Processing trade in Chinese interregional input–output tables

Duan, Yuwan; Dietzenbacher, Erik; Los, Bart; Yang, Cuihong

Published in:
Economic Systems Research

DOI:
10.1080/09535314.2021.2012430

Publication date:
2023

Processing trade in Chinese interregional input–output tables: construction and application

Yuwan Duan, Erik Dietzenbacher, Bart Los & Cuihong Yang

ECONOMIC SYSTEMS RESEARCH
2023, VOL. 35, NO. 4, 566–585
https://doi.org/10.1080/09535314.2021.2012430

Processing trade in Chinese interregional input–output tables: construction and application

Yuwan Duan[a], Erik Dietzenbacher[b], Bart Los[b] and Cuihong Yang[c,d]

[a]Central University of Finance and Economics, Beijing, People's Republic of China; [b]Faculty of Economics and Business, University of Groningen, Groningen, Netherlands; [c]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, People's Republic of China; [d]University of the Chinese Academy of Sciences, Beijing, People's Republic of China

ABSTRACT

We construct new interregional input–output tables for China, which can be used to analyze changes in the interindustry linkages within and between eight Chinese regions, and their consequences. We claim that analyses based on these tables yield more accurate results than analyses using existing interregional input–output tables for China, because our tables explicitly account for a typical feature of the Chinse economy: the importance of processing exports activities. These activities rely heavily on imported inputs and much less on inputs sourced from domestic regions. Accounting for such differences between processing exports and other production activities reduces aggregation biases. We illustrate the usefulness of the tables by computing supply chain fragmentation indices for China and quantifying the biases that are avoided by using our input–output tables instead of conventional ones. We make our tables (for 2002, 2007 and 2012) publicly available.

ARTICLE HISTORY
Received 23 May 2020
In final form 26 November 2021

KEYWORDS
Processing trade; interregional input–output tables; China; data construction; interregional trade

1. Introduction

The deepening of international production fragmentation that accelerated in the period 1995–2007 led to a boom in the trade of intermediate goods. Consequently, analyses of data on gross imports and exports no longer accurately reflect how trade affects value added and income. They were supplemented by studies adopting the so-called global value chain (GVC) perspective (Johnson, 2014, 2018; Timmer et al., 2013). To arrive at results using this GVC perspective, global intercountry input–output (IO) tables are necessary. Their construction is at the heart of several large research projects (see Tukker & Dietzenbacher, 2013, for an early overview). Many studies analyze the effects of the vast changes in the scale and structure of China's trade on the national economy (e.g. Aichele & Heiland, 2018; Chen et al., 2012; Duan et al., 2012, 2021; Koopman et al., 2012; Los et al., 2015). But can thorough analyses also be done for regional economies? This issue is important, given China's

CONTACT Yuwan Duan duanyuwan@126.com Central University of Finance and Economics, Beijing 100081, People's Republic of China

Supplemental data for this article can be accessed here. https://doi.org/10.1080/09535314.2021.2012430

substantial regional inequalities, in terms of both regional involvement in globalization and in the nation's economic growth. Reducing income inequalities across regions has always been high on the agenda of the Chinese government. A general conclusion of the existing literature is that China's interior benefits from China's export orientation by providing intermediates to the production of exported products in coastal regions (Duan & Yan, 2021; Jiang et al., 2017; Meng et al., 2017; Pei et al., 2017; Zhang et al., 2019).[1] Hence, issues regarding interregional trade in China are important. Decisions in this policy arena should be based on solid estimates.

In this paper, we introduce new interregional IO (IRIO) tables for China and make them publicly available. We argue that interregional supply-chain effects are captured more accurately in these tables than in currently available equivalents. We argue our new tables correct biases in earlier tables that ignore the prevalence of 'processing trade' activities, which are featured in China's economy. The term 'processing trade' refers to the three-step business activity of (i) purchasing all (or part of) the raw and auxiliary materials, components and parts, accessories and packaging materials abroad, after which (ii) these goods are processed and assembled, before (iii) the finished products are re-exported (Yang et al., 2015).

This feature of China's economy developed over time and became very unevenly distributed across regions.[2] According to China's Customs statistics, processing exports comprised about half of the gross exports since the early 1990s. From approximately 2002 onwards, however, this share started to decline, reaching 34% in 2016.

The left panel of Table 1 shows the shares of regional processing exports in the national totals. In 2012, for example, processing exports from East Coast and South Coast constituted 75% of China's total processing exports, while the inland provinces accounted for less than 15%.[3] Moreover, processing trade appears to play very different roles in the respective regions. The right panel of Table 1 gives the share of a region's processing exports in its total exports. It shows that processing exports in the South Coast, the most important exporting region of China, accounted for 48% of its total exports in 2012. In the inland region Northwest, however, this proportion is only 8%. From 2002 to 2012, the shares of processing exports have decreased for all regions, except the Central Regions and Southwest.

Imported materials for processing trade were exempted from taxes in China. This led to their heavy reliance on imported inputs (Yang et al., 2015). That is, in comparison to regular manufacturing production, processing exports involves only limited domestic activities. At the national level, this causes biased results in IO analyses that do not separate processing exports from other manufacturing production. For example, Pei et al. (2012) show how

[1] For example, Pei et al. (2017) demonstrate that interior regions tend to specialize in upstream activities (such as providing natural resources and raw materials) while coastal regions predominantly carry out downstream activities and export final products. Adopting an environmental perspective.

[2] Ma and Van Assche (2016) analyze the factors that affect the location choice of China's export processing plants.

[3] Northeast includes Heilongjiang, Jilin, and Liaoning; Northern Municipalities includes Beijing and Tianjin; North Coast includes Hebei and Shandong; East Coast includes Shanghai, Jiangsu, and Zhejiang; South Coast includes Guangdong, Fujian, and Hainan; Central Regions includes Shanxi, Henan, Hubei, Hunan, Anhui, and Jiangxi; Northwest includes Inner Mongolia, Shannxi, Ningxia, Gansu, and Xinjiang; Southwest includes Sichuan, Chongqing, Yunnan, Guizhou, Guangxi, Qinghai, and Tibet. See Appendix B of the Online Supplementary Information for a map. Note that Liaoning, Tianjin and Guangxi, which are coastal provinces, are respectively included in Northeast, Northern Municipalities and Southwest. Therefore, the share of processing exports in inland provinces is smaller than the sum of the shares in Northeast, Northern Municipalities, Central Regions, and Southwest.

Table 1. Distribution of processing exports across regions, 2002–2012.

	Distribution of processing exports (%)			Share of processing exports in all exports (%)		
	2002	2012	Change	2002	2012	Change
Northeast	4.1	2.9	−1.2	45.0	24.4	−20.6
Northern Municipalities	6.2	5.6	−0.6	57.9	21.4	−36.5
North Coast	6.0	6.8	0.8	42.3	37.3	−5.0
East Coast	24.5	34.4	9.9	43.4	40.4	−3.0
South Coast	56.6	40.8	−15.8	73.7	48.2	−25.5
Central Regions	1.1	4.7	3.6	16.0	38.2	22.2
Northwest	0.4	0.3	−0.1	16.5	7.8	−8.7
Southwest	1.0	4.7	3.7	23.6	40.7	17.1
Total	100.0	100.0	0.0	55.3	39.5	−15.8

Notes: The data are from Chinese customs, which provides Chinese annual trade statistics at the HS-8 digit level for different trade regime and cities. We aggregated them to the regional level.

it inflates China's exports contribution to economic growth.[4] Dietzenbacher et al. (2012) and Su et al. (2013) demonstrate how it overestimates the importance of international trade to China's emissions generation. Yang et al. (2015) show how it underestimates China's vertical specialization. Thus, similar biases in regional studies will also lead to misleading conclusions. So, what are the 'true' intra-regional and interregional linkages in China? Can we at least better articulately separate interregional flows for processing trade and other manufactured goods?

