

# Estimating domestic content in China's exports: Accounting for a dual-trade regime<sup>☆</sup>

Quanrun Chen<sup>a</sup>, Xikang Chen<sup>b</sup>, Jiansuo Pei<sup>c,\*</sup>, Cuihong Yang<sup>b</sup>, Kunfu Zhu<sup>d</sup>

<sup>a</sup> School of Statistics, University of International Business and Economics, China

<sup>b</sup> Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China

<sup>c</sup> School of International Trade and Economics, University of International Business and Economics, China

<sup>d</sup> Research Institute for Global Value Chains, University of International Business and Economics, China

## ARTICLE INFO

### JEL classification:

C67

C82

F14

F62

### Keywords:

Input-output tables

International fragmentation

Processing trade

Trade in value added

China

## ABSTRACT

This paper identifies the heterogeneity issue as a key challenge that is central to but not fully addressed when measuring global value chains. To resolve this issue, we propose an extended input-output model that is consistent with the theoretical framework of heterogeneous firms. Empirically, we use China as a prominent example of a country that is engaged in both normal trade and the processing trade under a dual-trade regime, and we synthesize methods for constructing China's extended input-output dataset for the period 1997 to 2015. Our results show that when alternative generic datasets are used, this is likely to result in overestimating the domestic content in China's exports by as much as 44%, compared to a model that uses an extended database that incorporates production heterogeneity, as does the one in this study. This paper's proposed methodology and rich dataset may be useful to a wider range of empirical applications.

## 1. Introduction

Conventional trade statistics face ever-growing challenges in an era of international-production fragmentation. Researchers have proposed new measures that could be used to account for the pattern and growth of this phenomenon. These include measures to account for vertical specialization (Hummels et al., 2001; Johnson and Noguera, 2012) and those that account for domestic content in exports (Chen et al., 2012; Koopman et al., 2012, 2014). The fundamental and central question for such research is whether to measure the value added that occurs along the global value chain (GVC) (see Timmer et al., 2014a; Los et al., 2015).<sup>1</sup> To that end, two elements are required: First, there is need for a method that is not only capable of tackling the double-counting problem that arises due to increasing trade in intermediates but also of being parsimonious enough for a wide audience to be able to understand it; and second, the data must be collected and maintained in a manner that allows the above-mentioned method to be applied.

In the 1930s, Leontief (1936) proposed a decomposition method

under an input-output (IO) framework that could be used to disentangle exports into domestic and imported content (i.e., vertical specialization; see, e.g., Hummels et al., 2001). In other words, such a method has existed for decades and has even been proliferating recently, largely extending the scope of related studies and enhancing our understanding of the method through a different lens. However, the limited frequency and timeliness of publishing IO databases, due to limited finances and human resources, has been slowing down the progress of new findings in the field (Daaniyall et al., 2017). Moreover, despite continuing efforts to construct multi-region IO databases, arguably, the quality of the data remains a bottleneck. This raises the question of the reliability of estimates that use different databases and whether there exist alternative databases that would permit us to more accurately estimate domestic content.

According to Los et al. (2016) and Johnson (2018), if our focus is on the domestic content in exports, then national IO tables are sufficient. Following this line of argument, our purpose in this study is twofold: first, to provide a better single-country database that accounts for China's

<sup>☆</sup> Equal contribution, alphabetical order of family names of the co-authors is applied.

\* Corresponding author. 10 Huixin Dongjie, Beijing, 100029, China.

E-mail address: [jspei@amss.ac.cn](mailto:jspei@amss.ac.cn) (J. Pei).

<sup>1</sup> World Development Report, 2020, the flagship report by World Bank recognizes the importance of global value chains with the theme "Trading for Development in the Age of Global Value Chains", examines development issues through the lens of GVCs and trade.

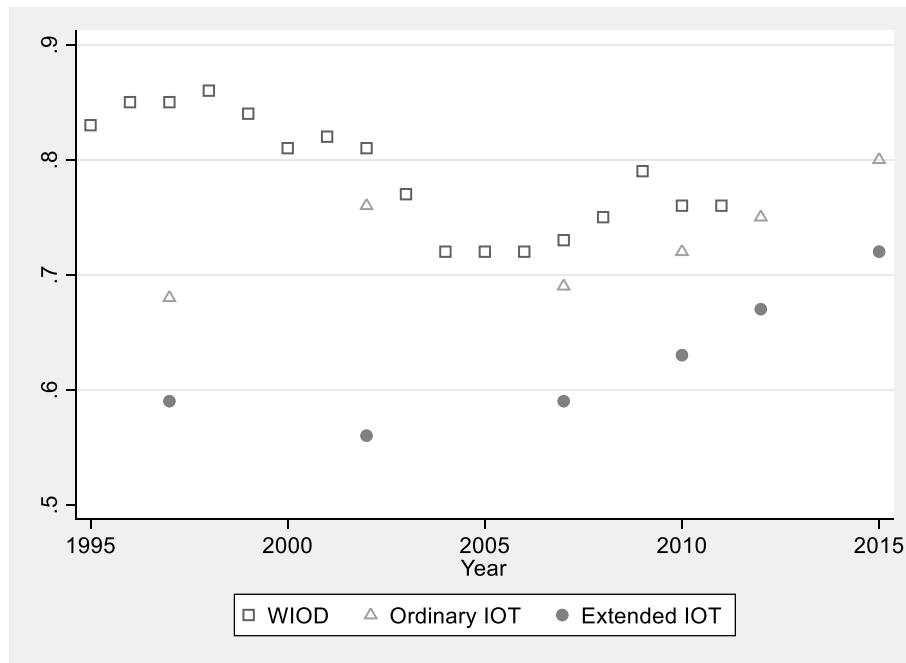


Fig. 1. Estimates for China's domestic content in exports, using alternative datasets, 1997–2015. Source: Authors' estimation using alternative datasets.

dual-trade regime (i.e., the coexistence of normal trade and the processing trade), thereby, complementing existing work (see, e.g., Zheng et al., 2018) and offering guidance on what can be accomplished by using such an extended database. Fig. 1, below, shows estimates of the domestic content in Chinese exports, using alternative datasets, where the World Input-Output Database (WIOD) and our preferred extended IO tables are presented. Clearly, there would be substantial overestimation if either ordinary IO tables or the WIOD national database, and not the extended IO tables, were employed. For example, for 2002, our preferred estimation method of using the extended IO table would report a domestic content per unit of exports of 0.56, while using the WIOD database would result in a value of 0.81, and using the ordinary IO table would result in a value of 0.76.

As was previously mentioned, estimating of the value added in exports is the core area of concern and the primary focus of much GVC-related research. One of the key challenges, here, is to improve the available information on the use of imports in firm-level production and also the trade data that shows this production's micro foundation; that is, this data should link manufacturers and traders by assigning imports intermediate goods to their proper end-users. This paper directly contributes to one vibrant strand of research and is related to several other streams of studies. First, it complements discussions on measuring global value chains (see Johnson, 2018, for an overview), in particular, those along the lines of combining micro data to improve the estimation of the import matrix (see also Yao et al., 2015). Essentially, the literature that discusses the measurement problem boils down to the need for a more transparent and accurate means of estimating import matrices (see Ahmad et al., 2011).<sup>2</sup> Better data quality would permit better

<sup>2</sup> As will be clear in section 3, three elements comprise the estimation of domestic content in exports; these are value added coefficients, exports, and the Leontief inverse. In fact, in most cases the value-added coefficients and exports are directly available from statistics, whereas the Leontief inverse must be estimated and, by construction, the only uncertainty comes from the domestic input coefficients because we need to split the input coefficients into their domestic ones and their imported counterparts (see Feenstra and Jensen, 2012, for an evaluation).

measurement results. Our study offers practical suggestions for gathering accurate IO data for other economies.

Second, this paper offers suggestions on how to use alternative datasets to evaluate the effects of international production fragmentation on domestic economies; it also provides refined methods that can be used to decompose gross exports in value-added terms. Recent literature has advanced Feenstra (1998) method, which uses an IO framework to estimate vertical specialization and is considered a better indicator for outsourcing. As was just mentioned, if the import matrix were subject to measurement error, then studies that build on this independent variable and/or its variants would not be warranted (see related work by Arnold et al., 2014). Third, our work also contributes to studies on quantifying the determinants that contribute to deeper globalization. In contrast to the second line of research, this measurement is now the dependent variable. As indicated in Kee and Tang (2016), a more accurate estimation of the domestic content in exports is preferably obtained by combining micro data with IO tables. We also extend the period under review from 2000 to 2007 to 1997 to 2015.

We also contribute to the work on computable general equilibrium (CGE) models as we provide much richer parameters. Conventional CGE models for Chinese trade policy analyses do not differentiate China's processing exports from the rest of the country's economy. Here, examples include the model developed by China's Development Research Center (the DRC model), which focuses on Chinese regions and uses the standard Global Trade Analysis Project (GTAP) model of Hertel and Tsigas (1997). Using a CGE model for China, economists have attempted to separate the country's processing trade from its normal trade (Wang, 2003; Ianchovichina, 2004; Ianchovichina and Martin, 2004). Recently, with the availability of Chinese trade data on the processing trade, Koopman et al. (2013) was able to split the processing trade sector from the rest of the Chinese economy and treat it as a separate economy in a GTAP-turned-GVC model. This split, however, was largely based on assumptions on key IO coefficients and did not further differentiate the production of normal exports from the production of goods for domestic use. Our proposed framework is an improvement of the core database along this line.

