Int. Journal of Economics and Management 20 (1): 97-109 (2026)

IJEM
International Journal of Economics and Management
Journal homepage: http://www.ijem.upm.edu.my

The Construction of Regional Input-Output Tables for Selangor Using Three Non-Survey Methods

MUHAMMAD DAANIYALL ABD RAHMAN[a*],
MUHAMMAD FIKRI IQMAL ABDUL NAJID[a], CHAKRIN UTIT[a],
FUTU FATURAY[b] AND SITI NURZAHIRA CHE TAHRIM[c]

[a]*School of Business and Economics, Universiti Putra Malaysia, Malaysia*
[b]*Fiscal Policy Agency, Ministry of Finance, Indonesia*
[c]*Selangor Research Institute, Malaysia*

ABSTRACT

Regional input-output table (RIOT) has been widely used to analyze impacts at the sub-national level. Although RIOT plays an extensive role in providing insights into regional development policies, it is rarely available for public use, particularly in developing countries. In Malaysia, Selangor is the most developed state and contributes nearly one-third to the national GDP. Although understanding inter-industry linkages is crucial for optimizing regional development, the absence of an input-output table for Selangor has limited the ability to identify key and high-value sectors. This study aims to estimate a Selangor RIOT using three methods, namely, the Simple Location Quotient (SLQ), the RAS technique, and the Cross Entropy (CE) method, and to assess the extent to which the estimated tables align with their analytical outcomes. The findings indicated that although the RAS and CE methods perform better than the SLQ method across all distance measures. All three estimates are relatively close. Furthermore, the statistical analysis demonstrated that the estimated Selangor RIOT from these methods produced smaller differences and high correlations, implying statistical accuracy and consistency.

JEL Classification: C67, D57, R10

Keywords: Input-output; Regionalization; Non-survey technique

Article history:
Received: 16 July 2025
Accepted: 12 September 2025

[*] Corresponding author: Email: daaniyall@upm.edu.my
DOI: http://doi.org/10.47836/ijeam.20.1.08

INTRODUCTION

Regional input-output table (RIOT) is an essential tool for understanding economic linkages within a sub-national economy, such as a district, state, or province. RIOT provides useful insights into supporting targeted sub-national development policies (Miller and Blair, 2009; Lahr and Mesnard, 2004), including inter-industry relationships, production and consumption structures, and multiplier effects, each of which is crucial for spatial-based planning (Huang et al., 2015; Patandianan and Shibusawa, 2020), industrial policy (Liu, 2019; Juhász et al., 2024) and environmental assessment, such as regional environmental footprints (Jensen et al., 2011; Islam et al., 2021). While their value for regional analysis is exceptional, RIOT is not readily available, where most national statistical offices commonly publish a national input-output table (IOT) aligned with the periodic economic census and resource availability (Kowalewksi, 2015).

For many developing countries, including Malaysia, the unavailability of survey-based and official RIOT leaves policymakers uninformed on regional economic strength. For example, Selangor, as the most industrialized and economically dominant state in Malaysia, which contributes nearly 30% of the national gross domestic product (GDP), has no published RIOT. Recently, the state government formulated the state's first five-year economic plan called Rancangan Selangor Pertama (RS-1), focusing on key sectors. Despite its importance for the national economy and the determination of key sectors, the absence of the RIOT presents a major obstacle for evidence-based regional planning and policy evaluation. Information on the state's unique industrial structures, economic interactions, and sectoral multipliers is mostly absent.

Researchers have developed various "regionalization" approaches to estimate RIOT from national data. The formation of the RIOT can be either as a single region (Utit et al., 2020) or in a multi-regional setting (Zheng et al., 2022). Each of the approaches has its own methodological characteristics. The regionalization approaches or so-called non-survey methods offer practical alternatives where detailed regional data is lacking (Flegg and Tohmo, 2013; Flegg and Tohmo, 2018; Tohmo, 2025). The most frequently employed techniques include the location quotient (LQ) and its variants (Lamonica et al., 2018; Flegg et al., 2021; Kwon and Choi, 2024), the cross entropy (CE) method (Robinson et al., 2001; Zheng et al., 2022), and the RAS iterative proportional fitting method (Lamonica et al., 2020). Each method has distinct assumptions and strengths. An ongoing debate continues regarding their comparative accuracy and suitability across different contexts (Jahn et al., 2020).

This study aims to construct an RIOT for Selangor, adopting three non-survey methods using recent economic census data. This study contributes by providing a practical RIOT for Selangor and an empirical comparison of method performance. Previous work was undertaken by Saari and Zakaria (2009) to develop a Selangor RIOT, but it is limited to a single technique application and used the 2000 national IOT as a benchmark year. This approach aims to extend the methodological scope and update the reference year. By doing so, the study aims to inform both academic research and practical policy applications in regional modelling.

REVIEW OF LITERATURE

The body of literature on regional input-output table (RIOT) construction and estimation is extensive. These studies cover a wide range of methodological techniques, comparative evaluations, and multi-disciplinary applications across sub-national regions. Analysis of previous studies reveals two key gaps, making these studies highly context dependent. From a methodological perspective, there is limited consensus on best practices for regionalization or non-survey methods. This lack of agreement underscores the need for continuous evaluation and comparison of these approaches. In developing countries, such as Malaysia, the application of RIOT has been increasing, yet comparatively less attention has been given to evaluating alternative regionalization techniques. Hence, these gaps motivate our work in this paper.

The lack of survey-based RIOT has led to the development of various regionalization or non-survey estimation methods. Moreover, a practical estimation tool is needed for regional studies and decision-making. The location quotient (LQ) approach and its variants, such as simple location quotient (SLQ), cross-industry location quotient (CILQ), and Flegg location quotient (FLQ) have been among the most widely used techniques (Flegg and Tohmo, 2018; Jahn et al., 2020; Anaman et al., 2024; Mardones and Correa, 2025). Although these methods are relatively simple to implement, the strong assumption imposing the similarity between national and sub-national economic structures can lead to inaccurate estimates, especially for regions with high specialization. To overcome this assumption, several studies suggest industry and regional size adjustments and a hybrid approach combination to the LQ methods (Kowalewksi, 2015; Flegg and Webber, 2000; and Flegg et al., 2021) Nevertheless, such adjustments do not constitute a universal remedy, especially for regions characterized by unique industrial profiles.

