Study on the Compilation of Time Series Input-Output Tables in China

Zhang Hongxia
School of Applied Economics
Renmin University of China

Contents

Research Background

- **Necessity of time series input-output tables**

Some countries' official input-output tables for consecutive years are unavailable for some reasons, while differences in sectoral categories and calibres in different compilation years also affect the comparability of official historical input-output data. This has led to several obstacles to the application of the input-output methodology.

Firstly, the input-output table data cannot complement other statistical data well enough to analyse economic characteristics as well as their changes in terms of both structural and aggregate data on a continuous time dimension.

Secondly, there are considerable differences in sectoral categories and calibres in official tables from one period to another, which fail to provide consistent comparative analysis and monitoring of structural changes in the economy over time.

Research Background

Thirdly, the absence of time series of input-output tables also brings obstacles to the combination of input-output model with other quantitative methods in the analysis of important socio-economic issues.

Fourthly, as shown by the compilation of input-output tables in other countries, most countries compile annual input-output serial tables to enable dynamic monitoring of long-term economic structures and industrial linkages.

Research Background

- **The main methodologies for the compilation of time series IO tables in the current literature**

The compilation of time series input-output tables usually involves the estimation of technical coefficient matrices and key aggregate indicators for the years with no official IO tables, and most of the studies that have been conducted are based on the known annual coefficient matrices and aggregate data. For example, a linear interpolation method was applied to estimate the coefficient matrix for years in between the two years with IO tables(Stadler et al., 2018); Wang et al. (2015) and Zheng et al. (2018) used the matrix transformation technique (MTT) to compile and forecast input-output tables for China for the period 1992-2020.

Research Background

SUT-RAS method is the methodology used by the World Input-Output Database (WIOD) to compile the time series supply-use tables or input-output tables for each country, where the initial estimation matrix and control indicators are obtained by interpolation or weighted averaging based on the SUT tables of the two known years (Dietzenbacher et al., 2013; Wu and Keiko , 2015).

Research Background

Lenzen et al. (2012) used forward and backward weighted averaging to estimate a smoothed serial input-output table over the entire time period, and this method was also used in the compilation of the Eora MRIO data (Lenzen et al., 2013); Based on a similar method, Wang et al. (2015) used Eora MRIO data as well as the Inter-regional Input-Output Tables, Provincial Input-Output Tables and the 2007 National Input-Output Tables in China to compile a nested inter-regional linkage table between China and 185 countries in the world for the period 1997 to 2011.

Research Background

- A common feature of the serial input-output tables of China in literature or provided by relevant institutions is that they primarily rely on mathematical methods to extrapolate from existing input-output tables.
- In practice, although no specific input-output surveys are conducted in non-survey years, national statistical data and data published by various government institutions also provide a wealth of important information that can be used for compilation purposes. The methodology published by the National Bureau of Statistics for compiling input-output extension tables for the years that ends with 0 and 5 is based on available statistical information and benchmark input-output tables. The full use of known information ensures that the closest input-output extension tables to economic reality can be obtained in years when no specific input-output survey is conducted.

Research Background

- This paper has developed a method for the compilation of time series input-output tables in China for the period 1981-2018, which is based on various types of statistical data and national economic accounting data. The method is based on the methodology used by the National Bureau of Statistics to compile input-output tables and input-output extension tables in China, and makes full use of the statistical information available, with the auxiliary of the necessary mathematical methods. This compilation method has two major advantages: firstly, it maintains the internal consistency between the compilation method and the method used by the National Bureau of Statistics to compile input-output extension tables, which enables the compiled tables to have better coherence and comparability with the data published by the National Bureau of Statistics; secondly, the tables are compiled on the basis of taking full advantage of known statistical information, rather than just making projection based on existing input-output tables, which enables the compiled tables to more fully reflect the actual economic situation in the year of compilation.

