j)} \ . \tag{5}$$

The RLQ is criticized for underestimating imports from other regions when the size of the region is small (e.g., Flegg and Webber, 1997).

Flegg et al. (1995) introduced Flegg's LQ (FLQ) to overcome the drawbacks of the SLQ and CILQ. It can be seen as a modification of the RLQ and is defined as

$$FLQ_{ij} = \begin{cases} CILQ_{ij} \times \lambda^* & i = j \\ SLQ_i \times \lambda^* & i \neq j \end{cases} \ , \tag{6}$$

where

$$\lambda^* = \left[\log_2\left(1+\frac{TRE}{TNE}\right)\right]^{\delta} \ , \tag{7}$$

and δ is a parameter ($0 \le \delta < 1$) that controls the degree of convexity in equation (7) (Flegg and Webber, 1997). The larger the value of δ, the lower the value of λ^*, which would lead to a large adjustment in local imports. The FLQ is the same as equation (4) when $\delta = 0$.

Flegg and Webber (2000) further developed the FLQ into the augmented FLQ (AFLQ). The AFLQ was devised in response to a comment by McCann and Dewhurst (1998), who pointed out the possibility of regional coefficients exceeding national coefficients ($r_{ij} > a_{ij}$) when there is regional specialization. The AFLQ is defined as follows:

$$AFLQ_{ij} = \begin{cases} FLQ_{ij} \times [\log_2(1+SLQ_j)] & SLQ_j > 1 \\ FLQ_{ij} & SLQ_j \le 1 \end{cases} . \tag{8}$$

When $SLQ_j > 1$, the degree of regional specialization of the purchasing sector increases, and consequently, the allowance of imports decreases (Figure 1).

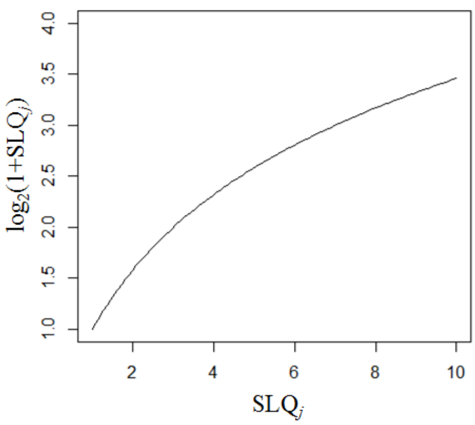

Figure 1. Relationship between the regional specialization and AFLQ

2.3. Empirical evidence

Several studies, including Flegg and Webber (2000), Tohmo (2004), Bonfiglio and Chelli (2008), Lehtonen and Tykkyläinen (2012), Flegg and Tohmo (2013), and Kowalewksi (2013), have examined the estimation accuracy of various LQs. Most reported that the SLQ and CILQ tend to underestimate imports from other regions and that the FLQ and AFLQ outperform the other LQs.

Flegg and Tohmo (2013) observed that the AFLQ is slightly superior to the FLQ, similar to Bonfiglio and Chelli (2008) who used the Monte Carlo simulation. Flegg and Webber (2000) used Scottish input-output data and confirmed that there is no difference between the AFLQ and FLQ in terms of estimation accuracy.

Compared to other LQs, a major difference between the FLQ and AFLQ lies in the parameter δ, the optimal value of which must be estimated empirically. According to Flegg and Webber (2000) and Flegg and Tohmo (2013), the optimal value of δ is 0.20 and 0.25 respectively, implying that the optimal value varies depending on the target data. As Flegg and Webber (1997) stated, more empirical evidence is required to better comprehend the characteristics of the FLQ and AFLQ.

3. Accuracy verification

3.1. Data

The present study uses two input-output tables for 2005: the national input-output table published by the Statistics Bureau of the Ministry of Internal Affairs and Communications, and the inter-regional input-output table published by the Ministry of Economy, Trade and Industry (METI). The latter input-output table divides Japan into nine regions, as shown in Figure 2, in order to create regional input-output tables. For employment data, we use the Establishment and Enterprise Census published by METI for 2006.

The numbers of industry sectors available in the national input-output table are 13, 34, or 108, but the corresponding numbers available in the inter-regional input-output table are 12, 29, or 53. To match the sector classification between these two input-output tables, we aggregate 53 sectors in the inter-regional input-output tables into 34 sectors. Since the classification corresponding to “other public services” in the 34-sector classification is integrated into “medical service, health, social security, and nursing care” in the inter-regional input-output table classification and cannot be separated, we integrate the former sector into the latter in both input-output tables. As none of the classifications in the employment data correspond to “activities not elsewhere classified” in the 34-sector classification, we exclude this sector from the two input-output tables. These adjustments eventually result in 31 sectors in our datasets.