Apart from the LQ methods, the recent application of information theory-based methods has gained attention as an alternative approach for regionalization techniques. The cross-entropy (CE) method, adapted from Robinson et al. (2001), provides flexibility in the RIOT estimation by incorporating prior information and constraints, which eventually improves the plausibility of estimated tables when limited regional data are available (Golan et al., 1994; Lamonica et al., 2020; Flegg et al., 2021). To the best of our knowledge, the application of this method is rather scarce in developing country cases, except for Zheng et al. (2022), who adopt the CE method to estimate city-level IOT for China. Nevertheless, for Malaysia, Saari et al. (2014) apply the CE method for updating the national IOT, a practice common in input-output research.

The RAS method, which relates to constrained matrix-balancing procedures, is also applied to regionalize the national IOT for estimating sub-national tables. Initially, Stone (1961) introduced RAS as a computational technique for the task of balancing a matrix to match specified row and column totals. It is also commonly used for IO updating procedures and available in multiple variants (Junius and Oosterhaven, 2003; Valderas-Jaramillo and Rueda-Cantuche, 2021; Lenzen et al., 2014y). Recent works that adopt RAS or its variants include Lamonica et al. (2020) for multiple countries' cases and Fournier Gabela (2020) applied to the Japanese IOT. In the Malaysian context, the RAS method is applied by Saari and Zakaria (2009) to estimate Selangor RIOT using the 2000 benchmark year. However, the study was limited to a single approach, while the current study extended the methodological scope and updated the base year.

In the context of developing countries, studies on sub-national IO construction have been growing. Extensive research is done for China, covering regional to city level IO development, demonstrating sprouting interest in the regional analysis using IO models (Flegg et al., 2015; Zheng et al., 2021; Zheng et al., 2022). In the Malaysian case, several studies are centred on regional IO development, such as Utit et al. (2020) for Sarawak, Saari et al. (2018) for Terengganu and Pahang, North Corridor Economic Region (NCER) (Hassan et al., 2017), and West Malaysia (Hassan et al., 2019). While these studies focus on the economic impacts of these regions, they also adopt SLQ or RAS methods as their regionalization techniques.

For a fast-growing state like Selangor, Saari and Zakaria (2009), and Mazan and Rashid (2012) developed Selangor RIOT using 2000 national IO as the benchmark year. Both studies use a single regionalization method, namely the RAS method, and do not systematically evaluate multiple non-survey techniques. Although a strong foundation of theoretical and methodological literature exists for RIOT estimation, gaps remain in comparative assessments and method validation, particularly in the Malaysian context. This study seeks to address these gaps by constructing the Selangor RIOT using three non-survey methods (SLQ, CE, RAS), applying updated economic census data, and systematically comparing their performance using distance-based measures. This approach aims to address practical needs in current policymaking while also contributing to the broader academic debate on regionalization techniques.

RESEARCH METHODOLOGY

This section provides the methods and materials used to meet the study objectives. First, the schematic structure of the regional input-output table (RIOT) is explained, illustrating the transaction flows of economic sectors within a single region. Next, the standard input-output (IO) model is derived, providing the basis for the analytical approach. Then, a brief explanation of the non-survey methods adopted in the study follows. Finally, the IO linkage effects and data sources are presented.

Input-Output Table Basic Structure

The Development of the Selangor RIOT relies on a single region IO setting, which focuses on a specific local area within a country. Such a framework is sufficient to assess inter-industry flows within the state economy, which offers distinct advantages in economic analysis. The basic structure of the Selangor RIOT is similar to the national input-output table (IOT), as shown in Table 1, which is presented in matrix form. According to the standard IO matrix framework, the $(n \times n)$ matrix $\mathbf{Z}$ represents intermediate input flows, where each element z_{ij} indicates the quantity of a commodity from sector i utilized by sector j in the production of final goods. To generate output for end consumers, production requires primary inputs as well. The $(1 \times n)$ vector m represents the imports across sectors, while the $(1 \times n)$ vector v denotes the value added in each sector.

According to the framework presented in Table 1, the independencies among production activities can be represented using the following material balance equation:

$$x = Zi + (c + i + g + e) = Zi + f \quad (1)$$

x refers to the total output vector, Z refers to the intermediate input matrix, i represents the summation vector, and f stands for the total final demand vector, which can be divided into four main components such as private consumption (c), investment (i), government consumption (g), and export (e). In sum, Eq. (1) shows that the total output of an industry equals the summation of intermediate demand flows and final demand components.

Table 1 Simplified structure of RIOT for Selangor

	Intermediate demand						Final demand				Total Output
	S_1	S_2	S_3	.	.	S_n	c	i	g	e	
Sector 1 (S_1) Sector 2 (S_2) Sector 3 (S_3) . . . Sector (S_n)	Z						f				x
Imports	m										
Value added	v										
Total input	x′										

Source: Illustrated by the authors

Standard Input-Output Model

The standard material balance equation written in Eq. (1) forms the basis for deriving a standard IO model. The model takes into account the direct and indirect effects of intermediate demand on the economic sectors. Hence, an $(n \times n)$ input coefficient matrix, A, is defined that is derived from dividing each intermediate demand flow by its total output, $A = Z\hat{x}^{-1}$. By treating the intermediate demand as an endogenous variable, while final demand as an exogenous variable, Eq. (1) can be rewritten as follows:

$$x = Ax + (c + i + g + e) = Ax + f \quad (2)$$

Eq. (2) can be further solved to arrive at the demand-driven IO model as follows:

$$x = (I - A)^{\wedge}(-1)\, f = Lf \quad (3)$$

where I is the identity matrix with one in the diagonal cells and zeros elsewhere. The matrix representation of $(I - A)^{-1}$is called the Leontief inverse matrix (L), which indicates the direct and indirect input requirements for each sector in order to satisfy its final demand.