In section 2, we explain the background for a new development of China's alternative IO tables; i.e., an extended dataset that incorporates

China's dual-trade regime.<sup>3</sup> To achieve that goal, one needs to develop new methods and collect additional (and crucial) micro data. These issues are discussed in section 3. Here, after briefly introducing the official IO data sources, two different yet interrelated perspectives are discussed in greater detail: i) data that is included and not included in the standard national IO table; and ii) the advantage of using the alternative IO table over the standard one. In section 4, we explore an application of these datasets, using our extended dataset to measure the domestic content in exports. We also present several aspects that can be further investigated by using this dataset, and section 5 concludes.

## 2. Background

It is widely known that China began its opening-up in 1978, whereas perhaps its real opening-up took off several years later; say, during the early 1990s, when Deng Xiaoping gave his famous “southern tour” speech.<sup>4</sup> Since then, China's macro-economic indicators, such as its gross domestic product (GDP), trade volume, and foreign direct investment have caught worldwide attention. After China's entry into the World Trade Organization (WTO), its trade growth soared. According to China Custom's trade statistics, the country recorded a \$323 billion trade surplus with the US in 2018, whereas in 1993 the value of this trade surplus was only \$6 billion.<sup>5</sup> This leads to the following questions.

To what extent did Chinese value added and employment benefit from this unprecedented export growth? Did the quality of China's exports improve and what were the consequences of its exporters' pricing behaviors (see, Mallick and Marques, 2016, 2017)?<sup>6</sup> To what extent have countries whose markets are China's largest targets for its exports, such as the US, lost in terms of employment?<sup>7</sup> And as China has become the world's largest carbon dioxide (CO<sub>2</sub>) emitting country, to what extent should its trading partners account for the increasing greenhouse gas emissions that are associated with importing goods that are “Made in China”? These are all relevant questions.

China has grown in importance in the world economy, having moved from being a negligible player in the early 1990s to becoming the world's largest exporter and second largest economy. The same is true in regard to the demand for China's data, in particular its IO data.<sup>8</sup> Several studies that have appeared in renowned international journals have used China's IO data; unfortunately, few of these studies acknowledge the strengths and weaknesses of this data.<sup>9</sup>

Moreover, fully exploring the IO work, including the compilation

<sup>3</sup> The dataset explicitly incorporates the processing trade in addition to normal trade; see early contributions in Chen et al. (2009), Dean et al. (2011), Koopman et al. (2012).

<sup>4</sup> For example, Naughton (2007) called it “the second phase of reform”; see also contributions by Lardy (2002).

<sup>5</sup> Such differences were larger, according to the US statistics these values were \$419 billion in 2018 and \$23 billion in 1993, respectively.

<sup>6</sup> They explicitly tackle the problem of export quality and exporters' pricing behaviors using trade data for China and India.

<sup>7</sup> This has led to debates regarding the potential threats China's exports pose to the US, collectively called the “China trade shock” (see e.g., Autor et al., 2013).

<sup>8</sup> We will not include aspects of the development of multi-country IO tables; one can refer to a recent special issue of the Global Multiregional Input-Output Frameworks (ESR, 2013).

<sup>9</sup> For instance, Hummels et al. (2001) used China's national IO table to estimate vertical specialization and found that China was average in this aspect. This was substantially revised by Yang et al. (2015) and Dean et al. (2011) via explicitly taking into account the processing trade, and similar refinements are made in the present study.

<sup>11</sup> The history of input-output compilations and applications in China up to the 1980s can be found in Chen (1989). More recently, Economic Systems Research, the official journal of the International Input-Output Association, published one special issue entitled “China's Growing Pains – Recent Input-Output Research in China on China” (ESR, 2008).

methods, the jargon and so forth, would be a daunting challenge for IO users, who, for example, only need the data to conduct empirical studies.<sup>11</sup> There was a need for a transition from what was said in the previous part of the paragraph.

With this concern in mind, this study aims to deal with two main questions from the point of view of data users. First, provided data users have obtained China's IO tables,<sup>12</sup> what are and are not included in these tables? Second, given the perceived limitations of China's standard national IO tables and any datasets that have been constructed based on them, do other datasets exist? And what are these alternative tables good for? Furthermore, given that the processing trade<sup>13</sup> takes up roughly one third up to half of China's total trade volume, are the standard IO tables still applicable? If not, what would be the alternative?

The processing trade has a rather different production structure in comparison with normal trade and other production. First, the processing trade is mainly conducted by foreign-invested enterprises and their production techniques can differ from those of other production firms, such as state-owned enterprises and private firms. Second, processing exports normally rely more on processing imports, which are duty-free. This implies that the national IO table shows an average production structure, which means when the processing trade is important, biased results would be inevitable if no distinction were made for the different trade regimes (processing and normal trade).

In light of these concerns, China's IO table should take each of the following two aspects into account. It must (i) distinguish between different trade regimes, and (ii) it should cover changes over time in order to evaluate past developments. In fact, Chen et al. (2009) developed the first non-competitive IO table that captures the processing trade for China for the year 1995. This table was set up with the aim of estimating the Sino-US trade surplus in terms of the value added.<sup>14</sup> Specifically, the production for the processing exports was differentiated from China's other production (either the production for domestic use or for normal exports); this resulted in a special framework that explicitly accounted for China's dual-trade regime.

This idea has been well received, and follow-up datasets were constructed and then applied to various policy questions (see, e.g., Yang et al., 2015; Chen et al., 2012; Dietzenbacher et al., 2012; Koopman et al., 2012; Dean et al., 2011). Moreover, the treatment of heterogeneous production technologies within an IO framework has been widely recognized and in keeping with the development of China's IO table. Despite institutional efforts, such as those of the OECD ICIO that distinguish the processing trade for China and global manufacturing for Mexico, national statistical agencies and research groups are embracing

<sup>12</sup> These are publicly available from the websites of the National Bureau of Statistics (NBS); see Table 1 for a sketch of the IO tables in China.

<sup>13</sup> Under China's dual-trade regime, there are mainly two types of trade: normal trade and the processing trade. The processing trade, as opposed to normal trade, uses a large share of the raw and auxiliary materials, parts and components, accessories, and packaging materials that are imported from abroad, duty free. The finished products that Chinese enterprises produce from these imported goods are re-exported. According to the special rule, the processing trade can only be used to produce processing exports. Roughly speaking, there are two types of processing trade: processing with purchased imports and processing and assembling imports. Chen et al. (2012) has explicitly tackled this problem (see also, Pei et al., 2011; Feenstra and Wei, 2010).

<sup>14</sup> This viewpoint has been endorsed by the OECD/WTO, which launched the “Made in the World” initiative and proposed “trade in value added” as a better approach for measuring international trade (OECD/WTO, 2012). In addition, the OECD has constructed extended inter-country IO tables, while paying special attention to China's processing trade and Mexico's global manufacturing sectors. The key parameters that are used to estimate the special inter-country IO table for China were provided by our research team (for the year 2010).

the idea of accounting for heterogeneous production technologies within IO tables (see, e.g., Michel et al., 2018, for the Belgium case).<sup>15</sup>

### 3. Methodology and data

The theory of heterogeneous firms (see, e.g., Melitz, 2003) gave rise to the study of firms within industries and this has advanced workhorse neo-classical trade models. Exactly the same idea was held in the development of an extended IO framework. Equally important has been how the availability of micro data has facilitated research that uses more-detailed information for production and trade. Building on China’s national IO table, supplemented with micro data (e.g., manufacturing firm surveys, detailed customs data), the development of a new dataset that incorporates heterogeneous production technologies has gone through several major steps; these are detailed in Appendix A.<sup>16</sup>

#### 3.1. Understanding the basics of a national IO table

IO tables have generally been of the so-called competitive type; i.e., imports are treated the same way as domestically produced goods (assuming perfect substitution; see Fig. 2).

To explain the problem without separating out imports from domestically produced goods, first, we specify the IO framework that would correspond to the official Chinese IO table. In Fig. 2, we have  $\mathbf{x} = \mathbf{Z}\mathbf{i} + \mathbf{f} + \mathbf{e} - \mathbf{m}$  as accounting for the identities found along the rows of the IO table, where  $\mathbf{i}$  indicates a summation vector that consists of ones. This IO framework first assumes that the supply of output  $\mathbf{x}$  follows total net demand  $\mathbf{Z}\mathbf{i} + \mathbf{f} + \mathbf{e} - \mathbf{m}$ . Next, it assumes that the endogenous intermediate demand for worldwide products is determined by total output, i.e.  $\mathbf{Z}\mathbf{i} = \mathbf{A}\mathbf{x}$ , in which the matrix of the technical coefficients is calculated by  $\mathbf{A} = \mathbf{Z}\hat{\mathbf{x}}^{-1}$  (the hat indicates a diagonal matrix). Consequently, the solution for this IO framework is  $\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1}(\mathbf{f} + \mathbf{e} - \mathbf{m})$ . The equation  $\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}$  is the famous Leontief inverse, whose typical element  $l_{ij}$  gives the (extra) output in industry  $i$  necessary for one (extra) final demand for product  $j$ .