Of course, in the case of China, accurately answering these questions requires interregional IO tables that differentiate between the production of processing exports and other production. Such tables are not readily available from any official agencies and need to be constructed. At the national level, several IO tables have been constructed (Chen et al., 2001; Koopman et al., 2012; Ma et al., 2015; and Su et al., 2013). Chen et al. (2019) and the OECD have distinguished China's production of processing exports from other production activities in international IO tables. While various institutions have compiled IRIO tables for China, none have systemically distinguished processing trade activities across China's regions.[5]

To fill this gap, we construct IRIO tables with separated processing exports (henceforth, 'IRIOP tables') for 2002, 2007, and 2012. We describe explicitly how we harmonize, reconcile and merge information from a wide range of data sources to arrive at these new tables.[6] We make the tables publicly available as online supplementary material. This enables an important foundation for extensive research on many important issues in China, some

[4] Pei et al. (2012) analyzed the contribution of changes in exports to China's value-added change between 2002 and 2007. They found to be 32% higher when the ordinary IO tables were used than when the tables capturing processing trade were used. Amiti and Freund (2010) found a significant skill upgrade in China's total exports between 1992 and 2005, but they found no evidence of skill upgrading when processing exports were excluded from total exports. See also Dean et al. (2011), Johnson and Noguera (2012), Koopman et al. (2012), and Upward et al. (2013).

[5] Pei et al. (2017) only split the processing exports from the total exports in the IRIO tables (without estimating the intermediate input coefficients and value added coefficients of processing exports industries). They admit that this is only a half-way treatment, and formal IRIO tables that fully distinguishing processing trade are required to conduct more accurate analyses.

[6] The IRIOP tables can be downloaded (as excel-files) from the Online Supplementary Information on the journal's website, or the corresponding author's personal homepage.

of which we already have briefly discussed. The tables can also be used to answer questions related to environmental policy, if linked to satellite accounts on energy use, carbon emissions, environmental pollutants, etc.

By way of illustration of the relevance of the IRIOP tables, we also show how disregarding the different production technologies and sourcing patterns of processing exports activities and how regular production activities lead to biases in computing recently developed indicators of supply chain fragmentation (Timmer et al., 2021). This has important implications for investment in infrastructure related to interregional transportation.

2. IRIOP tables and their construction

2.1. Structure of IRIOP tables

The general structure of an IRIOP table is shown in Figure 1, which depicts a two-region case to keep the exposition simple. Compared with a standard IRIO table, the unique feature of the IRIOP table is that the production of each sector in each region is divided into two types: the production of processing exports and ordinary production. Processing exports includes two types of export regimes: 'Processing & Assembly' (P&A) exports and 'Processing with Imported Materials' (PIM) exports.[7] Ordinary production then incorporates all production for the domestic market and the production of 'ordinary' exports (i.e. any exports but processing exports).

Before introducing the main parts of Figure 1, we introduce some notational principles regarding the use of indices. The subscript indices i, j, and k represent sectors, the subscript indices r, s, l, and w represent regions, and the superscripts represent the different types of production and products, that is, (output of) ordinary production activities (O), (output of) processing exports production activities (P), and (output of) foreign sectors, i.e. imports (M).[8] The cells in the IROP table contain monetary values (10,000 of RMB), expressed in current, basic prices.

An IRIOP table contains several parts: the matrix $\mathbf{Z}_{rs}^{OP}$ (with typical element z_{rsij}^{OP}) gives the intermediate deliveries from ordinary production in region r to processing exports production in region s[9]; the matrix $\mathbf{Z}_{s}^{MO}$ (with typical element z_{sij}^{MO}) gives the imports used as intermediate inputs for ordinary production in region s; the (row) vector $(\boldsymbol{v}_{s}^{P})'$ (with typical element v_{sj}^{P}) gives the amounts of value added generated in the production of processing exports in region s; the vector $\mathbf{x}_{s}^{O}$ (with typical element x_{si}^{O}) gives the gross output levels of ordinary production in region s; the vectors $\mathbf{e}_{r}^{P}$ (with typical element e_{ri}^{P}) and $\mathbf{e}_{r}^{O}$ (with typical element e_{ri}^{O}) contain processing exports and ordinary exports by region r, respectively; the vectors $\mathbf{c}_{rs}^{O}$ (with typical element c_{rsi}^{O}) and $\mathbf{f}_{rs}^{O}$ (with typical element f_{rsi}^{O}) represent

[7] P&A trade and PIM trade differ in terms of ownership and payment for the imported materials. Under P&A trade, materials and components are supplied by a foreign company and processed by a Chinese enterprise on a consignment basis. Ownership of raw materials and components remains with the foreign company. The Chinese enterprise (i) does not pay for the imported materials, and (ii) receives a processing fee. After processing and assembly, the finished products are owned by the foreign company, which distributes them further. In contrast, under PIM trade, a Chinese enterprise purchases the raw materials and components. It becomes the owner of the imported commodities. After processing and assembly, the Chinese enterprise exports the finished products to foreign customers.

[8] Bold-faced lowercase letters are used to indicate vectors, boldfaced capital letters indicate matrices, italicized lowercase letters indicate scalars (including elements of a vector or matrix). Vectors are columns by definition, row vectors are obtained by transposition, denoted by a prime (e.g. $\mathbf{x}'$). Diagonal matrices are denoted by a circumflex (e.g. $\hat{\mathbf{x}}$).

[9] For brevity, we give the meanings of parts of the table related to specific activities (e.g. O) or pairs thereof (e.g. OP). Similar meanings apply for similar parts of the table.

Figure 1. Schematic outline of the IRIOP table (2-region case).

		Intermediate use				Final use						*TOT*
		Region r		*Region s*		*Region r*		*Region s*		*INV*	*EXP*	
		P	*O*	*P*	*O*	*DFC*	*FCF*	*DFC*	*FCF*			
Region r	*P*	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{e}_r^P$	$\mathbf{x}_r^P$
	O	$\mathbf{Z}_{rr}^{OP}$	$\mathbf{Z}_{rr}^{OO}$	$\mathbf{Z}_{rs}^{OP}$	$\mathbf{Z}_{rs}^{OO}$	$\mathbf{c}_{rr}^O$	$\mathbf{f}_{rr}^O$	$\mathbf{c}_{rs}^O$	$\mathbf{f}_{rs}^O$	$\mathbf{q}_r^O$	$\mathbf{e}_r^O$	$\mathbf{x}_r^O$
Region s	*P*	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{e}_s^P$	$\mathbf{x}_s^P$
	O	$\mathbf{Z}_{sr}^{OP}$	$\mathbf{Z}_{sr}^{OO}$	$\mathbf{Z}_{ss}^{OP}$	$\mathbf{Z}_{ss}^{OO}$	$\mathbf{c}_{sr}^O$	$\mathbf{f}_{sr}^O$	$\mathbf{c}_{ss}^O$	$\mathbf{f}_{ss}^O$	$\mathbf{q}_s^O$	$\mathbf{e}_s^O$	$\mathbf{x}_s^O$
IMP		$\mathbf{Z}_r^{MP}$	$\mathbf{Z}_r^{MO}$	$\mathbf{Z}_s^{MP}$	$\mathbf{Z}_s^{MO}$	$\mathbf{c}_r^M$	$\mathbf{f}_r^M$	$\mathbf{c}_s^M$	$\mathbf{f}_s^M$	$\mathbf{q}^M$	$\mathbf{0}$	$\mathbf{m}$
VA		$(\mathbf{v}_r^P)'$	$(\mathbf{v}_r^O)'$	$(\mathbf{v}_s^P)'$	$(\mathbf{v}_s^O)'$							
TOT		$(\mathbf{x}_r^P)'$	$(\mathbf{x}_r^O)'$	$(\mathbf{x}_s^P)'$	$(\mathbf{x}_s^O)'$							

Notes: P = production of processing exports; O = other (or ordinary) production; DFC = domestic final consumption; FCF = fixed capital formation; INV = inventory changes; EXP = exports; TOT = gross sector outputs or total imports; IMP = imports; VA = value added.

the sectoral deliveries of ordinary products from region r for final consumption and fixed capital formation in region s, respectively; the vector $\mathbf{q}_r^O$ (with typical element q_{ri}^O) gives the sectoral inventory changes of ordinary products in region r; the vectors $\mathbf{c}_s^M$ (with typical element c_{si}^M) and $\mathbf{f}_s^M$ (with typical element f_{si}^M) give the sectoral imports used for final consumption and fixed capital formation in region s; the vector $\mathbf{m}$ gives the sectoral total imports by all regions. Note that re-exports are not included.