Non-Survey and Regionalization Techniques

Non-survey or hybrid methods rely on the national IOT as their primary data source, given the absence of detailed regional-level data (Flegg and Tohmo, 2013; Tohmo, 2025). The primary assumption in these approaches is that similar technological structures across regions are assumed to align with national-level production patterns. From this assumption, various methodologies are employed to estimate regional coefficients, primarily by adjusting regional production structures from the national production structures. This assumption is commonly used in regional economic modelling due to its practicality and simplicity (Miller and Blair, 2009). While this approach may be effective in data-limited contexts (Lamonica and Chelli, 2018), it may fail to capture the true heterogeneity of regional economies, thus leading to potential biases in regional IO estimations (Lahr, 2001). However, the latter argument is beyond the scope of the study. As per the discussion in the literature review section, this study employs three methods: namely simple location quotient, RAS, and cross entropy. Given the abundance of regionalization techniques available in the literature, these methods are selected as a result of their relative simplicity and robustness.

Simple Location Quotient

The location quotient (LQ) technique is commonly used in the regional economic literature in estimating RIOT (Klijs et al., 2016; Ponomarev and Evdokimov, 2021). There are several alternatives in the LQ technique, namely Simple Location Quotient (SLQ), Cross Industry Location Quotient (CILQ), Round's Location Quotient (RLQ), and Flegg's Location Quotient (FLQ). Each of these methods has its own strengths and weaknesses that attract a body of literature to measure the closeness of each estimate to the 'true values' of RIOT. Then, these comparative analyses give insight into the 'bias' of the results (Ponomarev and Evdokimov, 2021; Flegg et al., 2021; Kwon and Choi, 2024). Following previous studies, SLQ was chosen as the main regionalization technique due to its simplicity (Zheng et al., 2022) and evidence of its estimation superiority (Lamonica and Chelli, 2018).

Formally, the SLQ is given as follows:

$$SLQ_i^r = \left[\frac{x_i^r/_{x^r}}{x_i/_{x}}\right] \quad (4)$$

where x_i^r is the region (Selangor) total output for sector i, x_i is the national (Malaysia) total output for sector i, x^r is the total output for the region, and x is the total output at the national level.

The numerator in Eq. (4) indicates the proportion of region r's total output that is contributed by sector i. The denominator represents the proportion of total national output that is contributed by sector i, nationally. The interpretation of $SLQ_i^r > 1$ indicates a commodity whose production is relatively localised in region r, viewed as a measure of the ability of regional industry i to supply the demands placed upon it by other industries (and by final demand) in that region. If sector i is less concentrated in the region than in the nation ($SLQ_i^r < 1$), it is seen as less capable of satisfying regional demand for its output.

RAS Method

The second method employed is the RAS approach, which is used to balance the intermediate input after controlling the primary inputs and final demand estimates (Sargento et al., 2024). In the context of IO regionalization, the method serves as an iterative adjustment procedure that ensures the estimated RIOT conforms to known row and column margins—typically derived from partial regional data or control totals (Lamonica et al., 2020). According to Trinh and Phong (2013), the RAS adjustment functions as an iterative process. In this process, the sums of columns and rows (or rows and columns) are continuously adjusted until they match the specified margin totals, ensuring that the discrepancies in the balancing values of rows and columns are eliminated.

To conduct the procedure, several types of sectoral information by sector *n* are required, namely: (1) an initial estimate of inter-industry flow derived from the national IO table (denoted as $A^0 = a_{ij} = z_{ij}x_i^{-1}$); (2) total intermediate input (denote as $u_i^0 = \sum_{j=1}^{n} z_{ij}$) and target row sum u_i^1; and (3) total intermediate output (denoted as $v_j = \sum_{i=1}^{n} z_{ij}$) and target column sum v_j^1, each of which for target sum for (2) and (3) is regional information that is sourced from the economic census data. Following Lamonica et al. (2020), the RAS method can be mathematically derived by denoting r and s as the vector with entries, $r_i = \frac{u_i^0}{u_i^1}$; $s_j = \frac{v_j^0}{v_j^1}$; $i, j = 1,2, \dots, n$; and D(s) is the diagonal matrix whose entries are $d_{ii}(r) = r_i$; $d_{jj}(s) = s_j$; $i, j = 1,2, \dots, n$. The RAS procedure can be summarized by making two iterative rescaling of rows and columns.

For k=0 set:

$$\widehat{A}^k = A^0 D(s) \quad (5)$$

For k = 1, 2, 3, ..., repeat:

$$r_i^{(k)} = \frac{x_i}{\left(u'\widehat{A}^{(k-1)}\right)_i}, i = 1, 2, \dots, n \quad (6)$$

$$\widehat{A}^{(k)} = D(r^{(k)})\widehat{A}^{(k-1)} \quad (7)$$

$$\mathrm{s}_j{}^{(\mathrm{k})} = \frac{x_i}{\left(\widehat{\mathrm{A}}^{\mathrm{R}(\mathrm{k})}\mathrm{v}\right)_j}, i = 1, 2, \ldots, n \tag{8}$$

$$\widehat{\mathrm{A}}^{(\mathrm{k}+1)} = \widehat{\mathrm{A}}^{(\mathrm{k})}\mathrm{D}(\mathrm{s}^{(\mathrm{k})}) \tag{9}$$

until the following stopping criterion is satisfied:

$$|\mathrm{u}'\mathrm{A}^{(\mathrm{k}+1)} - \mathrm{x}| < \varepsilon \text{ and } |\mathrm{A}^{(\mathrm{k}+1)}\mathrm{v} - \mathrm{x}| < \varepsilon \tag{10}$$

for a positive tolerance parameter ε. In this study, $\varepsilon = 0.08$.