Analyzing the structural problems, such as the technical coefficients (Leontief’s “recipe”), poses no difficulties; but this would lead to biased estimations when investigating such issues as the effects of final demand shocks on total output. The reasoning here is intuitive: imported intermediates have their production linkages outside the system (no production roundabouts, say, in China); in the official IO table, however, no distinction is made between domestically produced inputs and imported inputs. Thus, the table is of little empirical use at this stage.<sup>17</sup>

To reckon with this problem, as a widely applied extension, the conventional IO table must be adapted. The so-called “non-competitive” type of table needs to be constructed; that is, it must be one that separates the intermediate and final imports from the domestically produced goods (see Fig. 3). Domestically produced and imported intermediate inputs, even those in the same sector classification, may not be perfect or even good substitutes for each other; however, “competitive” type IO tables lump domestically produced inputs and imported inputs together, and

<sup>15</sup> In addition to the OECD, under the APEC TiVA project, several economies endorsed the idea of developing IO tables that incorporate heterogeneous production technologies, such as for the US, Canada and ASEAN countries, despite China and Mexico. See <http://www.apectivagvc.org/> for details.

<sup>16</sup> For the standard procedure for compiling an IO table, please refer to the Department of Accounts, NBS (2009) for China’s 2007 benchmark table, and Meade (2010) for the US.

<sup>17</sup> This type of framework is not plausible for at least two reasons. First, it assumes that the endogenous intermediate inputs from the rest of the world were produced by Chinese industries. Second, it assumes that total imports are exogenous; i.e., they do not depend on the size of domestic intermediates and final demand, which may lead to inconsistency with the endogenous intermediate imports.

“non-competitive” types treat them differently (the so-called “Armington approach”). It is essential to make these distinctions and apply these treatments when studying the effects of exports on domestic factors (such as value added, employment, and emissions).

Given the limited information regarding which imports are to be classified as intermediates or final goods, the “proportionality assumption” is used to split the “competitive” table into a “non-competitive” one.<sup>18</sup> As the word suggests, we assume a common share of import use, irrespective of the intermediate or final use.

Define  $\mathbf{t}$  as the vector with foreign import coefficients that are calculated as  $t_i = m_i / (x_i - e_i + m_i)$ . Then, using the proportional method (see, e.g., Pei et al., 2011), the domestically produced intermediate inputs can be estimated as  $\mathbf{Z}^D\mathbf{i} = (\mathbf{I} - \hat{\mathbf{t}})\mathbf{A}\mathbf{x} = \hat{\mathbf{t}}^D\mathbf{A}\mathbf{x}$ , with  $\mathbf{I}$  as the identity matrix and  $\hat{\mathbf{t}}^D$  as a diagonal matrix with self-sufficiency ratios. The same method is used to estimate the domestically produced final demand, as  $\mathbf{f}^D = \hat{\mathbf{t}}^D\mathbf{f}$ . The accounting identity  $\mathbf{x} = \mathbf{Z}\mathbf{i} + \mathbf{f} + \mathbf{e} - \mathbf{m}$  can now be rewritten so as to cover only the deliveries of domestically produced goods. Thus, foreign imports  $\mathbf{m}$  are excluded in the rewritten accounting identity  $\mathbf{x} = \mathbf{Z}^D\mathbf{i} + \mathbf{f}^D + \mathbf{e}$ . After rearranging, we get  $\mathbf{x} = (\mathbf{I} - \hat{\mathbf{t}}^D\mathbf{A})^{-1}(\mathbf{f}^D + \mathbf{e})$ .

The strengths of the standard “non-competitive” IO table are obvious: it is transparent and easy to obtain. Importantly, it is in line with our economic intuition that only domestically produced products enter the circulation process; i.e., the Leontief inverse  $((\mathbf{I} - \hat{\mathbf{t}}^D\mathbf{A})^{-1})$ . In contrast to Fig. 2, the “non-competitive” framework has been widely used to analyze various issues, such as international trade (for example, vertical specialization, see Hummels et al., 2001), the energy issue (see Lin and Polenske, 1995), and the emissions embodied in foreign exports (Weber et al., 2008), and so forth.

One aspect that has not received much attention in Fig. 3, however, is the distinction between the different types of trade flows. China has become the world’s largest exporter and a major player in the fragmentation of international production. Statistics show that processing exports accounted for more than 50% of China’s total annual exports during the period 1996–2007 and no less than one-third in 2015. Therefore, when calculating the effects (value added generation, and/or emissions) due to China’s exports, it is important to make a distinction between exports from the processing sector and normal exports (see Fig. 4).

Analogous to the formulas in the previous sections, we have the augmented equation  $\tilde{\mathbf{x}} = \tilde{\mathbf{Z}}^D\mathbf{i} = \tilde{\mathbf{f}}^D + \tilde{\mathbf{e}}$ . Explicitly, following common practice (e.g., Yang et al., 2015), the solution to the extended formula is given by the following matrices:

$$\begin{bmatrix} \mathbf{x}^D \\ \mathbf{x}^P \\ \mathbf{x}^N \end{bmatrix} = \left[ \mathbf{I} - \begin{pmatrix} \mathbf{A}^{DD} & \mathbf{A}^{DP} & \mathbf{A}^{DN} \\ 0 & 0 & 0 \\ \mathbf{A}^{ND} & \mathbf{A}^{NP} & \mathbf{A}^{NN} \end{pmatrix} \right]^{-1} \begin{bmatrix} \mathbf{f}^{DD} + 0 \\ 0 + \mathbf{e}^P \\ \mathbf{f}^N + \mathbf{e}^N \end{bmatrix},$$

where,  $\tilde{\mathbf{L}} = \left[ \mathbf{1} - \begin{pmatrix} \mathbf{A}^{DD} & \mathbf{A}^{DP} & \mathbf{A}^{DN} \\ 0 & 0 & 0 \\ \mathbf{A}^{ND} & \mathbf{A}^{NP} & \mathbf{A}^{NN} \end{pmatrix} \right]^{-1} = \begin{bmatrix} \mathbf{L}^{DD} & \mathbf{L}^{DP} & \mathbf{L}^{DN} \\ 0 & \mathbf{I} & 0 \\ \mathbf{L}^{ND} & \mathbf{L}^{NP} & \mathbf{L}^{NN} \end{bmatrix}$  is the extended Leontief inverse, with partitioned matrices.

Note that in the IO framework final demand is usually treated as an

<sup>18</sup> Such an assumption is widely adopted in IO compilation practice, including for the US, the OECD and many other economies (see Pei et al., 2011); see also Feenstra and Jensen (2012) for an evaluation. Dean et al. (2011) propose combining the UN BEC method with the “proportionality assumption” to obtain a better estimation. This method is endorsed by many recent multi-regional IO projects; say, the WIOD project ([www.wiod.org](http://www.wiod.org)), see Timmer et al. (2014a). One problem regarding the UN BEC method is that about 14% of goods can be both intermediate and final goods. Further, even if the intermediates were to be separated from the rest, the assumption also needs to be made that they should be distributed into different sectors/industries, where the “proportionality assumption” is applied.



	Intermediate use	Final use			
		<i>DFD</i>	<i>EXP</i>	<i>IMP</i>	<i>TOT</i>
	<b>Z</b>	<b>f</b>	<b>e</b>	<b>-m</b>	<b>X</b>
<i>VA</i>	<b>v'</b>				
<i>TOT</i>	<b>x'</b>				

Notes: **Z** = (domestically produced and imported) intermediate deliveries matrix; *DFD* = domestic (worldwide) final demand; *EXP* = exports; *TOT* = gross industry outputs (and total inputs for each sector); *IMP* = imports; and *VA* = value added. China’s input-output table is expressed in monetary units (of 10,000 Yuan).

**Fig. 2.** The structure of China’s national input-output table, competitive type. Notes: **Z** = (domestically produced and imported) intermediate deliveries matrix; *DFD* = domestic (worldwide) final demand; *EXP* = exports; *TOT* = gross industry outputs (and total inputs for each sector); *IMP* = imports; and *VA* = value added. China’s input-output table is expressed in monetary units (of 10,000 Yuan). Source: Authors’ own illustration.

exogenous variable, while total output is endogenously determined.<sup>19</sup> Then, by pre-multiplying the extended value-added coefficients,<sup>20</sup> we can obtain the value added for each production class and sector (GDP can be obtained by summing up the total value added for these production classes and sectors).

3.2. The development of the extended IO table that captures firm heterogeneity

Fig. 3 is the most commonly used IO structure nowadays, although import matrices are estimated in most cases.<sup>21</sup> In fact, our procedure for constructing the extended IO framework is more like a “bottom-up” approach. That is, our starting point is Fig. 4 and by aggregating the corresponding matrices, Fig. 3 can be obtained.