2.2. Overview of available information

The availability of consistent and reliable data is often regarded as a major barrier to construct a new IO table (Peters et al., 2011). In the ideal case, data for all the variables in the IRIOP table are obtained through a series of comprehensive surveys. But conducting such surveys is extremely time-consuming and expensive, i.e. such data are not readily available. Therefore, we develop a semi-survey approach that uses a combination of survey data, some proportionality assumptions, and RAS.[10]

We primarily used four data sources. First, national IO tables that capture processing trade (NIOP, for more see Table A.1 in Appendix A of the Online Supplementary Information), a set of conventional IRIO tables, international trade statistics from China's Customs, and China's Regional Economic Accounts (REA). The Chinese Academy of Sciences (CAS)

[10] The RAS method is commonly used to bi-proportionally scale a matrix of unbalanced preliminary estimates of an unknown real matrix to prescribed row and column sums (Stone & Brown, 1962; Bacharach, 1970; Lenzen et al., 2009, 2014). It is a mainstream semi-survey technique and is widely used in estimating interregional trade flows (Mi et al., 2018; Zheng et al., 2019). While other complex forms of the algorithm exist, we stick to the conventional algorithm to balance each block of the tables.

and the National Bureau of Statistics (NBS) compiled the NIOP tables.[11] They include 42 sectors and are available for 2002, 2007, and 2012 (see Chen et al., 2001; Chen et al., 2012).[12]

The IRIO tables are derived from the national IO tables and provide all inter- and intra-regional deliveries in a country. The State Information Center (SIC) and the NBS (Zhang & Qi, 2012) compiled the IRIO tables that we used. They are available for 2002, 2007, and 2012, cover eight regions but just 17 sectors (see Appendix B of the Online Supplementary Information for the definition of the regions and Appendix C for the sector classification).[13] Provinces are grouped into regions not only based upon geographical proximity but also based on the similarity of their macroeconomic environments (e.g. stages of economic development). While the regions assure the capture of several key economic characteristics, they inevitably obscure much economic heterogeneity.

The third type of data that we used are the international trade statistics from China's Customs. For each province, they have detailed data on exports and imports based on the origins and destinations of the deliveries of goods.[14] The imports of a province thus give the value of the foreign commodities that are consumed or used in this province. The exports of a province indicate the delivery abroad of commodities for which the production, or the final assembly, or the original dispatch occurred in this province (NBS, 2017). The data are not only by commodity (at eight-digit Harmonized Commodity Description and Coding System, i.e. HS for short), but also by trade regime (e.g. P&A trade, PIM trade, and ordinary trade) and by firm type (e.g. foreign-invested enterprises and domestic enterprises). The HS eight-digit data are further aggregated using the NBS concordance table, which yields the trade data of processing trade and ordinary trade classified by the IO sector at the regional level.

The last data source we have used are the REAs published annually by the NBS. They provide value added for several broad industries (including agriculture, manufacturing, construction, trade and transport, and other service sectors) at the province level.[15] Moreover, the provincial totals for final consumption, fixed capital formation, and inventory changes are also taken from the REAs.

---

[11] Koopman et al. (2012) also constructed a 2007 NIOP table using quadratic optimization. We used the tables from the CAS and the NBS, since they are semi-official, publicly accessible and are available for more years.

[12] The original CAS and NBS NIOP tables are not 'bipartite' but rather are 'tripartite'. This is because 'other production' was split into two parts: production of domestic enterprises to satisfy domestic demand, and a combination of the production of non-processing exports and the production of foreign invested enterprises to satisfy domestic demand. Due to data limitations, however, this split of other production cannot be made at the regional level.

[13] Also, other institutes (e.g. the Development Research Centre of the State Council of China) have constructed Chinese IRIO tables, using different compilation methods and different classifications of provinces into regions. We have used the SIC-NBS tables not only because they are semi-official and publicly accessible but also because they adhere to the division of regions that is most common in China.

[14] In China's Customs statistics, two types of provincial trade data are available. One type is based on the source or destination of a delivery, the other type is according to the location of trading companies. Usually, the two types of trade data show significant differences. Products that are produced in one province are often exported (or imported) by a trading company that is located in a different province (Fu, 2004). For example, the exports of Guangdong by origin of goods are about 9.3% larger than the exports by location of trading companies in 2016 (NBS, 2017). In this paper, we have used data based on provincial trade statistics by origin or destination of a delivery. This is because in an interregional IO table we are particularly interested in knowing the region in which an imported product is used and in knowing the region in which the exported goods were produced.

[15] The correspondence between this broad five-category classification and the 17-sector classification in the interregional IO table is given in Appendix C of the Online Supplementary Information.

2.3. Inconsistency issues and some underlying construction principles

2.3.1. Inconsistencies

The use of different data sources implies conflicts between the sources. For example, ideally, aggregating the NIOP table over the two types of production and aggregating the IRIO table over the regions should both give the national IO table. This is not the case, however; all variables (except value added) show clear discrepancies.[16] This is not surprising because the NIOP and the IRIO tables are based on different data. The NIOP tables are based on the official national IO tables and the IRIO tables are largely based on provincial IO tables. In the rest of this subsection, we summarize the main inconsistency issues across data sources.

First, P&A imports are treated differently in NIOP tables from IRIO tables and national IO tables for 2007 and 2012. Recall that P&A imports are materials that are owned by foreign companies and that are supplied to Chinese enterprises to produce P&A exports. This production involves processing and assembly activities for which the Chinese enterprises receive a processing fee. The national IO tables and the IRIO tables record just these processing fees and not the P&A imports and exports themselves.[17] The NIOP tables, however, aim at reflecting the underlying technology of the processing sector. All imports (including P&A imports) used to produce processing exports are therefore recorded as intermediate inputs. As a result, imports, exports, and outputs are all larger in the NIOP tables than in the IRIO and the national IO tables. If we subtract the P&A imports from the corresponding items in the NIOP table, it is basically consistent with the official national IO table.[18]

Second, the value of trade differs considerably at the regional level between the IRIO table and China's Customs' statistics. Part of the explanation is that P&A imports are included in both the imports and the exports in China Customs' statistics, but not in the imports and exports in the IRIO tables. However, after deducting the P&A imports from Customs' statistics, considerable gaps still remain between the two sources. In the NIOP tables, on the other hand, merchandize trade is basically consistent with China's Customs statistics and with the official national IO tables (provided we correct for the different treatment of P&A imports).

Third, discrepancies exist also between IO tables and the REAs. For example, in 2007, Chinese GDP in the NIOP table is 3.5% lower than national value added (i.e. provincial GDP summed over the 31 provinces) in the REAs. This is not surprising, given the huge gaps between the sum of provincial GDP and nationwide total GDP mentioned by Holz (2004). Sources are not consistent at the regional level either. For example, value added in South Coast is 5.7% lower in the IRIO table than in the REAs.[19]

[16] For example, the total exports in 2007 are 24.6% larger in the NIOP table than in the IRIO table.

[17] China's national and regional IO tables treat processing trade following the SNA 2008 system. Chinese customs provide the value of total exports and P&A exports, while BOP (Balance of payment) provides the information for processing fee. The export value in official IO tables is arrived by minus P&A exports from total exports plus processing fee. It worth mentioning that though processing fee is accounted as part of service trade in BOP, it is treated as trade in goods in Chinese IO tables.

[18] This inconsistency issue does not exist for 2002. For that year, P&A imports were included both in the national IO table and the IRIO table. This was in line with the SNA 1993 guidelines.

[19] According to Zhang and Qi (2012), the value added in the IRIO table is obtained by adjusting the value added in provincial IO tables to the official national IO tables.

2.3.2. Underlying principles and choices

The inconsistencies discussed above imply that choices and adjustments need to be made in the actual construction of an IRIOP table. In this subsection, we describe some choices that we made and their underlying principles.

First, reliable and accurate trade statistics received the highest priority in connection to the separation of processing exports production from ordinary production – the distinguishing feature of IRIOP tables. This implies that the NIOP table (rather than the IRIO table), which are basically consistent with China's Customs' statistics (while the IRIO tables are not), was our preferred data source. This is because China's Customs is the only official institute to monitor international trade and is responsible for the trade statistics following strict regulations.

So, we use the NIOP table as a benchmark and distribute two types of production in it across regions. Thus, a spatially aggregate IRIOP table yields a NIOP table. The distribution is obtained from the IRIO tables, which reflect the input linkages across regions. Another advantage of our approach, which takes the NIOP table as the benchmark, is that all imported materials for processing exports (including P&A imports) are recorded as intermediate inputs in the IRIOP tables. Thus, they provide a full picture of the input structure of processing exports production.

We give the REAs the next highest priority; They are officially published and more authoritative than other data sources. Data on regional value added, final consumption, and fixed capital formation are taken from the REAs and not from the IRIO tables, which also embody these variables. But data from the REAs first must be rescaled, so they match with the NIOP table – the national benchmark for our approach.