Cross-Entropy Method

The third method employed in this study is the Cross-Entropy (CE) method. It applies the concept of minimizing the Kullback and Leibler (1951) measure of divergence (relative entropy) between two sets of probabilities. In the context of this study, it represents the distance between an initial IO matrix and a set of known constraints, rendering it suitable for contexts with limited regional data. In mathematical terms, given a national IOT intermediate delivery, Z^{N} and incomplete observations of regional transactions Z^{R}, the CE method seeks to update an initial estimate Z^{N} by minimizing:

$$\min, \mathrm{D}_{KL}(\mathrm{Z}^{\mathrm{R}}||\mathrm{Z}^{\mathrm{N}}) = \sum_{i,j} z_{ij}^{\mathrm{R}} \ln\left(\frac{z_{ij}^{\mathrm{R}}}{z_{ij}^{\mathrm{N}}}\right) \tag{11}$$

Subject to the regional supply and demand constraints:

$$\sum_j z_{ij} = r_i; \sum_i z_{ij} = r_j \tag{12}$$

where R_i and R_j represent observed regional output and input totals, respectively. This approach ensures that the resulting RIOT adheres to known marginal totals while incorporating prior knowledge in a systematic, probabilistically sound way. In addition, the method could also flexibly incorporate various types of prior information, including partial row and column totals, macroeconomic aggregates, inequality constraints, and explicit modelling of measurement errors (Robinson et al., 2001). In comparison with the RAS approach, which balances only known margins, the CE offers a more general framework capable of handling inconsistent or incomplete data—a common condition in regionalization.

Input-Output Linkage Effects

This study examined the input-output analytical outcomes produced by the Selangor RIOT. Input-output backward and forward linkage effects are applied to illustrate the interdependencies between sectors. Miller and Blair (2009) define backward (forward) linkage as an indicator of a sector's interconnection with upstream (downstream) sectors, from which it purchases or to which it sells output. Collectively, these effects offer insights into the role of industries' interconnectedness in generating growth across the economy. Backward and forward linkage analysis enables policymakers to identify sectors benefiting most from a given industry's growth and highlights the channels through which economic benefits are transmitted. The backward and forward linkages are derived from Leontief and Ghosh inverse matrices, respectively. The former has been explained in Eq. (3), while the latter is another approach that is based on a supply-driven model (Lenzen, 2003).

In short, the Ghosh model is denoted as:

$$\mathrm{x}' = \mathrm{i}\hat{\mathrm{x}}\mathrm{B} + \mathrm{d}' = \mathrm{x}'\mathrm{B} + \mathrm{d}' \tag{13}$$

where $\mathrm{i}'\hat{\mathrm{x}} = \mathrm{x}'$, $\mathrm{B}(\mathrm{B} = \mathrm{v}\hat{\mathrm{x}}^{-1})$ is referred to as the output coefficient matrix and d' is the vector of primary inputs (e.g. imports and value-added). Each element of the output coefficient matrix indicates the proportion of sectoral output that is delivered to other sectors. It presents the delivery of sector i (as a seller) to sector j (as the buyer) per unit of the sector i's output. In other words, this transaction represents the forward production relationships among sectors in the economy (Oosterhaven, 1988). This model primarily analyzes forward linkages, identifying sectors that act as key suppliers within the production network. According to Miller and Blair (2009), the formulas for calculating backward and forward linkage indicators are as follows:

$$\bar{b} = \frac{n i' A}{i' A i} = \frac{n i' L}{i' L i} \tag{14}$$

$$\bar{f} = \frac{n B i}{i' B i} = \frac{n G i}{i' G i} \tag{15}$$

Each formula represents a normalized total backwards (forward) linkage index, with an average value of unity as suggested by Ramussen (1956). Here, i and n indicate the summation vector and the total number of sectors, respectively. In this instance, sectors with "above average" backwards (forward) linkages have indices that are greater than one (stronger linkages) and those with "below average" backwards (forward) linkages have indices that are less than one (weaker linkages).

Data Sources and Processing

In this study, the Malaysian input-output table (IOT) for the year 2015, published by the Department of Statistics Malaysia (DOSM), is utilized as the benchmark table prior to estimating the RIOT Selangor. The 2015 benchmark year is selected because the primary data for the IOT comes from the 2015 Economic Census. DOSM had not published a new IOT or census dataset for other benchmark years. Additionally, regionalization of the national IoT IOTrequires information from state accounts statistics, household income and expenditure surveys, and trade statistics, all obtained at the state level, at a minimum at the most aggregated level. Some of these data are used as constraints to maintain consistency with the main aggregates. Table 2 outlines the data requirements for this study.

The collected economic census data are formatted according to the Malaysia Standard Industrial Classification (MSIC) 2008 with a 5-digit industrial code. The economic census data are aggregated into the IOT classification based on MSIC-to-IOT sectoral concordance mapping provided by DOSM. Through this process, the detailed sectoral-level information is systematically harmonized with the broader IOT classification. The resulting baseline structure for the Selangor RIOT comprises 119 aggregated sectors, representing a harmonized synthesis of the economic census data and the IOT classification framework.

Table 2 Data requirements for developing Selangor RIOT

Data	Classification	Source of Data
Malaysia Input-Output Tables	MSIC	National Accounts Statistics
State Accounts Statistics	MSIC	MyLocal Stats, Selangor
Value Added	MSIC	Economic Census 2015
Total Output	MSIC	Economic Census 2015
Intermediate Input	MSIC	Economic Census 2015
Household Final Demand	COICOP	Household Income Expenditure Survey (HIES) 2016
Gross Fixed Capital Formation	MSIC	Economic Census 2015
Export	SITC	Trade Statistics
Import	SITC	Trade Statistics

Note: MSIC refers to Malaysian Standard Industry Classification 2008, COICOP refers to Classification of Individual Consumption by Purpose and SITC refers to Standard International Trade Classification.