Motivated by the hotly debated issue of the Sino-US trade surplus, Chen et al. (2009) proposed estimating the domestic content embodied in foreign exports. Acknowledging China’s dual-trade regimes— i.e., the processing trade and normal trade—Chen et al. (2009) partitioned the IO table into two different parts: the IO framework that distinguishes the processing trade from normal trade (followed by Dean et al., 2011; Koopman et al., 2012). This was made possible due to China’s special regulation on the processing trade, which meant that processing imports

<sup>19</sup> In fact, Ghosh (1958) proposes a supply-driven IO framework and this has stimulated intense debate (de Mesnard, 2009; Guerra and Sancho, 2011; Oosterhaven, 1996, 2006; 2012); Dietzenbacher (1997) suggests interpreting Ghosh’s model as a Leontief price model.

<sup>20</sup> The value-added coefficient is only for illustrative purposes. In a broader context, under the generalized demand-driven IO model of Leontief, it can be the employment coefficient, the emissions coefficient or the energy intensity and so on; the procedure remains almost the same (due to the fixed-proportion nature of the IO analysis). Related theoretical and empirical discussions along this line can be found in Gurgul and Lach (2018c), and Miller and Temurshoev (2015).

<sup>21</sup> Reflecting on the issue of the processing trade, the NBS changed the way it records exports and imports in the IO tables, starting with the compilation of the 2007 IO table. Up until the compilation of this table, exports and imports were acquired from customs data without a distinction being made between the processing trade and normal trade. However, the pure processing trade does not show up in the NBS production data. Hence, there is inconsistency between these two datasets (Pei et al., 2011, is among the first to identify this issue and to resolve it in the context of accounting for import growth). From the 2007 IO table, the value of foreign trade is not the same as those found in the Customs statistics, as only the processing fee is recorded for the pure processing trade. Consequently, it is not appropriate to make any direct comparison of, say, the effects of trade on value-added generation by using IO tables before and after 2007 (but this can be recovered by adding back the value of the products, as is done in the datasets reported in this paper).

can only be used to produce processing exports (i.e., intermediate uses). Thus, it is possible to estimate an import matrix of processing imports. Later on, the extended framework was further developed while taking into account the part about the heterogeneity that exists within domestic production, so we have the full-fledged extended framework (Yang et al., 2015, discusses the differences between the original extended framework and the full-fledged one). The synthesized compilation method of the extended IO table is given in Appendix A.<sup>22</sup>

4. Applications and extensions

4.1. Application of the estimation of the value-added content in exports

In this section, we explore the advantage of using the extended IO framework, for example, estimating the domestic content in exports. As was previously stated, China has been recently characterized by the giant share of its processing exports in its foreign trade. Taking into account the special trade regime, the estimation of the impact of foreign exports on the Chinese economy should definitely account for this issue (e.g., the generation of value added, the emissions embodied in trade, and/or the energy usage embodied in trade).

As the domestic content of China’s exports (in fact, the vertical specialization is the other side of one coin) has been studied extensively these days (see e.g., Kee and Tang, 2016; Chen et al., 2012; Koopman et al., 2012), the present study does not intend to solely focus on this calculation. Rather, our main point is to show the advantages of using an alternative IO table to analyze certain economic issues.

As mentioned, the extended framework for China is a special table that takes into full account the different characteristics of processing exports. More formally, the value added that is generated in order to meet worldwide final demand for each of the four use categories in Fig. 4 is given by the following:

- i) the value added due to  $f^{DD}$  is given by  $(\hat{c}^D L^{DD} + \hat{c}^N L^{ND})f^{DD}$ ;
- ii) the value added due to  $f^N$  is calculated as  $(\hat{c}^D L^{DN} + \hat{c}^N L^{NN})f^N$ ;
- iii) the value added due to  $e^P$  is computed by  $(\hat{c}^D L^{DP} + \hat{c}^P + \hat{c}^N L^{NP})e^P$ ;

and.

- iv) the value added due to  $e^N$  is found by  $(\hat{c}^D L^{DN} + \hat{c}^N L^{NN})e^N$ ,

<sup>22</sup> Details on the construction of the dataset and raw data sources are given in Appendix A; and the final extended IO tables are provided as supplemental material, online.

	Intermediate use		Final use	
			<i>DFD</i>	<i>EXP</i>
	$Z^D$		$f^D$	$e$
	$Z^M$		$f^M$	$0$
<i>VA</i>	$v'$			
<i>TOT</i>	$x'$			

Notes:  $Z^D$  = domestically produced intermediate deliveries matrix;  $Z^M$  = imported intermediate deliveries matrix;  $f^D$  = domestic produced final uses;  $f^M$  = imported final uses. Other captions can be found in the notes for Figure 2.

**Fig. 3.** The structure of China’s non-competitive type national input-output table. Notes:  $Z^D$  = domestically produced intermediate deliveries matrix;  $Z^M$  = imported intermediate deliveries matrix;  $f^D$  = domestic produced final uses;  $f^M$  = imported final uses. Other captions can be found in the notes for Fig. 2.

	Intermediate use			Final use		
	<i>D</i>	<i>P</i>	<i>N</i>	<i>DFD</i>	<i>EXP</i>	<i>TOT</i>
<i>D</i>	$Z^{DD}$	$Z^{DP}$	$Z^{DN}$	$f^{DD}$	$0$	$x^D$
<i>P</i>	$0$	$0$	$0$	$0$	$e^P$	$x^P$
<i>N</i>	$Z^{ND}$	$Z^{NP}$	$Z^{NN}$	$f^N$	$e^N$	$x^N$
<i>IMP</i>	$Z^{MD}$	$Z^{MP}$	$Z^{MN}$	$f^M$	$0$	$x^M$
<i>VA</i>	$(v^D)'$	$(v^P)'$	$(v^N)'$			
<i>TOT</i>	$(x^D)'$	$(x^P)'$	$(x^N)'$			

Notes: The direction of the superscripts is from left to the right, e.g., ND denotes transactions from N to D) and their meanings are as follows: *D* = industries producing for domestic use only; *P* = industries producing processing exports; *N* = industries producing normal exports and the production of foreign-invested enterprises for domestic purposes; *M* = imported use. Other captions can be found in the notes for Figure 2.

**Fig. 4.** The structure of China’s input-output table that captures the processing trade. Notes: The direction of the superscripts is from left to the right, e.g., ND denotes transactions from N to D) and their meanings are as follows: *D* = industries producing for domestic use only; *P* = industries producing processing exports; *N* = industries producing normal exports and the production of foreign-invested enterprises for domestic purposes; *M* = imported use. Other captions can be found in the notes for Fig. 2.

where  $c^i$  gives the value-added coefficient of each production class; for example,  $c^P$  represents the value-added coefficient for processing exports.

Usually, calculations must be done using the ordinary national IO table for China, which sketches an average input structure. In order to highlight the consequences of this, we also calculated the value added that was generated to satisfy domestic final demand ( $f^D$ ) and exports ( $e$ ), by using the ordinary IO table (see Fig. 3).

Several observations follow from the aggregate results provided in Table 2. First, in line with previous findings (see Chen et al., 2012), the role of exports in generating value added had previously been overestimated (thus, China’s vertical specialization shares were underestimated; for similar findings, see Dean et al., 2011; Yang et al., 2015). Second, the value added generated by processing exports amounted to 19% of that generated by all exports, whereas the processing exports accounted for no less than 32% of the total exports in 2015. When compared to the normal exports, this implies that the processing exports generated relatively less value added.

Third, if the full-fledged extended IO table would not have been available—which is the case for many other economies in the world—we would have been forced to use the ordinary national IO table (or the national IO tables obtained from the WIOD). In that case, the ratio of the domestic content per unit of exports (i.e., the DVAR) would have been

reported as 0.80 in 2015 (which was 0.68 for 1997). This would have been an overestimation by about eight percentage points in 2015 (9 percentage points in 1997); in monetary terms, it would have been 1269 billion Yuan (more than 34% of the volume of China’s trade surplus in 2015). This is because the ordinary IO table was obtained by aggregating the extended IO table (using gross outputs as weights). Because the gross outputs of the processing exports and other production were relatively less than the gross outputs of the production for domestic uses, the average production (or input) structure and the value-added coefficients were very similar to those for domestic use only. In a more general sense, this observation also relates to the issue of aggregation bias, which may turn out to be quite noticeable for highly aggregated data (Kymn, 1990; Gurgul and Lach, 2018c).<sup>23</sup>

<sup>23</sup> It is noted that IO analysis normally presents point estimations, which do not explicitly present error bounds in the data tables. Moreover, much of the IO tables have been estimated using a combination of survey and non-survey techniques. Therefore, possible errors may occur in the IO tables. Admittedly, much work is still needed in this direction (see Rueda-Cantuche et al., 2013, for discussions of the stochastic supply-use table; and Yao et al., 2015, gives a range for the domestic content in China’s exports for 2007, given different assumptions).