Compilation Framework and Key Methodology

- Basic compilation framework

The basic thought of the data-centred compilation of time series input-output tables proposed in this paper is to refer to compilation method of the input-output tables and input-output extension tables published by the National Bureau of Statistics and to make full use of the statistical and accounting data published by the National Bureau of Statistics and various government institutions, with the necessary mathematical methods as a supplement; and to maintain consistency with the official GDP accounting data.

Compilation Framework and Key Methodology

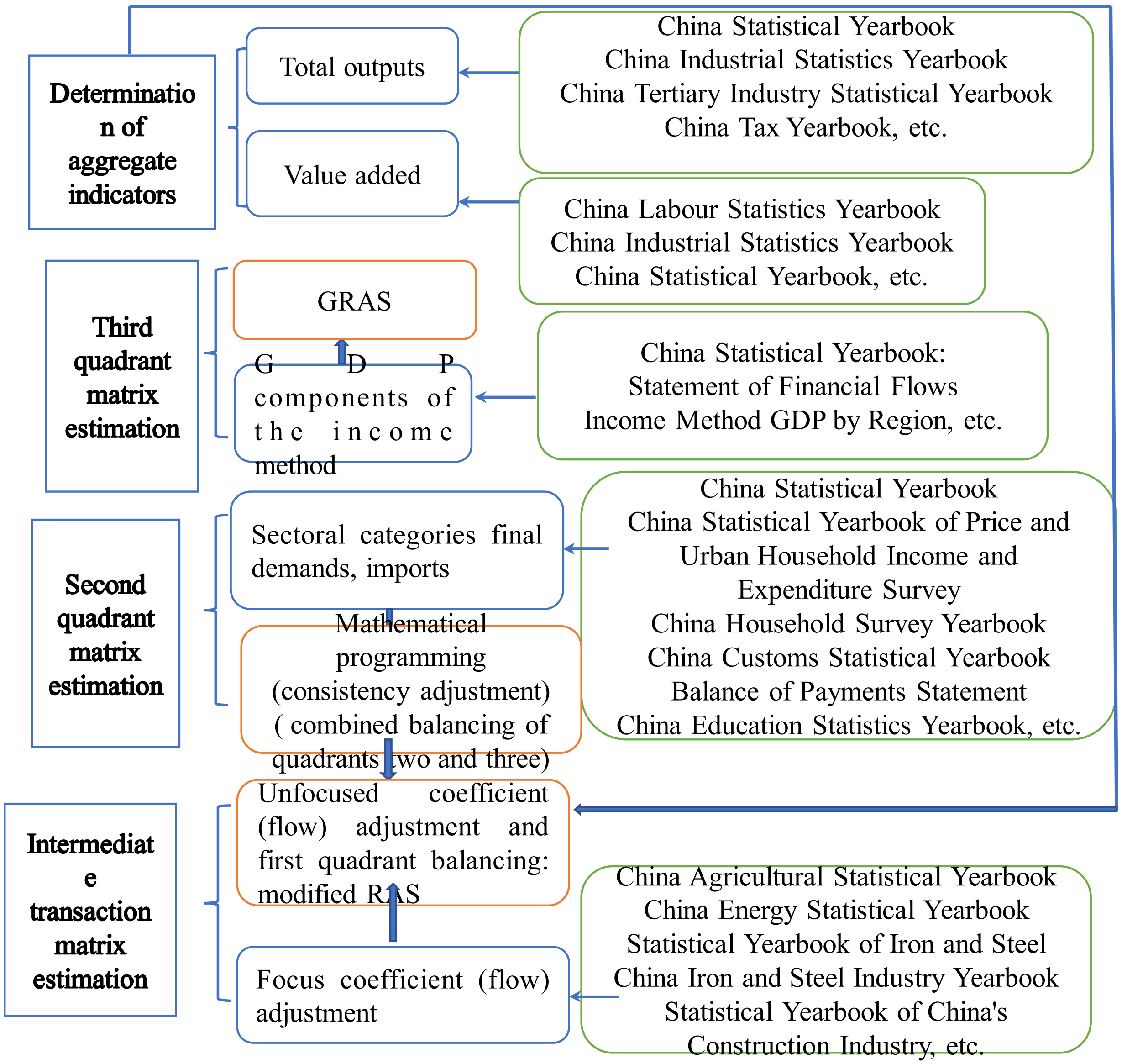
- The serial input-output tables compiled in this paper are product by product tables, with producer prices, which are consistent with the official input-output tables in China. Due to the long time span of the serial tables (from 1981 to 2018), the national economic accounting system in China has undergone significant adjustments during the period, with the gradual transition from the MPS system to the SNA system in the 1990s, and the input-output table compilation scheme of the National Bureau of Statistics also has significant changes, especially the differences between the table in 1987 and the input-output tables afterwards; meanwhile, in correspondence with the adjustment of the UN SNA system, the accounting system and the industrial classification of the national economy in China have also been adjusted several times, and accordingly, the sectoral categories of the input-output tables have also been modified.