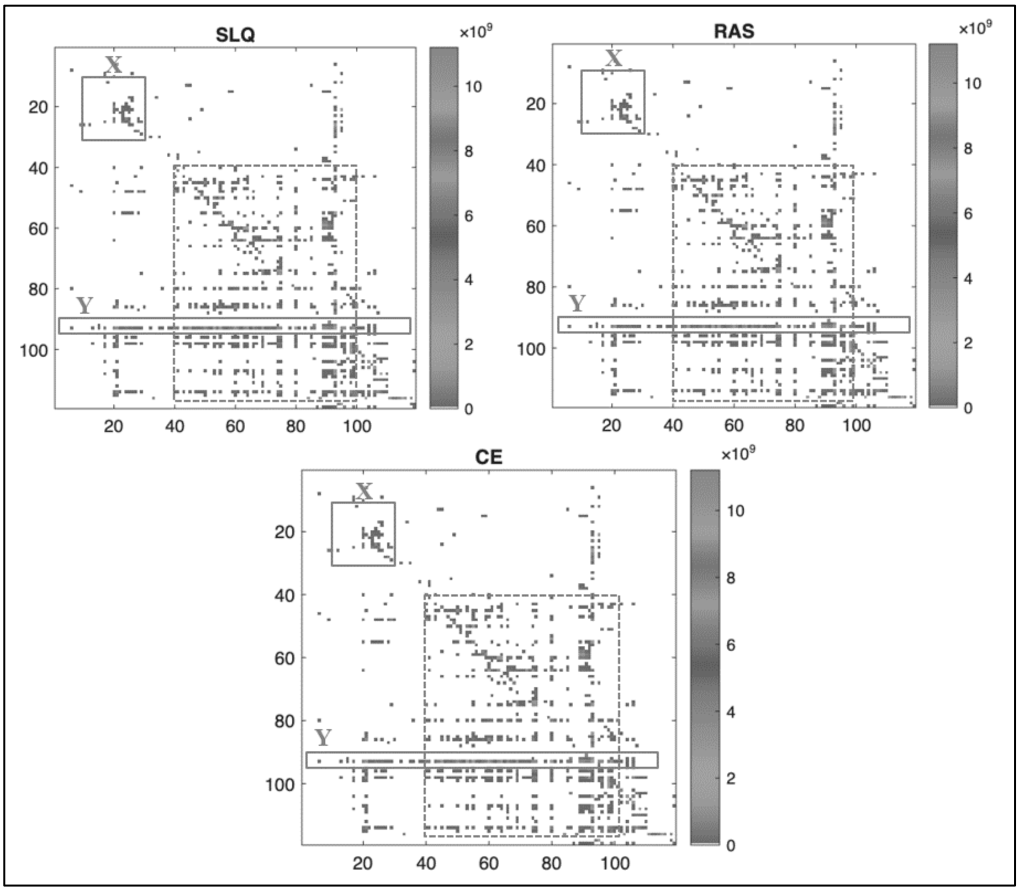
RESULTS AND DISCUSSION

This section presents the results of the construction of the Selangor RIOT through three non-survey methods. The comparative assessment of the estimates is first illustrated through visualization and matrix distance measures. Subsequently, an analysis employing IO demand-driven linkage measures is conducted to examine differences across the estimation approaches.

Comparative Assessment of Structural Distance

The use of heatmap visualization is common in IO research to compare the structural distance between estimated tables (Geschke et al., 2014; Abd Rahman et al., 2017). Figure 1 shows three heatmaps, each depicting the intermediate demand matrix (Z) for Selangor RIOT. Overall, the matrices appear relatively similar and consistent across the SLQ, RAS, and CE methods. For example, 'Area X' represents inter-industry flow within 20 sectors (Sector 10 - Sector 30), illustrating a similar structural distribution across the estimated tables. Likewise, 'Area Y', which corresponds to Sector

93 (Wholesale and Retails), exhibits most 'filled' transactions on the demand-side (row-wise), suggesting its central role in the Selangor economy. Meanwhile, the Z matrices for all estimated tables consistently highlight the concentration of inter-industry transactions from Sector 40 to Sector 100, encompassing both manufacturing and service sectors (as shown within the dotted square). This finding reflects the integrated nature of intermediate demand (distribution) and supply (production) within Selangor's mid-to-high complexity sectors.



Note: The heatmap images are generated in MATLAB software. The x-axis and y-axis represent sectoral coverage in the Selangor RIOT for each estimation method.

Figure 1 Heatmaps of intermediate demand flows for Selangor RIOT

To validate the initial qualitative assessment of the closeness of the estimated intermediate demand matrices, differences are tested using a set of statistical tools known as 'matrix distance measures' (Geschke et al., 2014). These measures are also frequently applied in the IO literature (Lamonica et al., 2020; Flegg et al., 2021) to evaluate distances between estimated IO tables derived from various regionalization approaches. Table 3 summarises the pair-wise comparison of the four distance measures: mean absolute difference (MAD), Euclidean metric distance (EMD), regression coefficient (1-R2), and Pearson's correlation coefficient (CORR). Notably, the estimated tables are found to be largely 'close', with all measures indicating minor differences in the intermediate demand matrices. The pairwise comparison indicates that RAS and CE show the smallest difference, as evidenced across all distance measures. This result demonstrates that RAS and CE yield the most comparable structures for the Selangor RIOT.

Table 3 Pair-wise comparison of distance measure for the intermediate demand matrix

Method	**Mean Absolute Difference (MAD)**	**Euclidean Metric Distance (EMD)**	**Regression Coefficient ($1-R^2$)**	**Pearson's Correlation Coefficient (CORR)**
SLQ-RAS	5.02×10^{-6}	1.01×10^{-4}	2.79×10^{-5}	1.40×10^{-5}
SLQ-CE	5.21×10^{-6}	1.02×10^{-4}	2.82×10^{-5}	1.41×10^{-5}
RAS-CE	2.50×10^{-7}	4.28×10^{-6}	4.65×10^{-8}	2.33×10^{-8}

Source: Author's own calculation

Subsequently, standard IO backward and forward linkage analysis is applied to the three estimated tables. This analysis is necessary because, in addition to testing matrix similarity, understand the extent to which analytical outcomes from the estimated matrices differ from some IO outcome comparative studies (Abd Rahman et al., 2021; Moran and Wood, 2014; Owen et al., 2014). The observed alterations between the IO tables result from the use of different data sources and variations in construction methodologies (Abd Rahman et al., 2021). In this case, policy implications could be substantial, where precautionary interpretation is necessary to avoid drawing misleading conclusions.

Table 4 presents the comparative results for the difference in linkage effects of the top and bottom 15 sectors. SLQ estimates serve as the point of reference, and sector-wise differences are calculated using relative changes (for example, $\frac{SLQ_i - RAS_i}{SLQ_i}$). Overall, estimated backward and forward linkages of the top and bottom 15 sectors across all three estimation methods show high consistency, with differences generally small in magnitude. Backward and forward linkage indicators derived from the SLQ, RAS, and CE provide sufficient insights regarding sectoral strength and importance. Minor discrepancies occur in some sectors, such as Glass and Glass products, which display slight divergence between forward linkages derived from RAS and CE. Similarly, the backward linkage for RAS and CE differ slightly for Other General-Purpose Machinery and Vegetables. Meanwhile, among the bottom 15 sectors, the Other Private Services sector shows the highest difference for both RAS and CE, compared to SLQ.