Third, we match the 42 sectors in the NIOP tables to the 17 sectors in the IRIO tables. We could have tried to disaggregate certain IRIO sectors, because more sector detail generates more accurate results (as argued by Su et al., 2010). But splitting sectors requires additional data that were not readily available and, further, might lead to results as biased as the aggregates if the additional data are sufficiently inaccurate. Therefore, we have chosen to aggregate the NIOP tables in line with the IRIO sector classification. (See Appendices B and C of the Online Supplementary Information for the details on regions and sectors.)

2.4. The construction approach

To construct the IRIOP tables, the variables in Figure 1 are estimated. We estimate them blockwise and balance the IO tables blockwise as well, instead of balancing the table all at once as many others have (Mi et al., 2018; Zheng et al., 2020). By doing so, we keep the table as consistent as possible with available data and minimize information loss. The estimation in Step 1 includes the processing exports and ordinary exports, domestic inventory changes, and outputs, all at region-sector level; Step 2 covers the imported final consumption and fixed capital formation; Step 3 deals with value added and Step 4 estimates the intermediate use of imports (i.e. the import matrices). These four steps together with the NIOP tables constrains the domestic intermediate-input matrix, while the domestic final-demand matrix ensures balanced IRIOP tables.

Step 5 yields the intermediate input use, the final consumption, and the fixed capital formation, all for domestically produced goods and services. Steps 3, 4 and 5 assure reasonable initial estimates for each variable upon which the RAS balances the table and assures its consistency with the official statistics. In the rest of this subsection, we sketch the main construction steps; details are in Appendix D in the Online Supplementary Information.

2.4.1. Step 1: exports, inventory changes, and outputs

This step estimates the exports (vectors $\mathbf{e}_r^P$ and $\mathbf{e}_r^O$ in the case of country r), inventory changes ($\mathbf{q}_r^O$), and outputs ($\mathbf{x}_r^P$ and $\mathbf{x}_r^O$) in IRIOP tables.

Exports. For manufacturing sectors, region-sector data for processing exports and for ordinary exports are obtained by aggregating data from China's Customs' statistics. The region-sector data are scaled to match the export data in the NIOP tables. Processing exports of services mainly are wholesale and retail trade margins but also are other services used when processing exports of manufactured goods. Estimates of processing exports of services at the regional level are regional distributions of the processing exports of services in the NIOP table and based upon each region's share of processing exports in manufacturing. Ordinary exports of services at the regional level are estimated via the regional distribution of ordinary exports of services from the NIOP tables by using the regional distribution of sector-specific services exports in the IRIO tables.

Domestic inventory changes. Note that a region's inventory changes are the inventory changes in the whole of China regarding products delivered by that particular region. We adopted a two-step procedure to estimate it at the region-sector level (the elements q_{ri}^O of the vector $\mathbf{q}_r^O$). First, we took regional totals of the inventory changes for region r from the REAs and distributed over the sectors using shares based on the sector-specific inventory changes for region r from the IRIO table (i.e. $\tilde{\mathbf{q}}_r^O$, where a tilde indicates that the variable is from an IRIO table). This yielded the initial estimate of q_{ri}^O. Next, we rescaled these initial estimates such that the aggregation over the regions equals the inventory changes in the NIOP table (i.e. $\bar{\mathbf{q}}^O$, where an overbar indicates that the variable is from a NIOP table).

Outputs. By definition, the output of processing exports production is the exports themselves (i.e. $\mathbf{x}_r^P = \mathbf{e}_r^P$). To arrive at the regional output of ordinary production ($\mathbf{x}_r^O$), we first estimate the domestic sales of regional ordinary production ($\mathbf{d}_r^O = \mathbf{x}_r^O - \mathbf{q}_r^O - \mathbf{e}_r^O$). At the national level, the domestic sales are obtained from subtracting the ordinary exports and inventory changes from the output of ordinary production in the NIOP tables, i.e. $\bar{\mathbf{d}}^O = \bar{\mathbf{x}}^O - \bar{\mathbf{q}}^O - \bar{\mathbf{e}}^O$. Next, we distributed these domestic sales by regions of origin using the regional distribution of domestic sales in the IRIO table, i.e. $\tilde{\mathbf{d}}_r = \tilde{\mathbf{x}}_r - \tilde{\mathbf{q}}_r - \tilde{\mathbf{e}}_r$. This then enabled estimates for $\mathbf{d}_r^O$, after which we defined regional gross output levels in the IRIOP table via $\mathbf{x}_r^O = \mathbf{d}_r^O + \mathbf{q}_r^O + \mathbf{e}_r^O$.

2.4.2. Step 2: imported final demands

This step estimates the vectors of imported final consumption ($\mathbf{c}_r^M$), imported fixed capital formation ($\mathbf{f}_r^M$), and imported inventory changes ($\mathbf{q}^M$). We distinguish between imports that are only used for producing processing exports (which we call processing imports) and imports for other purposes (which we call non-processing imports). We also distinguish between merchandise goods (sectors 1–15) and services (sectors 16 and 17). By definition, only non-processing imports can be used as final demand. Therefore, for merchandise goods, we adopted the revised BEC ('broad economic categories') (Dietzenbacher et al., 2013) to allocate the region-commodity-specific nonprocessing imports from China's Customs statistics into three use categories: 'final consumption', 'fixed capital formation', and 'intermediates'.[20] We then aggregated the data into the 17 IRIO sectors and scaled them to match with the imported final demands from the NIOP table. We thus obtained the regional imports used for consumption (c_{ri}^{M}) and for fixed capital formation (f_{ri}^{M}). In addition, we also obtained the total imports that are used as intermediate inputs in ordinary production ($\sum_j z_{rij}^{MO}$ in the IRIOP table).

Processing imports are, by definition, exclusively used as intermediate inputs in the production of processing exports. We aggregate the region-commodity-specific processing imports from China's customs statistics into the 17 IRIO sectors and scale them to match the processing imports from the NIOP table. For merchandise goods, this yields the region-commodity-specific processing imports in the IRIOP table (i.e. the totals $\sum_j z_{rij}^{MP}$). For services, we distribute the sector-level total imports data from the NIOP table over the regions using the allocation of merchandize imports. The imported inventory changes (q_i^{M}) are directly obtained from the NIOP table.

2.4.3. Step 3: value added

Values added for sector j in region s is obtained via RAS. The initial estimates follow from the value-added coefficients from the NIOP tables (i.e. $\bar{v}_j^P/\bar{x}_j^P$ and $\bar{v}_j^O/\bar{x}_j^O$). These are combined with region-sector outputs from step 1 (i.e. x_{sj}^P and x_{sj}^O). The column constraints follow directly from the NIOP table, i.e. $\sum_s v_{sj}^P = \bar{v}_j^P$ and $\sum_s v_{sj}^O = \bar{v}_j^O$. For the row constraints, we use the REA data. They include value added at the regional level (not distinguishing between P and O) for five broad sectors. One is manufacturing, IRIO sectors 2–14. Its total value added is $\ddot{v}_{s,manufacuring}$, which we split into value added by IRIO sector using region-sector-specific value added from the IRIO table ($\tilde{v}_{sj}$) to get IROP value added $\ddot{v}_{sj}$ for all 17 IO sectors. By summing, we can get national value added for industry j as $\bar{v}_j^P + \bar{v}_j^O$ (using the NIOP table), as well as $\sum_s \ddot{v}_{sj}$ (using REA data). Given the inconsistency between data from REAs and NIOP tables and our preference for benchmarking on NIOP data, we split $\bar{v}_j^P + \bar{v}_j^O$ by region using $\ddot{v}_{sj}$. This constrains rows $v_{sj}^P + v_{sj}^O = (\bar{v}_j^P + \bar{v}_j^O)\left(\ddot{v}_{sj}/\sum_r \ddot{v}_{rj}\right)$. Finally, we apply RAS using the initial estimates and the set of given row and column sums. Subsequently, we also estimate the components of value-added (i.e. compensation of labor, fixed asset depreciation, net indirect production taxes, and operating surplus) for all industries and all regions by both types of production (production of processing exports and ordinary production).

2.4.4. Step 4: import matrices

We use RAS again to estimate industry use of imports by region. As initial values we use information from the NIOP tables. We then assume that for the same production type (e.g. processing exports P), the national use of imported intermediates applies across all regions, although levels of use vary across them. The initial import matrices in the IRIOP table are $z0_{sij}^{MP} = \bar{z}_{ij}^{MP}[(x_{sj}^P - v_{sj}^P)/(\bar{x}_j^P - \bar{v}_j^P)]$, where x_{sj}^P and v_{sj}^P are obtained from the IRIOP table

[20] A major advantage of this revised BEC method is that it allows a good to go into more than one category.