Compilation Framework and Key Methodology

- In order to make full use of the statistical data information of each compilation year, generally, this paper has adopted the following compilation scheme: determining the corresponding sectoral categories and indicator estimation methods based on the information of different time periods and the compilation scheme of the official input-output tables in China, and completing the preliminary annual time series table in stages; then, identifying a unified and consistent sectoral category for the whole series on the basis of fully comparing the differences in sectoral categories and calibres of input-output tables in different time periods, and adjusting the sectoral categories and data consistency of the annual time series tables completed in stages to obtain a consistent series input-output table for the whole series from 1981 to 2018 according to the newly published updates of historical indicators of national economic accounts.

Table form of the serial input-output table in China

	Intermediate demand	Final demand	Imports	Total outputs
Intermediate inputs	Z	Y_1 Y_2 Y_3 Y_4 Y_5 Y_6	Y_7	X
Primary inputs	V_1 V_2 V_3 V_4			
	v			
Total inputs	X^T			


Determination of aggregate indicators
Total outputs
Value added
China Statistical Yearbook
China Industrial Statistics Yearbook
China Tertiary Industry Statistical Yearbook
China Tax Yearbook, etc.
China Labour Statistics Yearbook
China Industrial Statistics Yearbook
China Statistical Yearbook, etc.
Third quadrant matrix estimation
GRAS
GDP components of the income method
China Statistical Yearbook:
Statement of Financial Flows
Income Method GDP by Region, etc.
Second quadrant matrix estimation
Sectoral categories final demands, imports
Mathematical programming (consistency adjustment) (combined balancing of quadrants two and three)
China Statistical Yearbook
China Statistical Yearbook of Price and
Urban Household Income and
Expenditure Survey
China Household Survey Yearbook
China Customs Statistical Yearbook
Balance of Payments Statement
China Education Statistics Yearbook, etc.
Intermediate transaction matrix estimation
Unfocused coefficient (flow) adjustment and first quadrant balancing: modified RAS
Focus coefficient (flow) adjustment
China Agricultural Statistical Yearbook
China Energy Statistical Yearbook
Statistical Yearbook of Iron and Steel
China Iron and Steel Industry Yearbook
Statistical Yearbook of China's
Construction Industry, etc.

Compilation Framework and Key Methodology

- The main steps of compilation are as follows:

Step 1. Determine the aggregate control indicators, such as total outputs (X), value added (v), etc., for each sector;

Step 2. Estimate the initial matrices for the third quadrant and the second quadrant respectively;

Step 3. Adjust the two initial matrices in step 2, to ensure consistency between the second and third quadrants. This is because the second and third quadrants of the input-output table are production value-added for GDP and GDP calculated with expenditure approach respectively.

Compilation Framework and Key Methodology

The balancing process of the input-output table requires ensuring consistency between the second and third quadrants (i.e. the sum of the second quadrant is equal to the sum of the third quadrant). So the initial matrix of the second quadrant is adjusted using the production-accounted GDP of the third quadrant as the benchmark.