Table 4 Difference of linkage effects of top and bottom 15 sectors (SLQ as point of reference)

Sector	SLQ		RAS		CE	
	Backward	Forward	Backward	Forward	Backward	Forward
Top 15 sectors with highest backward linkages						
Cement, Lime and Plaster	1.394	1.401	1.98×10^{-5}	3.27×10^{-5}	1.83×10^{-5}	3.14×10^{-5}
Leather Products	1.384	0.961	1.29×10^{-5}	2.90×10^{-5}	1.14×10^{-5}	3.12×10^{-5}
Structural Metal Products & Tanks*	1.320	0.783	9.83×10^{-6}	3.39×10^{-6}	8.17×10^{-6}	2.25×10^{-6}
Non-Metallic Mineral Products	1.316	1.304	1.41×10^{-5}	2.88×10^{-5}	1.32×10^{-5}	2.76×10^{-5}
Glass and Glass Products	1.307	1.054	5.59×10^{-6}	6.80×10^{-7}	3.50×10^{-6}	1.73×10^{-6}
Soaps & Detergents*	1.307	0.695	6.59×10^{-6}	7.75×10^{-6}	1.01×10^{-5}	8.70×10^{-6}
Basic Iron and Steel	1.305	1.369	1.33×10^{-5}	1.01×10^{-5}	1.24×10^{-5}	8.84×10^{-6}
Dairy Products	1.301	0.624	3.33×10^{-6}	6.65×10^{-6}	5.27×10^{-6}	7.81×10^{-6}
Rubber Tyres and Tubes	1.297	1.023	2.80×10^{-5}	8.23×10^{-6}	3.05×10^{-5}	7.28×10^{-6}
Other General-Purpose Machinery	1.297	0.771	2.98×10^{-7}	2.27×10^{-6}	1.87×10^{-6}	3.63×10^{-6}
Vegetable & Animal Oils and Fats	1.288	0.881	6.79×10^{-5}	3.75×10^{-5}	7.63×10^{-5}	4.22×10^{-5}
Basic Metals*	1.283	0.807	3.05×10^{-6}	3.95×10^{-6}	4.72×10^{-6}	5.07×10^{-6}
Computers, & Office Equipment*	1.279	0.534	5.93×10^{-6}	7.10×10^{-6}	9.04×10^{-6}	8.34×10^{-6}
Weapons & Ammunition*	1.275	0.636	6.91×10^{-6}	7.34×10^{-6}	5.30×10^{-6}	8.56×10^{-6}
Quarrying of Stone, Sand and Clay	1.259	1.371	3.19×10^{-6}	3.40×10^{-5}	3.35×10^{-6}	3.27×10^{-5}
Bottom 15 sectors with lowest backward linkages						
Real Estate	0.716	0.736	2.07×10^{-6}	2.20×10^{-5}	1.16×10^{-6}	2.35×10^{-5}
Veneer Sheets and Wood-based*	0.705	1.004	2.98×10^{-6}	8.09×10^{-6}	3.50×10^{-6}	6.90×10^{-6}
Professional	0.700	1.255	5.42×10^{-6}	6.15×10^{-5}	9.31×10^{-6}	6.26×10^{-5}
Monetary Intermediation	0.690	1.062	4.23×10^{-6}	3.28×10^{-5}	5.10×10^{-6}	3.38×10^{-5}
Rental and Leasing	0.688	1.171	6.10×10^{-6}	4.53×10^{-6}	1.03×10^{-5}	4.42×10^{-6}
Scientific R&D*	0.677	0.530	3.13×10^{-6}	7.50×10^{-6}	6.10×10^{-6}	8.74×10^{-6}
Non-Profit Institutions*	0.669	0.530	1.39×10^{-5}	7.50×10^{-6}	1.24×10^{-5}	8.74×10^{-6}
Fruits	0.660	0.784	5.21×10^{-6}	3.17×10^{-5}	4.84×10^{-6}	3.58×10^{-5}
Vegetables	0.654	0.706	9.51×10^{-7}	6.01×10^{-5}	2.35×10^{-6}	6.73×10^{-5}
Public Administration	0.647	1.378	3.02×10^{-6}	2.77×10^{-5}	2.50×10^{-6}	3.35×10^{-5}
Fishing and Aquaculture	0.632	0.924	2.58×10^{-6}	7.78×10^{-5}	3.01×10^{-6}	8.71×10^{-5}
Other Private Services	0.628	0.866	4.43×10^{-4}	4.49×10^{-4}	4.45×10^{-4}	4.49×10^{-4}
Flower Plants	0.603	0.707	5.29×10^{-6}	1.00×10^{-5}	5.40×10^{-6}	1.14×10^{-5}
Ships, Boats, & Bicycles*	0.599	0.700	7.15×10^{-6}	7.83×10^{-6}	6.77×10^{-6}	8.88×10^{-6}
Arts & Entertainment*	0.588	0.660	1.58×10^{-6}	9.22×10^{-5}	1.28×10^{-6}	8.74×10^{-5}

Note: * sector names have been truncated to save space. Full result tables for the 119 sectors obtained from each regionalization technique are available upon request from the author.

Source: Author's own calculation.

Analysis suggests that the three regionalization techniques are comparable for constructing RIOT, with CE and RAS showing particularly close similarity. This result is expected, as CE and RAS employ a similar initial estimate, which is the national IOT structure, and apply balancing procedures subject to pre-determined constraints (regional data). On the other hand, SLQ requires 'trimming' the national structure by means of scaling the national IO input coefficients using sector- and region-specific ratios. This procedure will eliminate industries that are presumed to be less competitive for the region. Nonetheless, this can be argued that for the case of Selangor, which contributes almost 30% to national GDP and has achieved economic maturity across many sectors, except for resource-based sectors, the similarity of these approaches is somehow justified.