(and estimated earlier) and $\bar{z}_{ij}^{MP}$ is the import matrix of processing exports from the NIOP table.

There are three constraints that should be satisfied by z_{sij}^{MP} and z_{sij}^{MO}. First, summing z_{sij}^{MP} and z_{sij}^{MO} over the regions should yield the corresponding import matrix in the NIOP table. That is, $\sum_s z_{sij}^{MP} = \bar{z}_{ij}^{MP}$ in the case of production of processing exports. Second, aggregating z_{sij}^{MP} and z_{sij}^{MO} over the destination sectors (e.g. $\sum_j z_{sij}^{MP}$) yields the total imports of product i by region s for intermediate use. These totals were obtained in step 2 above. Third, for each region-sector the sum of imported inputs cannot be larger than the total amount of inputs that is required. In other words, the domestic intermediate inputs cannot be negative. RASing results in final estimates for z_{sij}^{MP} and z_{sij}^{MO}.

2.4.5. Step 5: domestic intermediate deliveries and final demands

In this step, we estimate the matrices with domestic intermediate deliveries ($\mathbf{Z}_{rs}^{OO}$ with elements z_{rsij}^{OO} and $\mathbf{Z}_{rs}^{OP}$ with elements z_{rsij}^{OP}) and the domestic final demand vectors ($\boldsymbol{c}_{rs}^{O}$ with elements c_{rsi}^{O} and $\mathbf{f}_{rs}^{O}$ with elements f_{rsi}^{O}). All these variables should be consistent with the variables estimated in prior steps to ensure balanced IRIOP tables. To this end, we adopt a hierarchical estimation method. We first estimate the sub-aggregates of the elements in the target matrix, and then use the sub-aggregates as column or row constraints to estimate the target matrix. By doing so, we minimize information loss.[21] The estimation procedure includes three parts, each of which uses RAS.

In step 5.1, we estimate the *total* intermediates and the *total* final demands of each product i, provided by each origin region r. The total is obtained by summing over the destination regions. The estimation is done by splitting the total region-sector-specific domestic sales that we obtained in Step 1 (i.e. d_{ri}^{O}) into the three use categories: total intermediate use (y_{ri}^{O}), total final consumption (ct_{ri}^{O}), and total fixed capital formation (ft_{ri}^{O}). Initial values for RAS are assigned by splitting the sector-level national intermediate use (or national final consumption, or national fixed capital formation) from the NIOP tables among origin regions using information from the IRIO tables. Meanwhile, y_{ri}^{O}, ct_{ri}^{O}, ft_{ri}^{O} are subject to two constraints. First, aggregating y_{ri}^{O}, ct_{ri}^{O}, and ft_{ri}^{O} across origin regions should exactly yield the corresponding elements at the national level in the NIOP table. Second, the sum of y_{ri}^{O}, ct_{ri}^{O}, and ft_{ri}^{O} must equal d_{ri}^{O} for any r and i.

In step 5.2, we estimate the domestic final demands (c_{rsi}^{O}, f_{rsi}^{O}) by using RAS again. The initial values for c_{rsi}^{O} and f_{rsi}^{O} are obtained from the IRIO tables. Meanwhile, the c_{rsi}^{O} and f_{rsi}^{O} are subject to two constraints. First, aggregating c_{rsi}^{O} and f_{rsi}^{O} over the destination regions s gives the totals that were estimated previously in part 5.1 (i.e. ct_{ri}^{O} and ft_{ri}^{O}). Second, aggregating c_{rsi}^{O} (f_{rsi}^{O}) over the origin regions and adding the imported final consumption c_{si}^{M} (imported fixed capital formation f_{si}^{M}) must yield the total consumption (fixed capital formation) consumed by each region, which is obtained by adapting NIOP data with information from the REAs.

[21] The hierarchical estimation here follows a similar logic to TRAS (three-stage RAS). However, there are too many additional constraints other than the row and column sums, so we use the hierarchical estimation to simplify the algorithm. The estimation can make sure the domestic intermediate deliveries and final demand are exactly consistent with the variables in the NIOP tables, and therefore will yield more accurate estimations with minimum information loss (Gilchrist & St Louis, 1999).

In step 5.3, we estimate the domestic intermediate deliveries (z_{rsij}^{OP} and z_{rsij}^{OO}). To this end, we first allocate the intermediate deliveries from the NIOP table over the destination regions (that is, we first estimate $zt_{sij}^{OP} = \sum_r z_{rsij}^{OP}$ and $zt_{sij}^{OO} = \sum_r z_{rsij}^{OO}$). Then, we allocate these estimates over the origin regions.

Based on previous steps, the IRIOP tables contain estimates for the total intermediate inputs used by each region s and sector j. That is, $x_{sj}^{P} - \sum_i z_{sij}^{MP} - v_{sj}^{P}$ and $x_{sj}^{O} - \sum_i z_{sij}^{MO} - v_{sj}^{O}$. These are further split over sectors of origin using the input structure in the IRIO tables. This yields the initial values for zt_{sij}^{OP} and z_{rsij}^{OO}. Further, zt_{sij}^{OP} and z_{rsij}^{OO} are subject to two constraints. First, aggregating zt_{sij}^{OP} (z_{rsij}^{OO}) over regions should give the national intermediates deliveries in the NIOP tables. Second, aggregating zt_{sij}^{OP} (z_{rsij}^{OO}) over the origin sectors should equal to $x_{sj}^{P} - \sum_i z_{sij}^{MP} - v_{sj}^{P}$ (or $x_{sj}^{O} - \sum_i z_{sij}^{MO} - v_{sj}^{O}$). Based on the initial values and constraints, the solution for zt_{sij}^{OP} and zt_{sij}^{OO} is obtained from the RAS procedure.

Then we estimate z_{rsij}^{OP} and z_{rsij}^{OO} by allocating zt_{sij}^{OP} and zt_{sij}^{OO} over the origin regions. For the initial values we use the IRIO tables for this allocation. Meanwhile, z_{rsij}^{OP} and z_{rsij}^{OO} are subject to two constraints. First, aggregating z_{rsij}^{OP} and z_{rsij}^{OO} over destination regions s and destination sectors j should yield the total intermediates supplied by each region-sector y_{ri}^{O} (obtained in step 5.1). Second, aggregating z_{rsij}^{OP} (z_{rsij}^{OO}) over origin regions should yield the sector-level intermediates consumed by each region-sector zt_{sij}^{OP} (zt_{sij}^{OO}). Finally, the RAS procedure results in the estimate of the domestic intermediate deliveries matrix. Note that we use the NIOP tables, which are balanced, as the benchmark. Therefore, as we ensure the sub-matrix in each block balanced and consistent with the NIOP tables, our estimations automatically ensure that we arrive at balanced IRIOP tables.

2.5. Validation analysis

The outcome of conducting all the above-described procedures is the set IRIOP tables for 2002, 2007, and 2012. They each include 17 sectors for 8 regions and distinguish in each region and sector between production of processing exports and ordinary production. The constraints we impose during the construction ensure that the IRIOP tables are balanced, are perfectly align with NIOP tables, and are maximally consistent with the information in the official statistics and the IRIO tables.

In this subsection, we compare the variables in the IRIOP tables with the existing IO tables to test the reliability of the IRIOP tables. As previously mentioned, when aggregating the IRIOP tables across regions, they exactly arrive at the NIOP tables. Therefore, the variables in the IRIOP tables are consistent with those in Chinese NIOP tables as well as the official national IO tables (at the national level). We therefore only compare the variables in the IRIOP tables with those in the SIC-NBS IRIO tables at the regional level. To this end, we aggregate the variables in the IRIOP tables across different production types, which yields variables with the same dimension as those included in the IRIO tables. After that, we calculate the correlation coefficients between value added, output, value-added coefficients, exports, domestic intermediate input and final use in the IRIOP tables and IRIO tables. Table 2 shows the dimensions of the variables as well as the correlation coefficients (multiplied by 100). The correlations are all very high, which suggests that the data in the IRIO

Table 2. Validation Analysis of IRIOP tables.

Variable	Description	2002	2007	2012
Value added levels	Vectors with 136 elements	99.9	100.0	99.2
Gross output levels	Vectors with 136 elements	99.6	99.5	98.8
Value added coefficients	Vectors with 136 elements	95.6	97.6	96.7
Domestic intermediate trade flows across regions and sectors	Matrices with 136∗136 elements	96.8	98.4	98.1
Exports levels	Vectors with 136 elements	93.4	91.0	96.7
Final demand	Matrices with 136∗8 elements	99.7	98.5	98.1

Note: All variables are at the region-sector level. The right three columns provide the correlation coefficients between the variables in the IRIOP tables and the SIC-NBS IRIO tables, multiplied by 100.

data have been little affected by the reconciliation procedures used to arrive at the IRIOP tables. This is reassuring. The table shows that the exports levels in the IRIOs have been affected more than have domestic variables. This is not surprising, given our purposeful choice to attach more importance to exports data in the NIOP and in customs data.