Compilation Framework and Key Methodology

Step 4. The compilation of the intermediate transaction matrix. The intermediate transaction matrix is the core of the input-output table and is estimated based on a combination of the important coefficients method and the RAS method. it also requires the row sum control numbers and column sum control numbers, which are respectively based on the total control indicators obtained in the first three steps and the estimated matrices in the second quadrant and the third quadrant.

Compilation Framework and Key Methodology

- The overall balancing relationship and logic in the compilation of the serial input-output tables is that firstly, the aggregate indicators and the third quadrant are compiled in line with the national accounts data; then, the second quadrant matrix determines the initial estimates based on GDP with expenditure method of the national accounts as well as the expenditure destination of the various categories of final demands, and further adjusts for consistency based on the aggregate values of the third quadrant (production method GDP), thus ensuring the consistency of the third quadrant with the second quadrant;

Compilation Framework and Key Methodology

finally, the rowwise and columnwise control number vectors of the intermediate transaction matrix are calculated from the total indicators and the data in the second and third quadrants, which, combined with the important coefficients, are used to estimate the intermediate flow matrix by using these control number vectors and the RAS method, thus ensuring that the balancing relationship between the first quadrant and the second and third quadrants holds.

Compilation Framework and Key Methodology

- Key methodology
  - Estimation method for the primary input matrix

Based on the available data base, estimates of the column sums and the row sums of the initial input matrix can be obtained. Firstly, the value added of each sector is given in the measure of the aggregate indicator, namely the total primary inputs of the subsectors, which is the column sum vector of the primary input matrix. Secondly, the available statistical information enables the estimation of the row sum of the primary input matrix, which is the national income method GDP for various sub-components such as labour compensation, depreciation of fixed assets, net production tax, operating surplus, etc. With both the row and column sums of the primary inputs matrix known, the GRAS method proposed by Junius and Oosterhaven (2003) has been used to estimate the elements of the primary inputs matrix.

Compilation Framework and Key Methodology

$$\min \sum_i \sum_j \left|V_{ij}\right| \ln \frac{V_{ij}}{V_{ij}^0}$$

$$s.t. \sum_j V_{ij} = v_i^R \qquad i = 1, \cdots, 4$$

$$\sum_i V_{ij} = v_j \qquad j = 1, \cdots, n$$

where V_{ij} denotes the quantity of the i-th primary input of the j-th sector, V_{ij}^0 denotes the initial value of the quantity of the i-th primary input of sector j, v_i^R denotes the total quantity of the i-th primary input (i.e. the row sum of the i-th primary input) and v_j denotes the quantity of value added of sector j. The initial values of the primary input matrix V can be derived from information of the sub-sectoral structure of the primary input matrix in the third quadrant of the known referable official input-output tables (including input-output tables and input-output extension tables).

Compilation Framework and Key Methodology

- Key methodology
  - **Estimation of the second quadrant matrix and method of adjusting for consistency of the second and third quadrants**

The second quadrant of the input-output table constitutes of sectoral rural consumption, urban consumption, government consumption, fixed capital formation, increase in inventories, exports and imports, i.e. the vector Y_i, i=1,⋯,7 in Table 1, where the front 6 items are final demand. These column vectors form the second quadrant matrix, the sums of which are components of GDP with expenditure approach. The compilation of the second quadrant of the time series input-output tables also makes full use of the statistical data and information closely related to the final use of products available in China, with reference to the compilation method of the input-output extension table published by the National Bureau of Statistics and using mathematical programming as an auxiliary method for the consistent adjustment of data in the second and third quadrants.

Compilation Framework and Key Methodology

The key issue to be addressed in the compilation of the second quadrant matrix is to achieve consistency between the second and third quadrants while making full use of publicly available statistical information from statistical offices, government institutions, etc.