**Table 1**  
Chinese national input-output tables, 1973–2017<sup>10</sup>.

Year	Type of table	Number of sectors/products	Comments
1973	Physical	61	First national table (trial)
1979	Physical	61	1973 table updated using the RAS method
1981	Value	21	MBS table
	Physical	146	MBS table
1983	Value	26	MBS table
	Physical	146	1981 updated physical table using the RAS method
	Value	22	1981 updated value table using the RAS method
1987	Value	117	First table based on special input-output surveys
1990	Value	33	1987 updated table using the RAS method
1992	Physical	151	Table based on special input-output surveys
	Value	118	Table based on special input-output surveys
1995	Value	33	1992 updated table using the RAS method
1997	Value	124	Table based on special input-output surveys
2000	Value	40	1997 updated table using the RAS method
2002	Value	123	Table based on special input-output surveys
2005	Value	42	2002 updated table using the RAS method
2007	Value	135	Table based on special input-output surveys
2010	Value	65	2007 updated table using the RAS method
2012	Value	139	Table based on special input-output surveys
2015	Value	65	2012 updated table using the RAS method
2017	Value	145	Table based on special input-output surveys

Note: MBS denotes material balance system. The IO tables are publicly and freely available from the National Bureau of Statistics (NBS). More on the RAS method can be found in Miller and Blair (2009).

<sup>10</sup> This table was reproduced and extended from Table 1 in Chen et al. (2005). In fact, the Institute of Systems Science at the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, has collaborated with the NBS in constructing China's IO table since the very beginning, i.e., since the very first IO table was developed in 1973. Table 1 presents a sketch on the development of China's national IO tables. More specifically, since 1987, it has been a routine job for the NBS to publish benchmark tables (i.e., based on special IO surveys) for the years ending in "2" and "7" (for example, 2007 is the latest benchmark IO table thus far), and to publish extended tables (i.e., using benchmark tables as starting points, relying on mathematical techniques such as the so-called RAS, which is supplemented with macro-economic data for updates, a common practice worldwide) for the years ending in "0" and "5" (for example, 2010 is the latest IO table in China); so, four tables were produced for one decade.

Fourth, when we compare the results from 1997 to 2015, we can see that the biases became smaller in relative terms. Two possible channels account for this observation: a) there was some substitution of domestic intermediate inputs for imported intermediates (as suggested in Kee and Tang, 2016), which led to the convergence of production structures for trade-oriented and domestic-oriented production activities (i.e., trade-oriented production relies increasingly more on domestic inputs); b) the share of the processing trade had declined since 2006, which makes the overall estimation bias smaller. That is, suppose the share of the processing trade were to drop to zero and the dual-trade regime were phased out, then there would be limited need to extend the table along this dimension.

Nonetheless, an eight percentage point bias in the aggregate results may hide very large biases at the sectoral level. To verify our statement,

Table 3 includes sector details. The "true" values for the value added generated by the processing and normal exports were estimated using the full-fledged IO framework, whereas the ordinary framework would have reported only the value added generated by exports (processing and normal exports are lumped together).

On average, the overestimation was about 11.6% (defined as the difference between the sum of the domestic content in exports that were estimated by using the ordinary IO table; i.e., the sum of the values in column (8) and that for processing exports (column (2)) and for normal exports (column (5)) over the sum of columns (2) and (5)). In particular, it is worth noting the overestimation of the contribution of exports to value added for so-called Machinery & Electronic Products.

For example, the overestimation of the contribution of exports to value added can be as high as 23% for Telecommunications Equipment, Computer and Other Electronic Equipment (sector 19), 22% for Transport Equipment (sector 17), and 21% for Common and Special Equipment (sector 16). The primary reason for this result is that the processing trade constitutes a large share of production in these industries, which is in line with previous studies (see Yang et al., 2015; Dean et al., 2011; Chen et al., 2012). In general, it was found that the bulk of the industries with relatively large overestimations consisted of those that produce manufactured products.

In sum, it is found that the extended framework, i.e., the extended IO table that incorporates China's processing trade, is useful for studies related to foreign trade as it gives full consideration to China's special trade regime. For a global scope of this analysis, China's extended IO table may be incorporated in multi-country tables, such as the WIOD database (see Chen et al., 2018, 2019, for an application) and the OECD database (which, along with the WTO, launched the "Made in the World" initiative).<sup>24</sup> In fact, the Ministry of Commerce responded to the WTO/OECD initiative by coordinating the NBS, China Customs and the State Administration of Foreign Exchange to contribute to the construction of China's extended IO table.<sup>25</sup>

#### 4.2. Extensions

Recent studies have proposed new methods to update and forecast national and global IO tables (see, e.g., Timmer et al., 2016; Zheng et al., 2018). These studies echo Wiedmann et al. (2011) proposition that in order to be relevant for policymaking, IO databases need to be created and updated in timely, continuous, consistent, and cost-effective ways. Combined with this strand of research, the extended IO framework for China can even be annualized. By so doing, it would open up more possible uses that would include, for example, dynamic and endogenous IO modeling (Pan, 2006; Gurgul and Lach, 2018b); analyses of key sectors (Miller and Lahr, 2001; Gurgul and Lach, 2018a, 2018c; Temurshoev, 2010; Temurshoev and Oosterhaven, 2014); accounting for emissions (Kander et al., 2015); the impact of the fragmentation of global production on labor markets and incomes (Timmer et al., 2014b); the effects of regional trade agreements on welfare; China's role in climate change (Chen et al., 2018), and so on.

<sup>24</sup> Along this line, advanced economies such as the US and Canada, together with developing economies such as Mexico and ASEAN nations, started to develop national IO tables that took into account heterogeneous production technologies (see also, Michel et al., 2018, using firm-level data to construct detailed export-heterogeneous national IO tables).

<sup>25</sup> This project began in 2012, with researchers from the Chinese Academy of Sciences, the University of Chinese Academy of Sciences, and the University of International Business and Economics. The extended IO tables used in this study are typical products of the joint work between government agencies and research institutes. Since its start, this work has attracted much attention from international agencies, such as the OECD. It is worth noting that the dataset is subject to revision, given further data and more advanced methods being made available.

**Table 2**  
Overview of aggregate results using alternative datasets, 1997 to 2015.

	Full-fledged extended IO table							Ordinary IO table		WIOD
	Processing exports			Normal exports			National exports			
	1	2	3	4	5	6	7	8	9	10
	gross	domestic content	DVAR	Gross	domestic content	DVAR	DVAR	domestic content	DVAR	DVAR
1997	826.02	240.33	0.29	828.30	728.22	0.88	0.59	1129.24	0.68	0.85
2002	1489.26	455.15	0.31	1605.99	1262.65	0.79	0.56	2352.51	0.76	0.81
2007	4662.76	1710.28	0.37	5536.57	4323.73	0.78	0.59	7047.78	0.69	0.73
2010	5011.32	1932.86	0.39	6818.41	5495.60	0.81	0.63	8561.51	0.72	0.76
2012	5445.66	2156.77	0.40	8695.38	7308.33	0.84	0.67	10597.03	0.75	
2015	4953.62	2042.06	0.41	10356.37	8943.56	0.86	0.72	12255.09	0.80	

Notes: Values are in billions of Yuan for both gross exports and the domestic content in exports, in current prices. The domestic value added ratio (DVAR) means the ratio of the domestic content per unit of exports. Using the full-fledged extended IO table, we compute the following columns:  $3 = 2/1$ ;  $6 = 5/4$ ;  $7 = (2 + 5)/(1 + 4)$ ; and using the ordinary IO table, we calculate columns 8 and 9, and have  $9 = 8/(1 + 4)$ ; while the WIOD (version 2013) national IO table is used to estimate the values in column 10.

Source: Authors' own calculation using alternative datasets.

**Table 3**  
Value-added estimates by sector, using different IO frameworks, 2015

IO sector	Full-fledged extended IO table							Ordinary IO table	
	Processing exports			Normal exports			National exports	8	9
	1	2	3	4	5	6	7		
	Gross	domestic content	DVAR	Gross	Domestic content	DVAR	DVAR	domestic content	DVAR
1	0	0		91.52	85.24	0.93	0.93	88.09	0.96
2	0	0		2.66	2.42	0.91	0.91	2.45	0.92
3	0	0		13.71	11.50	0.84	0.84	12.60	0.92
4	0	0		4.01	3.36	0.84	0.84	3.62	0.90
5	0	0	0.51	12.62	11.05	0.88	0.88	11.41	0.90
6	61.01	25.57	0.42	233.08	216.48	0.93	0.82	274.48	0.93
7	73.16	39.48	0.54	490.99	449.99	0.92	0.87	513.19	0.91
8	190.31	67.29	0.35	977.97	902.59	0.92	0.83	1034.17	0.89
9	50.58	32.90	0.65	329.05	290.01	0.88	0.85	332.54	0.88
10	260.91	133.17	0.51	301.02	241.58	0.80	0.67	461.35	0.82
11	49.82	3.25	0.07	67.95	49.00	0.72	0.44	84.71	0.72
12	311.82	132.75	0.43	774.87	632.80	0.82	0.70	915.23	0.84
13	18.83	8.55	0.45	316.46	282.38	0.89	0.87	299.30	0.89
14	74.50	22.31	0.30	445.31	343.88	0.77	0.70	415.41	0.80
15	88.98	44.47	0.50	401.00	338.36	0.84	0.78	411.80	0.84
16	260.54	104.03	0.40	470.08	383.51	0.82	0.67	590.71	0.81
17	131.50	49.51	0.38	258.49	214.55	0.83	0.68	323.27	0.83
18	277.35	153.01	0.55	316.63	258.07	0.82	0.69	484.83	0.82
19	531.54	197.23	0.37	768.75	627.64	0.82	0.63	1011.82	0.78
20	2463.40	996.47	0.40	822.57	577.33	0.70	0.48	1903.53	0.58
21	109.37	32.07	0.29	147.51	120.68	0.82	0.59	196.63	0.77
22	0	0		7.93	6.87	0.87	0.87	6.96	0.88
25	0	0		103.72	93.73	0.90	0.90	92.89	0.90
26	0	0		548.27	481.27	0.88	0.88	495.09	0.90
27	0	0		53.22	51.25	0.96	0.96	50.96	0.96
28	0	0		156.81	141.91	0.90	0.90	139.87	0.89
29	0	0		2240.18	2126.10	0.95	0.95	2098.17	0.94