3. Separating processing exports from ordinary production matters, also for regional analysis!

Chen et al. (2012), Dietzenbacher et al. (2012), Koopman et al. (2012) and Pei et al. (2012) show that disregarding processing trade leads to serious estimation biases of domestic value added (DVA) in Chinese exports at the national level. Similar biases are undoubtedly inherent to similar computations based on conventional IRIO. We, thus, now compare them to the IRIOP tables introduced in this paper. We compare these two sets of tables using a recently developed indicator, the Supply Chain Fragmentation (SCF) index (Timmer et al., 2021). When computed on the basis of interregional IO tables for a specific country, the SCF index gives the value of all shipments that cross borders (interregional trade in both goods and services) in all stages of production involved in producing one unit of final demand. We present the derivation (based on Timmer et al., 2021) in Appendix D of the Online Supplementary Information, and just present the mathematical expression in matrix notation here:

$$SCF_{\tilde{F}} = \frac{\mathbf{u}_{CN}^{'}(\mathbf{T} \circ \{\mathbf{A}\,[\overline{(\mathbf{I}-\mathbf{A})^{-1}(\tilde{\mathbf{F}}\mathbf{u}_{CM})}]\})\mathbf{u}_{CN}}{\mathbf{u}_{CN}^{'}\tilde{\mathbf{F}}\mathbf{u}_{CM}} \tag{1}$$

Equation 1 gives the SCF index for the part of final demand considered, the matrix $\tilde{\boldsymbol{F}}$. If applied to the interregional IO tables that are central to this paper, the number of regions is denoted by C, the number of sectors by N and the number of final demand categories by M, this matrix has dimensions $CN \times CM$. It contains the values of the final demand block in which we are interested, and zeros elsewhere. By selecting the non-zero elements of $\tilde{\boldsymbol{F}}$ appropriately, it can relate to, e.g. final demand for products from South Coast, or to exports from all Chinese regions simultaneously, or to final demand by users in East Coast. $\mathbf{u}$ is a column vector containing ones (the subscript indices give the number of elements). The $N \times N$ blocks along the main diagonal of the $CN \times CN$ matrix $\mathbf{T}$ contain zeros, while all other elements of $\mathbf{T}$ are equal to one. The symbol $\circ$ stands for elementwise (Hadamard) multiplication. The upper bar indicates a diagonal matrix.

The numerator of Equation 1 gives the value of all interregional trade (wherever in China) due to the bundle of final demand considered. The denominator gives the value of this final output. Consequently, the SCF index can be viewed as a weighted average of SCFs for the output of single sectors in single regions. The index has policy relevance, for example, because interregional trade requires transport infrastructure. If, for example, exports from East Coast grow and if a given amount of these exports generate increasing interregional trade between Northwest and Central Regions in upstream stages of production (due to relocation of stages of production), investment in roads and railways linking these regions might become necessary.

As discussed in the introduction, the input structures of processing exports production and ordinary production within a sector tend to be different, both in terms of the quantities of inputs per unit of output and in terms of where inputs are sourced. These input structures are represented by the matrix $\mathbf{A}$ in Equation 1, of which the typical element a_{rsij} gives the value of inputs of sector i in region r required per unit of output of sector j in region s. The heterogeneity between the two types of production is hidden when conventional IRIO tables are used, while it is explicit in our IRIOP tables. We are interested in getting some insights into the magnitude of the differences between SCF indices based on the two types of tables.

To this end, we compare the results based on the IRIOP tables with the results based on IRIO tables (obtained by aggregating the IRIOP tables over processing exports and ordinary production activities, on a sector-by-sector basis). We make the results for the two tables comparable (the IRIOP tables have twice as many rows and columns as the IRIO tables) by specifying the matrix $\tilde{\mathbf{F}}$ in Equation 1 appropriately: if we are interested in the SCF for a specific subset of final output, we keep the final output in the IRIOP of both processing exports and ordinary production in that subset unchanged, and set all other elements to zero.[22]

Table 3 presents a first set of results. The SCF indices it contains relate to the final output of each of the eight regions, aggregated over the final use categories. The upper panel considers final output of all sectors, the lower panel only considers final output of manufacturing sectors.

A first finding is that the SCFs generally increased in the period 2002–2012, implying that the domestic parts of value chains fragmented interregionally. In East Coast, for example, the IRIOP-based results in the upper panel indicate that in 2012, 1,000 RMB of demand for its final products generated 300 RMB of interregional trade anywhere in China, while the corresponding value for 2002 was only 160 RMB. East Coast is an exception, however, to the general finding that the SCF indices grew mostly in the early subperiod (2002–2007) and remained much more stable afterwards.

Another finding is that the SCF indices vary considerably over regions. The upper panel shows that supply chains for final output in Northwest were structured in such a way that they generated more than twice as much interregional trade as final output of North Coast,

[22] The matrix $\mathbf{T}$ in Equation 1 has twice as many rows and columns if SCFs are computed on the basis of the IRIOP tables rather than on the basis of the IRIO tables. The dimensions of the summation vectors $\mathbf{u}$ vary accordingly.

Table 3. Comparison of SCF indices, for the final output of the eight Chinese regions.

	IRIOP 2002	IRIOP 2007	IRIOP 2012	IRIO 2002	IRIO 2007	IRIO 2012	DIFF 2002	DIFF 2007	DIFF 2012
ALL SECTORS									
NE	0.16	0.27	0.25	0.15	0.27	0.25	0.00	0.00	0.00
NM	0.27	0.41	0.29	0.27	0.41	0.29	0.00	0.00	0.00
NC	0.17	0.32	0.16	0.17	0.32	0.16	0.00	0.00	0.00
EC	0.16	0.23	0.30	0.16	0.23	0.30	0.00	0.00	0.00
SC	0.17	0.25	0.21	0.18	0.26	0.21	0.01	0.01	0.01
CR	0.21	0.29	0.30	0.20	0.29	0.30	0.00	0.00	0.00
NW	0.32	0.40	0.41	0.31	0.39	0.40	−0.01	−0.01	0.00
SW	0.15	0.25	0.23	0.15	0.24	0.22	0.00	−0.01	0.00
ONLY MANUFACTURING									
NE	0.19	0.31	0.29	0.19	0.32	0.29	0.00	0.01	0.00
NM	0.23	0.39	0.25	0.26	0.41	0.25	0.02	0.02	0.00
NC	0.19	0.37	0.20	0.20	0.37	0.20	0.01	0.01	0.01
EC	0.16	0.22	0.32	0.17	0.23	0.33	0.01	0.01	0.01
SC	0.14	0.28	0.22	0.17	0.31	0.24	0.03	0.03	0.02
CR	0.26	0.35	0.37	0.25	0.35	0.38	0.00	0.00	0.00
NW	0.34	0.44	0.44	0.34	0.44	0.44	0.00	−0.01	0.00
SW	0.16	0.25	0.25	0.16	0.25	0.25	0.00	0.00	0.00

Notes: NE: Northeast; NM: North Municipality; NC: North Coast; EC: East Coast; SC: South Coast; CR: Central Regions; NW: Northwest; SW: Southwest. The values in columns labeled DIFF have been obtained by subtracting the SCF indices based on the IRIOP tables from the SCF indices based on the conventional IRIO tables for the same year.

in 2012. The main cause of this difference is related to what Northwestern final producers in ordinary production activities require in terms of intermediate inputs from other regions.

One might expect that the SCF indices for all final output as presented in the upper panel of Table 3 are affected by final demand for services, which tend to be produced much more locally than goods (see Timmer et al., 2021, for results in an international context). To study this, the lower panel of Table 3 documents the SCF indices if only final output of manufacturing sectors is considered (see Appendix C for the sectors included). Focusing on results for 2012, we indeed observe that the SCF indices for manufacturing output tend to be somewhat higher (North Municipality is the exception). The differences are modest, however. This appears to be largely due to the fact that the shares of processing exports production in total production tend to be much larger in manufacturing sectors than in services. As discussed before, processing exports production relies heavily on imported inputs, rather than on inputs produced in China. Hence, the final output of processing exports tends to have lower SCF indices than the final output of ordinary production.