Compilation Framework and Key Methodology

- Key methodology
  - **Estimation of the second quadrant matrix and method of adjusting for consistency of the second and third quadrants**

Based on a comprehensive consideration of the data base and the methodology for the compilation of input-output extension tables published by the National Bureau of Statistics, the basic idea of the second quadrant matrix estimation method has been identified as follows.Firstly, the initial values of sub-sectoral data for each type of final demand as well as imports, i.e. the column vector Y_i, i=1, ⋯, 7 in the second quadrant, are estimated based on the available data information;

Compilation Framework and Key Methodology

then the control numbers for each column in the second quadrant are set based on the components given by the GDP with expenditure approach, and the total value added in the third quadrant of the input-output table (i.e. GDP) is used as the total control number to build a mathematical programming model for the consistency adjustment of the second and third quadrants.

Compilation Framework and Key Methodology

- Key methodology
  - **Estimation of the second quadrant matrix and method of adjusting for consistency of the second and third quadrants**

Due to the difference in accounting methods, the results of the expenditure method GDP and the production method GDP accounting, although theoretically supposed to be equal, differ in the actual accounting because of errors and other factors. Expenditure method GDP and production method GDP are reflected in the second and third quadrants of the input-output table respectively. In accordance with the rowwise and columnwise balancing relationship of the input-output model, the sum of the second quadrant equals the sum of the third quadrant. Hence, in order to ensure the balances of the input-output table and the consistency of the second and third quadrants, the mathematical programming model in this paper has been constructed for consistency adjustment after the initial estimates for each sub-sectoral final demand were obtained.

Compilation Framework and Key Methodology

- Key methodology
  - **Estimation of the second quadrant matrix and method of adjusting for consistency of the second and third quadrants**

The aim of the consistency adjustment is to ensure that the sum values of the second and third quadrants of the input-output tables be equal, i.e. that they are both equal to the production method GDP published by the National Bureau of Statistics. A non-linear mathematical programming model for the minimisation of absolute deviations is established as follows：

$$\min \sum_{k=1}^{7} \sum_{i=1}^{n} w_k \left| y_{ik} - y_{ik}^0 \right|$$

$$\text{s.t.} \quad \sum_{i=1}^{n} y_{ik} = y_k \quad k = 1,\ 2 \ldots 5$$

$$\sum_{i=1}^{n} (y_{i6} - y_{i7}) = x\_nx$$

$$\sum_{i=1}^{n} \left(\sum_{k=1}^{6} y_{ik} - y_{i7} + \varepsilon_i \right) = va$$

$$(1 - \gamma^k) y_{ik}^0 \le y_{ik} \le (1 + \gamma^k) y_{ik}^0 \quad \gamma^k \in (0,1)$$

$$\bar{\varepsilon}_i^1 \le \varepsilon_i \le \bar{\varepsilon}_i^2 \quad i = 1, \cdots, n$$

$$y_{ik} \ge 0 \quad k \ne 5$$

Compilation Framework and Key Methodology

- Key methodology
  - **Method for compiling the intermediate transaction matrix Z in the first quadrant**

The key point in the compilation of the intermediate transaction matrix is to estimate the technical coefficient matrix. The National Bureau of Statistics (NBS) has adopted a combination of the important coefficients method and the RAS method in the compilation of input-output extension table, which makes full use of available statistics to achieve the best estimate of the intermediate transaction matrix in the absence of an input-output survey. Hence, the estimation of the intermediate transaction matrix in the compilation of time series IO tables in this paper is mainly based on the method of compiling input-output extension tables in China published by the National Bureau of Statistics.

Compilation Framework and Key Methodology

Let $A = A^1 + A^2$,where A^1 is the matrix of key coefficients whose entries of the important coefficients are greater than 0 and other coefficients equal to 0; A^2 refers to the non-key coefficient matrix, whose elements take values with key coefficients equal to 0 and non-key coefficients greater than or equal to 0. The key coefficients are first determined and estimated to obtain A^1, and the transaction matrix corresponding to the key coefficients is further calculated, i.e.$Z^1 = A^1\hat{X}$.