Note: Values are in billions of Yuan, in current prices, for gross exports and the domestic content in exports. The DVAR means the ratio of the domestic content per unit of exports (note that sectors 23 and 24 do not export, thus omitted here). Using the full-fledged extended IO table, we compute values in the following columns,  $3 = 2/1$ ;  $6 = 5/4$ ;  $7 = (2 + 5)/(1 + 4)$ ; and using the ordinary IO table, we calculate values in columns 8 and 9, and have  $9 = 8/(1 + 4)$ .

Source: Authors' own estimation using the full-fledged extended IO table and the ordinary IO table.

## 5. Concluding remarks

The global-value-chain research that has emerged over the past two decades has a clear policy orientation. The “Made in the World” initiative the WTO and OECD jointly launched aims to establish the inclusion of

trade in value-added databases (i.e., TiVA database). In fact, the analytical framework is relatively mature, and deriving the domestic content in exports is straightforward. However, there is still a great deal of uncertainty in determining the use of import information, given the current data. This is the key problem and it needs further research.



Regardless of whether the purpose is to estimate the value-added content of exports within an input-output framework or the desire is to go one step further and formulate trade policies along the lines of the global value chain (see discussions in Dollar et al., 2019), the key to accomplishing either objective lies in accurately estimating how imports are used.

Pioneered by Chen et al. (2009) and continued in Chen et al. (2012), the structure of the Chinese IO table has undergone dramatic changes over the past decade so as to reflect the unique feature of Chinese foreign trade. About half of the country’s foreign trade falls under the processing trade regime. The IO framework that accounts for China’s dual-trade regime has a very rich trade structure. The separation of the processing trade, normal trade and domestic production in the Chinese IO table is theoretically supported by the theory of firm heterogeneity (Melitz, 2003).

As the need for China’s IO data has grown in recent years, this paper takes the viewpoint of IO data users and, for the first time, a database that includes heterogeneous production has been made available to the public. Previous studies have documented that the implications of such a distinction for estimates of “trade in value added” (e.g., Michel et al., 2018; Chen et al., 2009, 2012) and the “emissions embodied in exports” (Dietzenbacher et al., 2012) are substantial. Recently, the OECD even began the joint Trade by Enterprise Characteristics project with Eurostat,

with the aim of the systematic construction of data material that would explicitly distinguish between the production structures of the units within industries that export and of those units that only deliver to domestic consumers. Hence, re-examining the socio-economic and environmental aspects of trade growth while also taking into account industries’ heterogeneous production structures within economies will prove to be fruitful.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgement**

We are very grateful to two anonymous referees and the editor, Prof. Sushanta Mallick, for their valuable suggestions and helpful comments. Thanks also go to comments from experts at the Hong Kong Conference on Recent Advances in International Trade, Finance and Business 2018. Financial support by the National Natural Science Foundation of China (Nos. 41675139, 72042003) and the Program for Excellent Talents, UIBE, China (No. 17JQ11) are gratefully acknowledged.

**Appendix A. The compilation method of an extended IO table, incorporating heterogeneous production technologies: the China case**

Without loss of generality, we use China as an example to illustrate the procedure of extending the ordinary IO table to account for various heterogeneities (see also Dietzenbacher et al., 2012). The basic idea is to synthesize the compilation method of an extended IO table. To extend the official IO table to account for different input uses across trade modes, we use the official IO table and include information from other available data sources. Likewise, we split the official IO table from China’s NBS into blocks by trade modes. This process ensures that all of the aggregate numbers and the balance conditions in the official IO table are met by the estimated new IO table that incorporates the processing trade.

The key variables in our extended framework are listed below, including both the variables to be estimated and those we obtained from external data sources.

Variables ( $l = D, N, k = D, P, N$ )	Definition	#of unknowns
<i>variables to be estimated</i>		
$z_{ij}^{lk}$	Intermediate good $i$ produced by type- $l$ firms and used by trade mode $k$ in sector $j$	$6n^2$
$z_{ij}^{mk}$	Intermediate good $i$ imported to be used by trade mode $k$ in sector $j$	$3n^2$
$v_j^k$	Value added by sector $j$ for trade mode $k$	$3n$
$y_i^l$	Final goods used domestically produced by type- $l$ firms in sector $i$	$2n$
<i>variables with available data</i>		
$x_j, v_j, y_i, e_i, m_i$	Output, value added, final demand, and exported and imported intermediates by sector, from the official IO table	
$z_{ij}$	Goods $i$ used as intermediate inputs in sector $j$ , without distinguishing trade mode from the official IO table	
$x_j^l, v_j^l$	Output and value added by type- $l$ firms in sector $j$ , known from the Annual Surveys of Industrial Production (ASIP) data, the official IO table, and China Statistical Yearbook	
$y_i^M$	Imported final goods of sector $i$ , estimated from trade statistics and the official IO table	
$m_i^k$	Normal and processing imported intermediate inputs of sector $i$ , known from the merged dataset (ASIP with the trade data) and the official IO table	
$e_i^l$	Normal and processing exports of sector $i$ , known from the combined dataset and the official IO table	

We now obtain the data for output, exports, value added, imports for final demand, and imported intermediate inputs, by type and sector.<sup>26</sup> These data will be used as controls for our estimation. To be more specific, sector-level gross output and value added by each firm type are obtained from the ASIP data and the China Statistical Yearbook.<sup>27</sup>

<sup>26</sup> Reflecting on the issue of the processing trade, the NBS changed the way it records exports and imports in the IO tables, starting with the compilation of the 2007 IO table. Up until the 2007 IO table, exports and imports were acquired from customs data without a distinction being made between the processing trade and normal trade. However, the pure processing trade does not show up in the NBS production data. Hence, there is inconsistency between these two datasets (Pei et al., 2011, is among the first to identify this issue and to resolve it in the context of accounting for import growth). From the 2007 IO table, the values for foreign trade is not the same as those found in the Customs statistics, as only the processing fees are recorded for the pure processing trade. Consequently, it is not appropriate to make any direct comparison of, say, the effects of trade on value-added generation by using the IO tables that were developed before and after 2007 (but this can be recovered by adding back the value of the products, as is done in the datasets reported in this paper).

<sup>27</sup> The ASIP dataset only includes industrial firms. For other industries, such as construction and transportation, we can obtain the sector-level value-added data from the Statistical Yearbooks. The shares of FIEs (Foreign Invested Enterprises) in gross output and value added are estimated based on the total sales of FIEs or the total registered capital of these FIEs.

For example,  $x_j$  can be divided into total output of China-owned enterprises, COEs ( $x_j^C$ ) and FIEs ( $x_j^F$ ).  $e_j^{CP}$  and  $e_j^{CN}$  denote the processing exports and normal exports of COEs, respectively; and the output of COEs for domestic use,  $x_j^D$ , denotes the difference between output and exports ( $x_j^D = x_j^C - e_j^{CP} - e_j^{CN}$ ). The outputs of the processing exports are denoted by  $x_j^P = e_j^{CP} + e_j^{FP}$ , where  $e_j^{FP}$  is the processing exports of FIEs. The normal exports and others (the domestic use of FIEs) are denoted as  $x_j^N = x_j^F - e_j^{FP} + e_j^{FN}$ . Using the same approach, we can also separate value added  $v_j$  into  $v_j^D$ ,  $v_j^P$ ,  $v_j^N$ .

Our estimation procedure is performed by using a quadratic programming model (see Koopman et al., 2012). As indicated above, our method involves estimating six inter-industry-between-firm-type domestic input transactions for matrix  $Z^k$ , three inter-industry-between-firm-type imported input transactions for matrix  $Z^{Mk}$ , three sector-level value added vectors  $V^k$ , and two domestic final demand vectors  $Y^l[fx]$ .