This brings us to the results related to the bias when using conventional IRIOs. The results in the rightmost part show these biases. A positive value indicates that the SCF index is overestimated if the conventional IRIO is used, while a negative value points at the opposite. With the exception of the results for manufacturing sectors in the South Coast, the biases are very modest. The result for South Coast gives a hint at the cause of the generally small biases. Looking at final output as a whole might hide heterogeneity. Exports rely partly on processing exports production with low SCF indices. Domestic final demand, however, is exclusively on ordinary production with higher SCF indices. If computed on the basis of a conventional IRIO, the elements of the **A**-matrix with intermediate input coefficients in Equation 1 are weighted averages of the input coefficients for both types of production. The biases in SCF indices for exports can thus be expected to be positive, while those for domestically sold final output are most likely negative. In South Coast, a relatively

Table 4. Comparison of SCF indices, by the final output of the eight Chinese regions, by final use category.

	IRIOP 2002	IRIOP 2007	IRIOP 2012	IRIO 2002	IRIO 2007	IRIO 2012	DIFF 2002	DIFF 2007	DIFF 2012
ONLY MANUFACTURING, consumption									
NE	0.21	0.29	0.20	0.19	0.27	0.20	−0.02	−0.02	−0.01
NM	0.32	0.48	0.30	0.24	0.44	0.29	−0.08	−0.04	0.00
NC	0.16	0.36	0.17	0.14	0.34	0.16	−0.02	−0.02	−0.01
EC	0.18	0.27	0.33	0.16	0.24	0.30	−0.02	−0.04	−0.03
SC	0.27	0.51	0.31	0.19	0.40	0.25	−0.08	−0.11	−0.06
CR	0.25	0.31	0.33	0.24	0.30	0.32	−0.01	−0.01	−0.01
NW	0.36	0.38	0.37	0.35	0.38	0.36	−0.01	−0.01	−0.01
SW	0.15	0.17	0.18	0.14	0.16	0.18	−0.01	−0.01	−0.01
ONLY MANUFACTURING, Capital goods									
NE	0.25	0.34	0.31	0.19	0.27	0.20	−0.06	−0.07	−0.12
NM	0.33	0.50	0.34	0.24	0.44	0.29	−0.09	−0.06	−0.05
NC	0.29	0.45	0.25	0.14	0.34	0.16	−0.14	−0.10	−0.08
EC	0.19	0.31	0.40	0.16	0.24	0.30	−0.04	−0.07	−0.11
SC	0.31	0.42	0.41	0.19	0.40	0.25	−0.12	−0.02	−0.17
CR	0.32	0.46	0.48	0.24	0.30	0.32	−0.08	−0.16	−0.16
NW	0.47	0.61	0.63	0.35	0.38	0.36	−0.12	−0.23	−0.27
SW	0.22	0.38	0.36	0.14	0.16	0.18	−0.08	−0.22	−0.18
ONLY MANUFACTURING, Exports									
NE	0.13	0.27	0.30	0.19	0.32	0.32	0.06	0.05	0.02
NM	0.17	0.35	0.20	0.25	0.39	0.21	0.08	0.05	0.01
NC	0.13	0.30	0.16	0.18	0.37	0.20	0.05	0.07	0.04
EC	0.13	0.19	0.28	0.17	0.22	0.33	0.04	0.03	0.05
SC	0.10	0.24	0.16	0.16	0.31	0.23	0.06	0.08	0.07
CR	0.24	0.32	0.29	0.26	0.37	0.39	0.02	0.05	0.10
NW	0.32	0.41	0.43	0.34	0.41	0.47	0.02	0.00	0.04
SW	0.17	0.24	0.22	0.19	0.27	0.27	0.02	0.03	0.05

Notes: NE: Northeast; NM: North Municipality; NC: North Coast; EC: East Coast; SC: South Coast; CR: Central Regions; NW: Northwest; SW: Southwest. The values in columns labeled DIFF have been obtained by subtracting the SCF indices based on the IRIOP tables from the SCF indices based on the conventional IRIO tables for the same year.

large share of final output is produced to meet export demand, so the positive bias reported for that region gives some support for this explanation.

The results in Table 4 confirm the explanation more systematically. In this table, we focus on results by type of final use, rather than on all final demand aggregated over uses. The results are presented for final output of manufacturing sectors. For exports, we find sizable positive biases, especially for the coastal region and in 2012 also for the Central Regions, which is a part of China that the government targeted for the development of processing exports activities in later stages of China's rise as a manufacturing giant. Not taking the differences between the two types of production into account clearly leads to overestimation of the amount of interregional trade due to further export rises for the Central Regions.

As expected, we find the opposite for the two final uses of manufactured products in China itself. Especially for capital goods, the negative biases in the SCF indices when computed on the basis of conventional IRIOs are often very substantial. For household consumption, we also find negative biases but these tend to be smaller. The differences in sizes of the biases between consumption goods and capital goods are a consequence of the differences in the composition of the two product bundles.

The results we present in this empirical illustration shows that the biases that can be avoided by using the IRIOP tables instead of conventional IRIO tables can either be small

(Table 3) or large (Table 4). For some research questions, using IRIOs does not cause problems, but for other questions, the results can be quite inaccurate. Hence, it is up to users to figure out whether the distinction between the two types of production in the IRIOP tables is relevant for their specific purposes or not.

4. Summary and conclusion

Processing trade is important to China's economy, and it is spatially unevenly distributed. Any good analysis requires adequate data at a minimum. We, therefore, develop a new interregional input–output table for China that explicitly distinguishes the production of processing exports from ordinary production. The new tables – for 2002, 2007, and 2012 – contain 17 sectors, including eight regions. After describing the available data, we detail how the information from the different data sources can be harmonized and reconciled. Then we give a step-by-step approach to the construction of the tables. The approach is not only applicable to China. It may also be adopted for countries like Mexico that also engage in a lot of processing trade activity that is unevenly distributed across regions.

We expected that failing to separate processing exports would yield misleading conclusions for regions as well as for a nation. So, we use the new table to investigate whether separating the production of processing exports from ordinary production matters when computing interregional supply chain fragmentation indices. To this end, we compare the empirical results derived from the new tables with the results derived from an aggregated version of them. The aggregate combines the two types of production and, thus, results from it mimics results from ordinary interregional input–output tables that do not disentangle processing trade. We found that inherent biases vary in size and depend on the question at hand.

Besides yielding more accurate empirical results, IRIOP tables have other advantages. They can enable answers to questions that cannot be secured from ordinary IRIO tables. An example of such a question is 'How is the value added generated by processing exports distributed over regions and how much does processing trade contribute to regional growth, regional inequality, as well as the regional environment changes?' By making the IRIOP tables publicly available, we believe these (and related) questions can be addressed via improved data. One challenge to constructing a Chinese IRIO table is the absence of official interregional transactions data. We estimate regional transaction flows by using existing IO tables in China. A better method might estimate them using firm-level survey data, for example, the value-added tax invoice data that record transactions among firms. We will leave this for future research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Beijing Social Science Fund: [Grant Number 21JJA004].

References

Aichele, R., & Heiland, I. (2018). Where is the value added? Trade liberalization and production networks. *Journal of International Economics*, *115*, 130–144. https://doi.org/10.1016/j.jinteco.2018.09.002

Amiti, M., & Freund, C. (2010). The anatomy of China's export growth. In R. C. Feenstra, & S. Wei (Eds.), *China's growing role in World trade* (pp. 35–56). University of Chicago Press.

Bacharach, M. (1970). *Biproportional matrices and input-output change*. Cambridge University Press.

Chen, Q., Zhu, K., Liu, P., Chen, X., Tian, K., Yang, L., & Yang, C. (2019). Distinguishing China's processing trade in the World input-output table and quantifying its effects. *Economic Systems Research*, *31*(3), 361–381. https://doi.org/10.1080/09535314.2018.1534225

Chen, X., Cheng, L. K., Fung, K. C., & Lau, L. J. (2001). *The estimation of domestic value-added and employment induced by exports: an application to Chinese exports to the United States*. Working Paper No. 94305, Department of Economics, Stanford University, California. http://www.stanford.edu/ ljlau/Presentations/Presentations/010618.pdf.