Figure 2. Nine regions of Japan (Source: METI report)

Furthermore, imports from foreign countries in both input-output tables are competitive imports; commodities that are domestically produced and imported from other countries cannot be distinguished from each other. To remove competitive imports from the total transaction tables, we use an approximation method discussed in Miller and Blair (2009). For each sector, we define the degree of self-sufficiency as

$$c_i = 1 - \frac{m_i}{\sum_j z_{ij} + f_i} \quad , \tag{9}$$

where m, z, and f denote the import, intermediate input, and final demand respectively. Then,

Table 1. Descriptive statistics

	Input Coefficients		TRE/TNE	SLQ	
	mean	sd		mean	sd
Hokkaido	0.0093	0.0228	0.0418	0.799	0.551
Tohoku	0.0083	0.0192	0.0743	1.047	0.457
Kanto	0.0120	0.0287	0.4086	1.008	0.207
Chubu	0.0096	0.0245	0.1129	1.131	0.480
Kinki	0.0101	0.0230	0.1642	1.002	0.282
Chugoku	0.0092	0.0259	0.0590	1.062	0.320
Shikoku	0.0074	0.0161	0.0305	0.967	0.456
Kyusyu	0.0095	0.0242	0.0991	0.867	0.341
Okinawa	0.0084	0.0246	0.0096	0.691	0.519

estimates of the domestic input coefficients can be obtained by multiplying equation (9) with technological coefficients, which include imported commodities from abroad: $a_{ij} = c_i \times a_{ij}^*$. The descriptive statistics of input coefficients, region size (*TRE*/*TNE*), and degree of regional specialization (SLQ) of each region are shown in Table 1.

3.2. Results

Following Flegg and Tohmo (2013) and Kowalewksi (2013), this study employs four measures for assessing accuracy:

$$\gamma_1 = \frac{\sum_j \sum_i (\hat{r}_{ij} - r_{ij})^2}{n^2} \quad , \tag{10}$$

$$\gamma_2 = \sqrt{\sum_j \sum_i \left|\hat{r}_{ij} - r_{ij}\right|^2} \quad , \tag{11}$$

$$\gamma_3 = \frac{\sum_j \sum_i \left|\hat{r}_{ij} - r_{ij}\right|}{\sum_j \sum_i r_{ij}} \quad , \tag{12}$$

$$\gamma_4 = \sqrt{\frac{\sum_j \sum_i (\hat{r}_{ij} - r_{ij})^2}{\sum_j \sum_i r_{ij}^2}} \quad , \tag{13}$$

where r_{ij} denotes the actual regional coefficients obtained from the regional input-output table, $\hat{r}_{ij}$ denotes the estimated regional coefficients using LQs, and *n* is the number of sectors. γ_1 is also known as the mean squared error (MSE), which measures the variance of the estimation error. γ_2 is the Euclidean metric difference, and γ_3 is the mean absolute deviation as a percentage of the mean value of r_{ij} (Flegg and Tohmo, 2013). γ_4 is known as Theil's inequality measure (Theil et al., 1966). One of the merits of using γ_1 or γ_4 is that the MSE component in these measures can be decomposed into proportions of bias, variance, and covariance (Jorgenson et al., 1970; Leuthold, 1975; Flegg and Tohmo, 2013):

$$\frac{1}{n^2}\sum_j \sum_i (\hat{r}_{ij} - r_{ij})^2 \equiv \{m(\hat{r}_{ij}) - m(r_{ij})\}^2 + \{sd(\hat{r}_{ij}) - sd(r_{ij})\}^2 + 2(1-c) \times sd(\hat{r}_{ij}) \times sd(r_{ij}) \, , \tag{14}$$

where *m* (•) is the arithmetical mean, *sd* (•) is a standard deviation, and *c* is a correlation coefficient between $\hat{r}_{ij}$ and r_{ij}.