Compilation Framework and Key Methodology

Then, based on the estimation results of the total control indicators, the second quadrant and the third quadrant, the row control number vector *r* and column control number vector *s* of the intermediate transaction matrix are obtained, and the corresponding row and column control vectors of the transaction matrix $A^2\hat{X}$ corresponding to the non-key coefficients are: $\bar{r} = r - A^1\hat{X}i$,$\bar{s} = s - i^T A^1\hat{X}$.The initial estimate of the transaction matrix corresponding to the non-key coefficients is $Z^{20} = \bar{A}^2\hat{X}$, given that the non-key coefficient matrix for the reference year is $\bar{A}^2$. By adjusting Z^{20} via RAS method, the transaction matrix corresponding to the non-key coefficients is obtained as Z^2. Finally, the intermediate flow matrix for the year of compilation is $Z = Z^1 + Z^2$.

Estimation of Key Indicators

- It mainly includes aggregate control indicators such as sectoral total output and value added, initial estimates of important variables such as disaggregated final demand vectors and imports
  - Estimation methods for sectoral total outputs

The calculation methods provided in the methodology of the compilation of input-output extension tables published by the National Bureau of Statistics have been taken into account in the compilation of the total output indicators.

Estimation of Key Indicators

Generally speaking, the total outputs are compiled either directly as per the data from the statistical sources for those industries where the Chinese Statistical Yearbook or other statistical sources give indicators of total output, or by adjusting for the accounting calibre. For industries for which total output data are not given in the China Statistical Yearbook or other statistics, extrapolation is made according to the information given in the statistics on reference indicators such as sales output, main business income or total wages. Take industry as an example:

$$\text{Total industrial output for the year of compilation} = \text{Reference indicator for the year of compilation} \div \frac{\text{Reference indicator for the year of reference}}{\text{Total industrial output for the year of reference}}$$

Estimation of Key Indicators

- Estimation of secteral value added

Gross Domestic Product (GDP) has been the core indicator of the SNA, which, on a sectoral basis, is the value added of each sector. The statistical yearbooks published after the implementation of the SNA in China give more detailed data on the accounting of GDP. However, for sectoral value added data, the China Statistical Yearbook currently only provides value added by major sectoral categories. Of these, the information for the service industry has a more detailed disaggregation, and the China Tertiary Industry Statistical Yearbook for some years also provide a more detailed breakdown of the value added of the service industry. The industrial sector, on the other hand, lacks a breakdown of value added by sector, with data only available for major sub-categories such as extractive industries, manufacturing industries, and electricity, gas and water production and supply industries. Hence, the measurement of value added indicators lies in the imputation of value added in industrial sub-sectors. The data bases of the industrial sectors also vary over time, and this paper has designed different estimation methods for different data bases.

Estimation of Key Indicators

- Estimation methods for second quadrant indicators

In the case of consumption, for example, for urban and rural consumption, this paper has adopted the method of estimating the structure first and then estimating the sub-product consumption vector based on the aggregate. The urban household expenditure survey and the rural household expenditure survey, for which data are available, have provided data on per capita consumption expenditure by major categories as a reference for adjustment. The consumption structure by product category is first given an initial estimate based on the column vectors of urban and rural residents' consumption in the input-output table for the reference year. This is then adjusted by reference to the consumption structure of the sub-categories in that year to determine the sub-sectoral consumption structure of urban and rural residents in the year of compilation, and finally the initial estimates of the consumption of rural residents and urban residents are estimated by combining the total consumption of rural residents and the total consumption of urban residents given in the expenditure method GDP in the year of compilation.