Finally, the consistent estimates of backward and forward linkage indicators across SLQ, RAS, and CE methods allow a robust consensus on the role of key sectors within the Selangor economy. The results obtained in this study provide confidence in selecting analytical outcomes derived from these approaches. These outcomes enable policymakers to design policy interventions based on methodologically sound indicators, regardless of the estimation techniques adopted. Regarding linkage indicators, the observed consistency allows policymakers to prioritize sectors for investment, tax incentives, or infrastructure support, since the sectors are central to both upstream (backward linkage) and downstream (forward linkage) economic activity (Miller and Blair, 2009).

CONCLUSION

The use of an input-output table (IOT) for regional economic evaluation is emerging in developed and developing countries. Literature frequently notes the absence of an official regional IOT (RIOT), a gap that requires regional analysts to apply various regionalization techniques to estimate the RIOT. This study estimates the Selangor RIOT using three IO regionalization approaches: SLQ, RAS, and CE. Estimates obtained from the three methods demonstrate high similarity, as indicated by minimal differences across multiple matrix distance measures, including MAD, EMD, regression coefficient, and Pearson’s correlation coefficient. Backward and forward linkage effects derived from the Selangor RIOT show a high degree of consistency, demonstrating each method in capturing the Selangor economy.

Based on the consistent estimates of backwards and forward linkage indicators obtained by these three methods, policymakers in Selangor are well-positioned to formulate targeted regional development strategies with a high degree of methodological confidence. The robustness of these results enables the identification of sectors that play a pivotal role in both supplying inputs to other industries and generating downstream economic activity. For example, the results suggest that the manufacturing sector – such as Cement, Lime and Plaster; Non-Metallic Mineral Products; Glass and Glass Products, and Basic Iron and Steel – indicates a relatively high backwards linkage effect to the state economy. This outcome reinforces the strong role of these sectors in stimulating growth among supplying industries. Accordingly, policy interventions could prioritize these key sectors for strategic investment, skills development programs, and infrastructure upgrades to amplify sectoral spillover effects. These approaches would not only strengthen the structural interdependencies within the Selangor economy but also drive value creation throughout the supply chain.

Despite the reliability of the estimates, limitations remain that future studies could address. First, the absence of an actual survey-based Selangor RIOT restricts the comparative analysis from fully assessing the extent to which the estimates reflect the ‘real’ structure of the Selangor economy. Second, primary inputs and final demand are controlled, while intermediate inputs are allowed to vary. Nonetheless, regional studies recommends applying regionalization methods to additional input-output components. Regionalization methods can also support the development of inter-regional or multi-regional IOT, enabling sectoral flow information to be expanded into smaller spatial units, such as districts or cities.

ACKNOWLEDGEMENT

This study was funded by Selangor State Government under the Geran Penyelidikan Negeri Selangor 2023 [Grant Number: SUK/GPNS/2023/PEM/04]. The authors would like to thank anonymous referees for their valuable comments for the earlier version of this paper. Special thanks also extend to Mr. Amirul Syafiq Azlan for his assistant.

REFERENCES

Abd Rahman, M. D., Los, B., Geschke, A., Xiao, Y., Kanemoto, K., & Lenzen, M. (2017). A flexible adaptation of the WIOD database in a virtual laboratory. *Economic Systems Research*, *29*(2), 187–208.

Abd Rahman, M. D., Los, B., Owen, A., & Lenzen, M. (2021). Multi-level comparisons of input–output tables using cross-entropy indicators. *Economic Systems Research*, *35*(1), 75–94.

Anaman, K. A., & Shaibu, A. F. (2024). Development of regional input-output tables for Northern Ghana: An analysis using location quotient methods. *Cogent Social Sciences*, *10*(1), 2340429.

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

Flegg, A. T., & Tohmo, T. (2013). Regional input–output tables and the FLQ formula: A case study of Finland. *Regional Studies*, *47*(5), 703-721.

Flegg, A. T., & Tohmo, T. (2018). *The regionalization of national input–output tables: A review of the performance of two key non-survey methods*. In K. Mukhopadhyay (Ed.), Applications of the input-output framework (pp. 347-386). Springer.

Flegg, A. T., Huang, Y., & Tohmo, T. (2015). Using CHARM to adjust for cross-hauling: The case of the province of Hubei, China. *Economic Systems Research*, *27*(3), 391-413.

Flegg, A. T., Lamonica, G. R., Chelli, F. M., Recchioni, M. C., & Tohmo, T. (2021). A new approach to modelling the input–output structure of regional economies using non-survey methods. *Journal of Economic Structures*, *10*, 1-31.

Fournier Gabela, J. G. (2020). On the accuracy of gravity-RAS approaches used for inter-regional trade estimation: evidence using the 2005 inter-regional input–output table of Japan. *Economic Systems Research*, *32*(4), 521-539.

Golan, A., Judge, G., & Robinson, S. (1994). Recovering information from incomplete or partial multisectoral economic data. *The Review of Economics and Statistics*, *76*(3), 541–549.

Hassan, M. K. H., Azmi, N. A., Arip, M. A., & Liew, V. K-S. (2017). Regional input-output table: The case of North Corridor Economic Region (NCER) in Malaysia. *International Journal of Business and Social Science*, *8*(12), 65-72.

Hassan, M. K. H., Noor, Z. M., Ismail, N. W., Radam, A., & Rashid, Z. A. (2019). The contribution of various sectors in West Malaysia to the economic growth: An input-output analysis. *International Journal of Academic Research in Business and Social Sciences*, *9*(1), 221–234.

Huang, W., Corbett, J. J., & Jin, D. (2015). Regional economic and environmental analysis as a decision support for marine spatial planning in Xiamen. *Marine Policy*, *51*, 555-562.

Islam, K. N., Kenway, S. J., Renouf, M. A., Wiedmann, T., & Lam, K. L. (2021). A multi-regional input-output analysis of direct and virtual urban water flows to reduce city water footprints in Australia. *Sustainable Cities and Society*, *75*, 103236.

Jahn, M., Flegg, A. T., & Tohmo, T. (2020). Testing and implementing a new approach to estimating interregional output multipliers using input–output data for South Korean regions. *Spatial Economic Analysis*, *15*(2), 165-185.

Jensen, C. D., McIntyre, S., Munday, M., & Turner, K. (2011). Responsibility for regional waste generation: A single-region extended input–output analysis for Wales. *Regional Studies*, *47*(6), 913–933.