Table 2. Accuracy assessment of LQs

	SLQ	CILQ	RLQ	FLQ	AFLQ
$\gamma_1 \times 10^3$	0.243	0.384	0.307	0.190	0.377
				(δ = 0.226)	(δ = 0.888)
γ_2	0.478	0.597	0.538	0.409	0.576
				(δ = 0.281)	(δ = 0.813)
γ_3	0.538	0.627	0.598	0.485	0.676
				(δ = 0.230)	(δ = 0.614)
γ_4	0.634	0.809	0.726	0.536	0.754
				(δ = 0.279)	(δ = 0.804)

Table 3. Decomposition of the MSE

	$\gamma_1 \times 10^3$	bias	variance	covariance
SLQ	0.243	0.011	0.063	0.169
		(4.7%)	(25.9%)	(69.4%)
CILQ	0.384	0.016	0.152	0.216
		(4.2%)	(39.6%)	(56.2%)
RLQ	0.307	0.015	0.110	0.182
		(4.9%)	(35.8%)	(59.3%)
FLQ (δ = 0.226)	0.190	0.008	0.027	0.162
		(4.3%)	(14.3%)	(85.2%)
AFLQ (δ = 0.888)	0.377	0.036	0.164	0.178
		(9.6%)	(43.4%)	(47.1%)

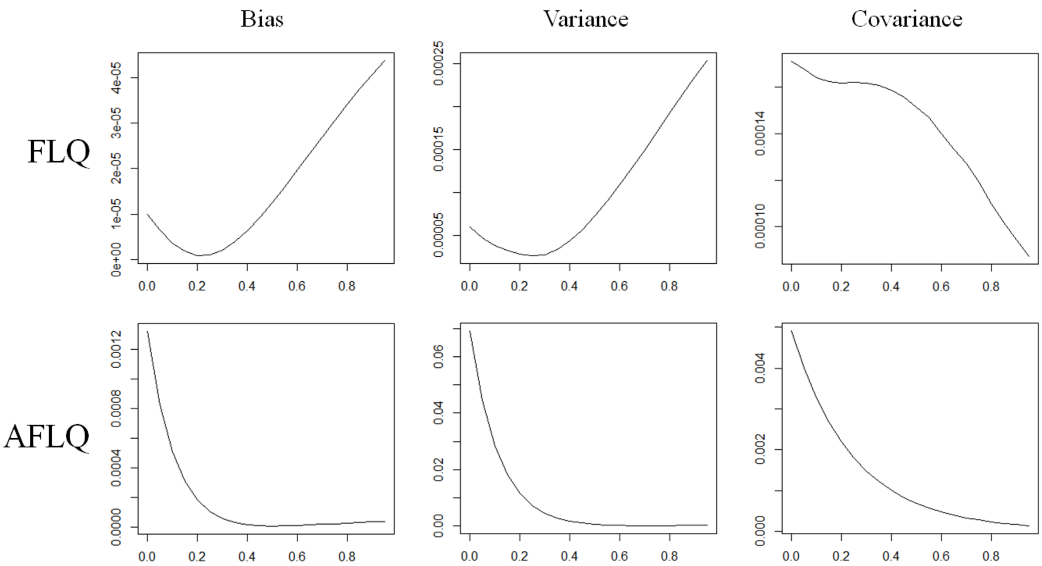

Figure 3. Relationship between MSE components and δ

Table 2 shows the accuracy assessment of LQs with the four previously described measures. For the FLQ and AFLQ, the optimal values of δ are chosen to minimize each criterion. From Table 2, the FLQ has the smallest values in all measures, indicating the best accuracy of all LQs. On the other hand, the AFLQ does not perform as well as the FLQ. The SLQ and RLQ have higher accuracy than the AFLQ for all criteria, and the AFLQ only outperforms the CILQ, which scores the worst in the accuracy assessment. In the order of descending accuracies, the result is as follows: FLQ, SLQ, RLQ, AFLQ, and CILQ. The optimal value of δ in the FLQ and AFLQ varies with each measure, but the value of δ in the FLQ lies within the allowable range of around 0.2 to 0.3, as indicated by abovementioned existing studies. The optimal value of δ in the AFLQ, however, varies more widely (from 0.614 to 0.888), and these values are much larger than those reported by the existing studies.

The results of the decomposition of the MSE are shown in Table 3. In all LQs, the covariance component accounts for the highest percentage of the MSE, and the bias component accounts for the lowest percentage. This indicates that bias is not an important factor in the estimation error. According to Flegg and Tohmo (2013), "the covariance component captures the lack of a perfect correlation between $\hat{r}_{ij}$ and r_{ij} distributions, and the variance component arises when these distributions have different standard deviations." A reason for the superior performance of the FLQ is that it succeeds in reducing all the components of the MSE. In Flegg and Tohmo (2013), the biggest reduction was witnessed in the covariance component. However, the biggest reduction in our study occurs in the variance component. On the other hand, the poor performance of the AFLQ can be attributed to the increase in all the MSE components, the biggest increase being in the bias component.