Estimation of Key Indicators

For fixed capital formation, the similar method is used, estimating the sub-product structure first and then compute the fixed capital formation. The China Statistical Yearbook provides data on fixed asset investment by composition into several major categories such as construction and installation works, purchase of equipment, tools and appliances, other costs, etc. The structure of capital formation by major categories can be taken as a reference and then combined with the sub-sectoral structure of fixed capital formation given in the reference input-output table to determine the sub-sectoral structure of fixed capital formation for the year of compilation. Finally, the initial values of the sub-sectoral fixed capital formation vectors are estimated by combining the data on gross fixed capital formation given by the expenditure method GDP.

For imports and exports, the available statistics provide a good data base and are compiled in accordance with the method of input-output tables and input-output extension tables published by the National Bureau of Statistics.

Results and Comparative Analysis

- The time series input-output tables of China have been completed for the period of 1981 to 2018, with 37 sectors, in current price.
- The time series input-output tables we have compiled are compared with several other types of serial tables available in the literature, and the differences in the basic characteristics are shown in the table below.

Basic characteristics of the serial input-output tables in China in different databases

Database	Methodology	Number of sectors	Periods	Classification criteria	Type of IO table
世界投入产出数据库（WIOD）	RAS, Interpolation method	35 56	1995–1999 2000–2014	ISIC	Industry x Industry
中国产业生产率项目（CIP）	RAS, Interpolation method	37	1981–2010 the breakdown of value added ranges from 1987 to 2010	China National Account and EU_EKLEMS Category	Industry x Industry
Eora 全球多区域投入产出数据库（Eora MRIO）	RAS, Forward and backward circular weighted average method	123	1990–2015	National economic category	product x Iproduct
中国投入产出表序列（CIOD）	MTT，Liner interpolation method	42	1992–2020	National economic category 业	Products x Products
本文序列投入产出表（本文SIOT）	Statistical data-based estimation, RAS and GRAS, Mathematical programming	Full serial consistent categories: 37	1981–2018	Input-Output Tables Sectoral Categories by National Bureau of Statistics	Products x Products

Results and Comparative Analysis

- In terms of basic compilation methods, the several types of China's serial input-output tables in the literature are compiled in such a way that the aggregate indicators are mainly interpolated and averaged according to China's input-output tables in the benchmark years, and the methods of estimating technical coefficient matrix are mainly RAS method or matrix transformation, without taking full advantage of the various public released statistics and information in the years with no official published IO tables in China.

Results and Comparative Analysis

The method of compilation of the serial tables for China established in this study refers to the method used by the National Bureau of Statistics to compile input-output tables and input-output extension tables. The aggregate indicators and the second and third quadrant indicators are estimated by making full use of statistical information for China, while the technical coefficient matrix is a combination of key coefficient estimation based on the use of statistical information and the RAS method, and the consistency adjustment combines the mathematical programming, and GRAS method. Generally speaking, the method of compiling the sequential tables in this paper is characterised by the full use of available statistical information.

Results and Comparative Analysis

- In terms of sectoral categories, the serial table sectoral categories for this project are determined in accordance with the official input-output table compilation scheme and sectoral categories published by the National Bureau of Statistics, which are consistent with the official input-output tables of China. The serial tables have been prepared with both a uniform and consistent sectoral classification for the entire period, for 37 sectors, and a sectoral category specific to the characteristics of the different time periods (as determined by the sectoral categories of the official Input-Output Tables for the different time periods and the compilation manual).

Results and Comparative Analysis

In terms of time span, the time series input-output table of China compiled in this study covers the entire history of China's development from the early stage of reform and opening up to the present from 1981 to 2018, which covers a longer time span compared to the existing literature.In terms of the type of input-output table, the serial input-output table compiled in this study is a symmetrical input-output table of product x product, which is consistent with the official input-output table published in China.

Data Extensions

- The construction of a Serial Input-Output Database (SIOD) in China:

On the basis of time series input-output tables (competitive) for the years 1981 to 2018,

  - Compile an annual serial import matrix, non-competitive input-output table;
  - Compile comparable price series input-output tables (previous year prices)
  - Satellite accounts: other socio-economic accounting indicators, and resource and environmental indicators

Thank you！

zhanghx_c@126.com