Juhász, R., Lane, N., & Rodrik, D. (2023). The new economics of industrial policy. *Annual Review of Economics*, *16*, 213-242

Junius, T., & Oosterhaven, J. (2003). The solution of updating or regionalizing a matrix with both positive and negative entries. *Economic Systems Research*, *15*(1), 87-96.

Klijs, J., Peerlings, J., Steijaert, T., & Heijman, W. (2016). *Regionalising Input-Output Tables: Comparison of Four Location Quotient Methods*. In Matias, Á., Nijkamp, P., Romão, J. (eds) Impact Assessment in Tourism Economics. (pp. 43–65), Springer.

Kowalewksi, J. (2015). Regionalization of national input–output tables: Empirical evidence on the use of the FLQ formula. *Regional Studies*, *49*(2), 240–250.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86.

Kwon, H., & Choi, S. G. (2024). An alternative approach to estimating regional input–output tables: the KFLQ method. *The Annals of Regional Science*, *72*(2), 561-578.

Lahr, M. L. (2001). Reconciling domestication techniques, the notion of re-exports, and some comments on regional accounting. *Economic Systems Research*, *13*(2), 165-179.

Lamonica, G. R., & Chelli, F. M. (2018). The performance of non-survey techniques for constructing sub-territorial input-output tables. *Papers in Regional Science*, *97*(4), 1169-1203.

Lamonica, G. R., Recchioni, M. C., Chelli, F. M., & Salvati, L. (2020). The efficiency of the cross-entropy method when estimating the technical coefficients of input–output tables. *Spatial Economic Analysis*, *15*(1), 62-91.

Lenzen, M., Moran, D. D., Geschke, A., & Kanemoto, K. (2014). A non-sign-preserving RAS variant. *Economic Systems Research*, *26*(2), 197-208.

Liu, E. (2019). Industrial policies in production networks. *The Quarterly Journal of Economics*, *134*(4), 1883-1948.

Mardones, C., & Correa, M. (2025). Methodological proposal to approximate the sectoral impacts of a carbon tax at the regional level–the case of Chile. *Economic Systems Research*, *37*(1), 52-75.

Mazan, D., & Rashid, Z. A. (2012). Constructing regional input-output table: A case study of Selangor, Malaysia. *Business and Management Quarterly Review*, 21-32.

Miller, R. E., & Blair, P. D. (2009). *Input-output analysis: foundations and extensions*. Cambridge University Press.

Moran, D., & Wood, R. (2014). Convergence between the EORA, WIOD, EXIOBASE, and OPENEU's consumption-based carbon accounts. *Economic Systems Research*, *26*(3), 245–261.

Oosterhaven, J. (1988). On the plausibility of the supply-driven input-output model. *Journal of Regional Science*, *28*(2), 203-217.

Owen, A., Steen-Olsen, K., Barret, J., Wiedmann, T., & Lenzen, M. (2014). A structural decomposition approach to comparing MRIO databases. *Economic Systems Research*, *26*(3), 262–283.

Patandianan, M. V., & Shibusawa, H. (2020). Evaluating the spatial spillover effects of tourism demand in Shizuoka Prefecture, Japan: an inter-regional input–output model. *Asia-Pacific Journal of Regional Science*, *4*(1), 73-90.

Peters, G. P., & Wood, R. (2023). Dynamic input-output modeling for climate policy: A cross-entropy regionalization approach. *Journal of Cleaner Production*, *412*, 137292.

Ponomarev, Y. Y., & Evdokimov, D. Y. (2021). Construction of truncated input–output tables for russian regions using location quotients. *Studies on Russian Economic Development*, *32*(6), 619-630.

Rasmussen, P.N. (1956). *Studies in Intersectoral Relations*. North-Holland, Amsterdam, Netherlands.

Robinson, S., Cattaneo, A., & El-Said, M. (2001). Updating and estimating a social accounting matrix using cross-entropy methods. *Economic Systems Research*, *13*(1), 47-64.

Saari, M. Y., & Abdul Rashid, Z. (2009). Pembangunan jadual input-output wilayah dan analisis ke atas struktur industri Selangor. *International Journal of Management Studies*, *16*(1), 1-30.

Saari, M. Y., Habibullah, M. S., Utit, C., & Maji, I. K. (2018). Economic impacts of petroleum industry in states of Pahang and Terengganu. *Jurnal Ekonomi Malaysia*, *52*(2), 149-161.

Saari, M. Y., Hassan, A., Abd Rahman, M. D., & Mohamed, A. (2014). Evaluation of the relative performance of RAS and cross-entropy techniques for updating input-output tables of Malaysia. *Malaysian Journal of Economic Studies*, *51*(2), 217-229.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423 and 623–659.

Steen-Olsen, K., Owen, A., & Hertwich, E. G. (2022). Environmental footprint modeling using cross-entropy-based MRIO regionalization. *Journal of Industrial Ecology*, *26*(3), 704–720.

Tohmo, T. (2025). Estimating SFLQ-based regional input-output tables for South Korean regions. *National Accounting Review*, *7*(1), 125-142.

Trinh, B., & Phong, N. V. (2013). A short note on RAS method. *Advances in Management and Applied Economics*, *3*(4), 133-137

Utit, C., Saari, M. Y., Abd Rahman, M. D., Habibullah, M. S., & Norazman, U. Z. (2020). Regional economic impacts of natural resources: The case of petroleum, forestry, and logging in Sarawak. *International Journal of Business and Society*, *21*(2), 898-916.

Valderas-Jaramillo, J. M., & Rueda-Cantuche, J. M. (2021). The multidimensional nD-GRAS method: Applications for the projection of multiregional input–output frameworks and valuation matrices. *Papers in Regional Science*, *100*(6), 1599-1624.

Zheng, H., Bai, Y., Wei, W., Meng, J., Zhang, Z., Song, M., & Guan, D. (2021). Chinese provincial multi-regional input-output database for 2012, 2015, and 2017. *Scientific Data*, *8*, 244.

Zheng, H., Többen, J., Dietzenbacher, E., Moran, D., Meng, J., Wang, D., & Guan, D. (2022). Entropy-based Chinese city-level MRIO table framework. *Economic Systems Research*, *34*(4), 519-544.