Chen, X., Cheng, L. K., Fung, K. C., Lau, L. J., Sung, Y. W., Zhu, K., Yang, C., Pei, J., & Duan, Y. (2012). Domestic value added and employment generated by Chinese exports: A quantitative estimation. *China Economic Review*, *23*(4), 850–864. https://doi.org/10.1016/j.chieco.2012.04.003

Dean, J. M., Fung, K. C., & Wang, Z. (2011). Measuring vertical specialization: The case of China. *Review of International Economics*, *19*(4), 609–625. https://doi.org/10.1111/j.1467-9396.2011.00969.x

Dietzenbacher, E., Los, B., Stehrer, R., Timmer, M., & de Vries, G. (2013). The construction of world input–output tables in the WIOD project. *Economic Systems Research*, *25*(1), 71–98. https://doi.org/10.1080/09535314.2012.761180

Dietzenbacher, E., Pei, J., & Yang, C. (2012). Trade, production fragmentation, and China's carbon dioxide emissions. *Journal of Environmental Economics and Management*, *64*(1), 88–101. https://doi.org/10.1016/j.jeem.2011.12.003

Duan, Y., Dietzenbacher, E., Los, B., & Yang, C. (2021). How much did China's emergence as "the world's factory" contribute to its national income? *China Economic Review*, *69*, 101658. https://doi.org/10.1016/j.chieco.2021.101658

Duan, Y., & Yan, B. (2021). Has processing trade made China's exports cleaner? A regional level analysis. *Energy Economics*, *96*, 105150. https://doi.org/10.1016/j.eneco.2021.105150

Duan, Y., Yang, C., Zhu, K., & Chen, X. (2012). Does the domestic value added induced by China's exports really belong to China? *China & World Economy*, *20*(5), 83–102. https://doi.org/10.1111/j.1749-124X.2012.01303.x

Fu, X. (2004). Limited linkages from growth engines and regional disparities in China. *Journal of Comparative Economics*, *32*(1), 148–164. https://doi.org/10.1016/j.jce.2003.10.004

Gilchrist, D. A., & St Louis, L. V. (1999). Completing input–output tables using partial information, with an application to Canadian data. *Economic Systems Research*, *11*(2), 185–194. https://doi.org/10.1080/09535319900000013

Holz, C. (2004). China's statistical system in transition: Challenges, data problems, and institutional innovations. *Review of Income and Wealth*, *50*(3), 381–409. https://doi.org/10.1111/j.0034-6586.2004.00131.x

Jiang, X., Duan, Y., & Green, C. (2017). Regional disparity in energy intensity of China and the role of industrial and export structure. *Resources, Conservation and Recycling*, *120*, 209–218. https://doi.org/10.1016/j.resconrec.2016.12.013

Johnson, R. C. (2014). Five facts about value-added exports and implications for macroeconomics and trade research. *Journal of Economic Perspectives*, *28*(2), 119–142. https://doi.org/10.1257/jep.28.2.119

Johnson, R. C. (2018). Measuring global value chains. *Annual Review of Economics*, *10*(1), 207–236. https://doi.org/10.1146/annurev-economics-080217-053600

Johnson, R. C., & Noguera, G. (2012). Accounting for intermediates: Production sharing and trade in value added. *Journal of International Economics*, *86*(2), 224–236. https://doi.org/10.1016/j.jinteco.2011.10.003

Koopman, R., Wang, Z., & Wei, S. (2012). Estimating domestic content in exports when processing trade is pervasive. *Journal of Development Economics*, *99*(1), 178–189. https://doi.org/10.1016/j.jdeveco.2011.12.004

Lenzen, M., Gallego, B., & Wood, R. (2009). Matrix balancing under conflicting information. *Economic Systems Research*, *21*(1), 23–44. https://doi.org/10.1080/09535310802688661

Lenzen, M., Moran, D., Geschke, A., & Kanemoto, K. (2014). A non-sign-preserving RAS variant. *Economic Systems Research*, *26*(2), 197–208. https://doi.org/10.1080/09535314.2014.897933

Los, B., Timmer, M., & de Vries, G. (2015). How important are exports for job growth in China? A demand side analysis. *Journal of Comparative Economics*, *43*(1), 19–32. https://doi.org/10.1016/j.jce.2014.11.007

Ma, A. C., & Van Assche, A. (2016). Spatial linkages and export processing location in China. *The World Economy*, *39*(3), 316–338. https://doi.org/10.1111/twec.12284

Ma, H., Wang, Z., & Zhu, K. (2015). Domestic content in China's exports and its distribution by firm ownership. *Journal of Comparative Economics*, *43*(1), 3–18. https://doi.org/10.1016/j.jce.2014.11.006

Meng, B., Fang, Y., Guo, J., & Zhang, Y. (2017). Measuring China's domestic production networks through trade in value-added perspectives. *Economic Systems Research*, *29*(1), 48–65. https://doi.org/10.1080/09535314.2017.1282435

Mi, Z., Meng, J., Zheng, H., Shan, Y., Wei, Y., & Guan, D. (2018). Data descriptor: A multi-regional input-output table mapping China's economic outputs and interdependencies in 2012. *Scientific Data*, *5*(1), 1–12. https://doi.org/10.1038/s41597-018-0002-5

National Bureau of Statistics (NBS). (2017). *China statistics yearbook 2016*. China Statistics Press.

Pei, J., Oosterhaven, J., & Dietzenbacher, E. (2012). How much do exports contribute to China's income growth? *Economic Systems Research*, *24*(3), 275–297. https://doi.org/10.1080/09535314.2012.660746

Pei, J., Oosterhaven, J., & Dietzenbacher, E. (2017). Foreign exports, net interregional spillovers and Chinese regional supply chains. *Papers in Regional Science*, *96*(2), 281–298. https://doi.org/10.1111/pirs.12188

Peters, G. P., Andrew, R., & Lennox, J. (2011). Constructing an environmentally-extended multi-regional input–output table using the GTAP database. *Economic Systems Research*, *23*(2), 131–152. https://doi.org/10.1080/09535314.2011.563234

Stone, R., & Brown, A. (1962). *A computable model of economic growth*. Chapman and Hall.

Su, B., Ang, B. W., & Low, M. (2013). Input–output analysis of CO_2 emissions embodied in trade and the driving forces: Processing and normal exports. *Ecological Economics*, *88*, 119–125. https://doi.org/10.1016/j.ecolecon.2013.01.017

Su, B., Huang, H. C., Ang, B. W., & Zhou, P. (2010). Input–output analysis of CO2 emissions embodied in trade: The effects of sector aggregation. *Energy Policy*, *32*, 166–175.

Timmer, M. P., Los, B., Stehrer, R., & de Vries, G. J. (2013). Fragmentation, incomes and jobs: An analysis of European competitiveness. *Economic Policy*, *28*(76), 613–661. https://doi.org/10.1111/1468-0327.12018

Timmer, M. P., Los, B., Stehrer, R., & de Vries, G. J. (2021). Supply chain fragmentation and the global trade elasticity: A new accounting framework. *IMF Economic Review*, *69*, 656–680. https://doi.org/10.1057/s41308-021-00134-8

Tukker, A., & Dietzenbacher, E. (2013). Global multiregional input–output frameworks: An introduction and outlook. *Economic Systems Research*, *25*(1), 1–19. https://doi.org/10.1080/09535314.2012.761179

Upward, R., Wang, Z., & Zheng, J. (2013). Weighing China's export basket: The domestic content and technology intensity of Chinese exports. *Journal of Comparative Economics*, *41*(2), 527–543. https://doi.org/10.1016/j.jce.2012.07.004

Yang, C., Dietzenbacher, E., Pei, J., Chen, X., Zhu, K., & Tang, Z. (2015). Processing trade biases the measurement of vertical specialization in China. *Economic Systems Research*, *27*(1), 60–76. https://doi.org/10.1080/09535314.2014.955463

Zhang, Y., & Qi, S. (2012). *China's multi-regional input-output tables for years 2002 and 2007*. China Statistics Press. in Chinese.

Zhang, Z., Duan, Y., & Zhang, W. (2019). Economic gains and environmental costs from China's exports: Regional inequality and trade heterogeneity. *Ecological Economics*, *164*, 106340. https://doi.org/10.1016/j.ecolecon.2019.05.020

Zheng, H., Ou, J., Meng, J., Mi, Z., Song, M., Shan, Y., Ou, J., & Guan, D. (2019). Linking city-level input-output table to urban energy footprint. Construction framework and application. *Journal of Industrial Ecology*, *23*, 781–795.

Zheng, H., Zhang, Z., Wei, W., Song, M., Dietzenbacher, E., Wang, X., Meng, J., Shan, Y., Ou, J., & Guan, D. (2020). Regional determinants of China's consumption-based emissions in the economic transition. *Environmental Research Letters*, *15*(7), 074001. https://doi.org/10.1088/1748-9326/ab794f