Figure 3 illustrates the relationship between the parameter δ (x-axis) and each of the components of the MSE (y-axis). The bias and variance components are downward convex, whereas the covariance component is monotonically decreasing. There are some optimal values for δ within the parameter space [0, 1) that minimize the bias and variance components. However, the covariance component decreases as δ increases. Furthermore, for an increasing δ, the rates of change of the covariance component in the FLQ and AFLQ are in direct contrast.

4. Conclusion

The present study applied various LQs to data from Japanese input-output tables, to evaluate and compare the estimation accuracy of LQ techniques. Using four different types of criteria, we showed that the FLQ outperforms other LQs in all criteria. Moreover, in contrast to the evidence reported by existing studies, the AFLQ does not perform as well as expected. The optimal values of δ in the FLQ obtained in this study are similar to those in Flegg and Webber (2000), Bonfiglio and Chelli (2008), and Flegg and Tohmo (2013), implying the robustness of the FLQ. On the contrary, the optimal value of δ in the AFLQ is much larger than the results from previous studies, indicating that it is frail compared to the FLQ.

A detailed factor analysis of the estimation error is needed to provide insight into the further development of regionalization techniques for input-output tables. In response to McCann and Dewhurst's (1998) observation that regional coefficients depend on the spatial distribution of economic activities and interindustry linkage, it would be worthwhile to consider the geographical location of each region and the spatial distribution of industries.

In this study, we focused only on accuracy among LQ techniques. However, other non-survey methods also exist; Kronenberg (2009) and Nakano and Nishimura (2013) considered the existence of cross-hauling to estimate imports from other regions. Comparing accuracies between these methods is desirable to examine the pros and cons of each non-survey method.

References

Bonfiglio, A., and Chelli, F.: Assessing the behavior of non-survey method for constructing regional input-output tables through a Monte Carlo simulation, *Economic Systems Research*, Vol.20, pp.243–258, 2008.

Flegg, A.T. and Tohmo, T.: Regional input-output tables and the FLQ formula: A case study of Finland, *Regional Studies*, Vol.47, pp.703–721, 2013.

Flegg, A.T., Webber, C.D.: On the appropriate use of location quotients in generating regional input-output tables: reply, *Regional Studies*, Vol.31, pp.795–805, 1997.

Flegg, A.T. and Webber, C.D.: Regional size, regional specialization and the FLQ formula, *Regional Studies*, Vol.34, pp.563–569, 2000.

Flegg, A.T., Webber, C.D., and Elliott, M.V.: On the appropriate use of location quotients in generating regional input-output tables, *Regional Studies*, Vol.29, pp.547–561, 1995.

Jorgenson, D.W., Hunter, J., and Ishag Nadiri, M.: The predictive performance of econometric models of quarterly investment behavior, *Econometrica*, Vol.38, pp.213–224, 1970.

Kowalewksi, J.: Regionalization of national input-output tables: Empirical evidence on the use of the FLQ formula, *Regional Studies*, DOI:10.1080/00343404.2013.766318, 2013.

Kronenberg, T.: Construction of regional input-output tables using nonsurvey methods: the role of cross-hauling, *International Regional Science Review*, Vol.32, pp.40–64, 2009.

Lehtonen, O. and Tykkyläinen, M.: Estimating regional input coefficients and multipliers: Is the choice of a non-survey technique a gamble?, *Regional Studies*, DOI:10.1080/00343404.2012.657619, 2012.

Leuthold, R.M.: On the use of Theil's inequality coefficients, *American Journal of Agricultural Economic*, Vol.57, pp.344–346, 1975.

McCann, P. and Dewhurst, J.H.L.: Regional size, industrial location and input-output expenditure coefficients, *Regional Studies*, Vol.32, pp.435–444, 1998.

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

Nakano, S. and Nishimura, K.: A nonsurvey multiregional input-output estimation allowing cross-hauling: partitioning two regions into three or more parts, *The Annals of Regional Science*, Vol.50, pp.935–951, 2013.

Riddington, G., Gibson, H., and Anderson, J.: Comparison of gravity model, survey and location quotient-based local area tables and multipliers, *Regional Studies*, Vol.40, pp.1069 – 1081, 2006.

Round, J.I.: An interregional input-output approach to the evaluation of nonsurvey methods, *Journal of Regional Science*, Vol.18, pp.179–194, 1978.

Smith, P. and Morrison, W.I.: *Simulating the Urban Economy: Experiments with Input-Output Techniques*. Pion, London, 1974.

Theil, H.: *Applied Economic Forecasting*, North-Holland Publishing Company, Amsterdam, 1966.

Tohmo, T.: New developments in the use of location quotients to estimate regional input-output coefficients and multipliers, *Regional Studies*, Vol. 38, pp.43–54, 2004.