Assuming there are  $n$  sectors, our estimation will involve  $9n^2$  unknowns for the intermediate inputs,  $3n$  unknowns for the value added, and  $2n$  unknowns for the final demand. First, we make conjectures about their values based on the trade statistics, the ASIP, and the official benchmark IO table. These conjectured values are then used as initial values in our estimation.

For domestic final demand, we set the initial value  $y0_i^l$  as the following residual:

$$y0_i^l = \frac{x_i^l - e_i^l}{x_i - e_i} (y_i - y_i^M), \tag{A1}$$

where  $\frac{x_i^l - e_i^l}{x_i - e_i}$  indicates the fraction of domestic use (exclusive of exports) by type- $l$  firms out of the total domestic use supplied by sector  $i$ .

The initial values of value added,  $v0_j^k$ , are obtained from decomposing  $v_j^C$ ,  $v_j^F$  [fx] based on the merged data of the ASIP and the Customs trade statistics.

The initial values of the use of intermediate imports are generated by allocating imported intermediates product  $i$  ( $[fx]m_i^l$ ) in proportion to input  $i$ 's usage in sector  $j$

$$z0_{ij}^{Mk} = \left( \frac{z_{ij}}{\sum_j z_{ij}} \right) m_i^k \tag{A2}$$

where  $[fx]m_i^k$  can be estimated from the detailed trade statistics, using the UN BEC classifications<sup>28</sup>; whereas  $z_{ij}$  is obtained from the official IO table. The fraction in the brackets denotes the proportion of sector  $i$ 's intermediate products used by sector  $j$ . This method is used for compiling the tables for the period before 2007. For the tables for the year 2007 and afterwards, the intermediate import uses are directly obtained through the census data (i.e., import end-use survey, see Appendix B for the questionnaire used) provided by the General Administration of Customs of China.<sup>29</sup>

The initial values for domestically produced intermediates are generated following two steps. First, the total domestic product  $i$  used as intermediate inputs in sector  $j$  are computed as a residual of the total intermediate inputs minus the imported intermediates:

$$z_{ij}^d = z_{ij} - \sum_k z0_{ij}^{Mk} \tag{A3}$$

Second, we obtain  $z0_{ij}^{lk}$  by assuming a proportional use of  $z_{ij}^d$ .

$$z0_{ij}^{lk} = \frac{x_i^l - e_i^l}{x_i - e_i} \frac{x_j^k}{x_j} z_{ij}^d \tag{A4}$$

where  $\frac{x_j^k}{x_j}$  denotes the fraction of the output of type  $k$  firms in sector  $j$ .

Clearly, these initial guesses are not guaranteed to satisfy various economic and statistical restrictions. Therefore, following Koopman et al. (2012), we cast the estimation problem as a constrained optimization procedure, where the optimization program is specified to minimize the following objective function:

$$\begin{aligned} \text{Min } S = & \sum_j \sum_i \sum_k \frac{|z_{ij}^{lk} - z0_{ij}^{lk}|}{z0_{ij}^{lk}} + \sum_j \sum_i \sum_k \frac{|z_{ij}^{Mk} - z0_{ij}^{Mk}|}{z0_{ij}^{Mk}} \\ & + \sum_j \sum_k \frac{|v_j^k - v0_j^k|}{v0_j^k} + \sum_i \sum_l \frac{|y_i^l - y0_i^l|}{y0_i^l} \end{aligned} \tag{A5}$$

where the  $z$ 's,  $v$ 's and  $y$ 's are the variables to be estimated; whereas the  $z0$ 's,  $v0$ 's and  $y0$ 's are the initial values that have been specified above. Equation (A5) minimizes the sum of the differences of the estimated values from their initial values. Furthermore, the results of this minimization problem should satisfy the accounting identities of both the official IO table and the extended table, in addition to the other regularities we specify below.

Constraint set 1: row sum identities required by the IO table

<sup>28</sup> Normal imports are for intermediate input use and final use, while the processing imports, by definition, are for intermediate input use only. Hence, all imports for final uses are obtained from normal imports.

<sup>29</sup> The NBS constructed a concordance table between the international trade statistics (which are classified according to the Harmonized System Code) and the input-output classification, where the HS Code is used by customs officials around the world to determine the duties, taxes and regulations that apply to these products.

$$\sum_j^n \sum_k z_{ij}^{jk} + y_i^l = x_i^l - e_i^l \tag{A6}$$

$$\sum_j^n \sum_k z_{ij}^{Mk} + y_i^M = m_i \tag{A7}$$

Constraint set 2: column sum identities required by the IO table

$$\sum_i^n \sum_l (z_{ij}^{lk} + z_{ij}^{Mk}) + v_j^k = x_j^k \tag{A8}$$

Constraint set 3: adding up the conditions for intermediate inputs

$$\sum_l \sum_k z_{ij}^{lk} + \sum_k z_{ij}^{Mk} = z_{ij} \tag{A9}$$

Constraint set 4: adding up the conditions for import intermediate inputs

$$\sum_j^n z_{ij}^{Mk} = m_n^k \tag{A10}$$

Constraint set 5: adding up the conditions for value added and final use

$$\sum_k v_j^k = v_j \tag{A11}$$

$$\sum_l y_i^l + y_i^m = y_i \tag{A12}$$

Constraint set 6: non-negativity constraints

$$z_{ij}^{lk}, z_{ij}^{Mk}, v_j^k, y_i^l \geq 0 \tag{A13}$$

The economic meaning of the six sets of constraints are straightforward. Equations (A6)-(A7) pertain to the row sum identities for the expanded IO account, which is also the market clearing condition. These equations state that the total gross output of sector *i* should be equal to the sum of final demand and exports, plus domestic use as intermediates by all production types across all sectors.

Similarly, the total imports should be equal to the imported intermediate inputs that are used across production types and sectors, plus the imports that are delivered to the final users. Equation (A8) provides the column sum identities, maintaining that the total gross inputs of sector *j* should be equal to the intermediated inputs supplied by all of the production types across sectors and the imported intermediate inputs, plus the primary inputs. Equations (A9) to (A12) are a set of adding-up constraints that ensure the consistency of the solution from the model that includes the official statistics on sector-level trade and the transactions within the relevant industries. Finally, Equation (A13) indicates the non-negativity condition.

The nice thing is that the NBS provides survey data for some key numbers that could be used to split the national IO table and it further supplements this data with the import-use survey conducted by China Customs.<sup>30</sup> These data are used to cross-check to ensure the accuracy of the extended IO dataset.

**Appendix B. The Customs import uses survey questionnaire.**

Commodity code: XXXX.XXXX Imports (thousand USD): XXXX Proportion: XX.X%.

Primary classification	Secondary classification (input-output sector)		Amount ratio (%)	Note
Intermediate uses	01	Agriculture		
	02	Coal mining, washing and processing		
	03	Crude petroleum and natural gas products		
	...	...		
	...	...		
	...	...		
	...	...		
	...	...		
	...	...		
	...	...		
	65	Public management and social administration		
Final uses		Household consumption		
		Gross capital formation		

Source: Reproduced from the original Customs import uses survey questionnaire by China Customs.

<sup>30</sup> The NBS survey targeted production entities for 2012 in a module that was added to their input-output survey. The structure was adopted for the compilation of the 2015 tripartite table (which was not entirely satisfactory). The Customs survey mainly targets trading companies by asking for the import use so it could supplement the split of the competitive table into a non-competitive type (a copy of the questionnaire is provided in Appendix B). Yao et al. (2015) fully explored the pros and cons of these surveys.

## References

- Ahmad, N., Escaith, H., Miroudot, S., Webb, C., Yamano, N., 2011. Trade in Value Added: Concept, Methodologies and Challenges. OECD/WTO, June.
- Arnold, J., Javorcik, B., Lipscomb, M., Mattoo, A., 2014. Services reform and manufacturing performance: evidence from India. *Econ. J.* 126, 1–39.
- Autor, D., Dorn, D., Hanson, G., 2013. The China syndrome: local labor market effects of import competition in the United States. *Am. Econ. Rev.* 103 (6), 2121–2168.
- Chen, Q., Löschel, A., Pei, J., Peters, G., Xue, J., Zhao, Z., 2019. Processing trade, foreign outsourcing and carbon emissions in China. *Struct. Change Econ. Dynam.* 49, 1–12.
- Chen, Q., Zhu, K., Liu, P., Chen, X., Tian, K., Yang, L., Yang, C., 2018. Distinguishing China's processing trade in the world input-output table and quantifying its effects. *Econ. Syst. Res.* <https://doi.org/10.1080/09535314.2018.1534225>.
- Chen, X., 1989. Input–output techniques in China. *Econ. Syst. Res.* 1 (1), 87–96.
- Chen, X., Cheng, L., Fung, K., Lau, L., 2009. The estimation of domestic value-added and employment induced by exports. In: Cheung, Y., Wong, K. (Eds.), *China and Asia: Economic and Financial Interactions*. Taylor and Francis Group/Routledge, pp. 64–82.
- Chen, X., Cheng, L., Fung, K., Lau, L., Sung, Y., Zhu, K., Yang, C., Pei, J., Duan, Y., 2012. Domestic value added and employment generated by Chinese exports: a quantitative estimation. *China Econ. Rev.* 23, 850–864.
- Chen, X., Guo, J., Yang, C., 2005. Extending the input–output model with assets. *Econ. Syst. Res.* 17 (2), 211–225.
- Daaniyall, M., Rahman, A., Los, B., Geschke, A., Xiao, Y., Kanemoto, K., Lenzen, M., 2017. A flexible adaptation of the WIOD database in a virtual laboratory. *Econ. Syst. Res.* 29, 187–208.
- de Mesnard, L., 2009. Is the Ghosh model interesting? *J. Reg. Sci.* 49 (2), 361–372.
- Dean, J., Fung, K.C., Wang, Z., 2011. Measuring vertical specialization: the case of China. *Rev. Int. Econ.* 19, 609–625.
- Department of National Accounts, N.B.S., 2009. *The Compilation Methods for China's 2007 Benchmark Input-Output Table*. China Statistics Press, Beijing.
- Dietzenbacher, E., 1997. In vindication of the Ghosh model: a reinterpretation as a price model. *J. Reg. Sci.* 37 (4), 629–651.
- Dietzenbacher, E., Pei, J., Yang, C., 2012. Trade, production fragmentation, and China's carbon dioxide emissions. *J. Environ. Econ. Manag.* 64, 88–101.
- Dollar, D., Khan, B., Pei, J., 2019. Should high domestic value added in exports be an objective of policy? In: Dollar, et al. (Eds.), *Global Value Chain Development Report 2019: Technological Innovation, Supply Chain Trade, and Workers in a Globalized World*. World Trade Organization, Switzerland, pp. 141–153, 1211 Geneva 2.
- ESR, 2008. China's growing Pains – recent input–output research in China on China. *Econ. Syst. Res.* 20 (2) (Special Issue).
- ESR, 2013. Global multiregional input-output frameworks. *Econ. Syst. Res.* 25 (1) (Special Issue).
- Feenstra, R., 1998. Integration of trade and disintegration of production in the global economy. *J. Econ. Perspect.* Fall 31–50.
- Feenstra, R., Jensen, B., 2012. Evaluating estimates of materials offshoring from US manufacturing. *Econ. Lett.* 117 (1), 170–173.
- Feenstra, R., Wei, S. (Eds.), 2010. *China's Growing Role in World Trade*. University of Chicago Press, Chicago.
- Ghosh, A., 1958. Input-output approach in an allocation system. *Economica New Series* 25 (97), 58–64.
- Guerra, A., Sancho, F., 2011. Revisiting the original Ghosh model: can it be made more plausible? *Econ. Syst. Res.* 23, 319–328.
- Gurgul, H., Lach, L., 2018a. Some remarks on a social network approach to identifying key sectors. *Econ. Syst. Res.* 30, 120–135.
- Gurgul, H., Lach, L., 2018b. On using dynamic IO models with layers of techniques to measure value added in global value chains. *Struct. Change Econ. Dynam.* 47, 155–170.
- Gurgul, H., Lach, L., 2018c. Sectoral linkages at the beginning of the 21st century: the role of Polish economy in global production structures. *Commun. Post Commun. Stud.* 51 (4), 299–314.
- Hertel, T., Tsigas, M., 1997. Structure of GTAP. In: Hertel, T. (Ed.), *Global Trade Analysis: Modeling and Applications*. Cambridge University Press, Cambridge, UK.
- Hummels, D., Ishii, J., Yi, K., 2001. The nature and growth of vertical specialization in world trade. *J. Int. Econ.* 54, 75–96.
- Ianchovichina, E., 2004. Trade policy analysis in the presence of duty drawbacks. *J. Pol. Model.* 26, 353–371.
- Ianchovichina, E., Martin, W., 2004. Impacts of China's accession to the WTO. *World Bank Econ. Rev.* 18, 3–27.
- Johnson, R., 2018. Measuring global value chains. *Ann. Rev. Econ.* 10, 207–236.
- Johnson, R., Noguera, G., 2012. Accounting for intermediates: production sharing and trade in value added. *J. Int. Econ.* 86 (2), 224–236.
- Kander, A., Jiborn, M., Moran, D., Wiedmann, T., 2015. National greenhouse-gas accounting for effective climate policy on international trade. *Nat. Clim. Change* 5, 431–435.
- Kee, H., Tang, H., 2016. Domestic value added in exports: theory and evidence from China. *Am. Econ. Rev.* 106 (6), 1402–1436.
- Koopman, R., Tsigas, M., Riker, D., Powers, W., 2013. The implications of using value-added trade data for applied trade policy analysis. In: Elms, D., Low, P. (Eds.), *Global Value Chains In a Changing World*, 109–134. World Trade Organization, Geneva.
- Koopman, R., Wang, Z., Wei, S., 2012. Estimating domestic content in exports when processing trade is pervasive. *J. Dev. Econ.* 99, 178–189.
- Koopman, R., Wang, Z., Wei, S., 2014. Tracing value-added and double counting in gross exports. *Am. Econ. Rev.* 104 (2), 459–494.
- Kyym, K., 1990. Aggregation in input–output models: a comprehensive review, 1946–71. *Econ. Syst. Res.* 2 (1), 65–93.
- Lardy, N., 2002. *Integrating China into the Global Economy*. Brookings Institution Press, Washington, 2002.
- Leontief, W., 1936. Quantitative input-output relations in the economic system of the United States. *Rev. Econ. Stat.* 18, 105–125.
- Lin, X., Polenske, K., 1995. Input–output anatomy of China's energy use changes in the 1980s. *Econ. Syst. Res.* 7 (1), 67–84.
- Los, B., Timmer, M., de Vries, G., 2015. How global are global value chains? A new approach to measure international fragmentation. *J. Reg. Sci.* 55, 66–92.
- Los, B., Timmer, M., de Vries, G., 2016. Tracing value-added and double counting in gross exports: comment. *Am. Econ. Rev.* 106 (7), 1958–1966.
- Mallick, S., Marques, H., 2016. Does quality differentiation matter in exporters' pricing behaviour? Comparing China and India. *China Econ. Rev.* 40, 71–90.
- Mallick, S., Marques, H., 2017. Export prices, selection into exporting and market size: evidence from China and India. *Int. Bus. Rev.* 26 (6), 1034–1050.
- Meade, D., 2010. The U.S. Benchmark IO Table: History, Methodology, and Myths, Paper Presented at the 18<sup>th</sup> *Inforum World Conference*. Hikone, Japan. Sepp. 5–10.
- Melitz, M., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71 (6), 1695–1725.
- Michel, B., Hambÿe, C., Hertveldt, B., 2018. The Role of Exporters and Domestic Producers in GVCs: Evidence for Belgium Based on Extended National Supply-and-Use Tables Integrated into a Global Multiregional Input-Output Table. NBER Working Paper #25155.
- Miller, R., Blair, P., 2009. *Input-Output Analysis: Foundations and Extensions*, second ed. Cambridge University Press, Cambridge, UK.
- Miller, R., Lahr, M., 2001. A taxonomy of extractions. In: Lahr, M.L., Miller, R.E. (Eds.), *Regional Science Perspectives in Economic Analysis*. Elsevier Science, Amsterdam, pp. 407–441.
- Miller, R., Temurshoev, U., 2015. Output upstreamness and input downstreamness of industries/countries in world production. *Int. Reg. Sci. Rev.* 40 (5), 443–475.
- Naughton, B., 2007. *The Chinese Economy: Transitions and Growth*. MIT Press, Cambridge, Massachusetts.
- OECD/WTO, 2012. *Measuring Trade in Value Added: an OECD-WTO Joint Initiative*.
- Oosterhaven, J., 1996. Leontief versus ghoshian price and quantity models. *South. Econ. J.* 62 (3), 750–759.
- Oosterhaven, J., 2006. On the plausibility of the supply-driven input-output model. *J. Reg. Sci.* 28 (2), 203–217.
- Oosterhaven, J., 2012. Adding supply-driven consumption makes the Ghosh model even more implausible. *Econ. Syst. Res.* 24 (1), 101–111.
- Pan, H., 2006. Dynamic and endogenous change of input-output structure with specific layers of technology. *Struct. Change Econ. Dynam.* 17, 200–223.
- Pei, J., Dietzenbacher, E., Oosterhaven, J., Yang, C., 2011. Accounting for China's import growth: a structural decomposition for 1997–2005. *Environ. Plann.* 43, 2971–2991.
- Rueda-Cantuche, J., Dietzenbacher, E., Fernández, E., Amores, A., 2013. The bias of the multiplier matrix when supply and use tables are stochastic. *Econ. Syst. Res.* 25 (4), 435–448.
- Temurshoev, U., 2010. Identifying optimal sector groupings with the hypothetical extraction method. *J. Reg. Sci.* 50 (4), 872–890.
- Temurshoev, U., Oosterhaven, J., 2014. Analytical and empirical comparison of policy-relevant key sector measures. *Spatial Econ. Anal.* 9, 284–308.
- Timmer, M., Erumban, A., Los, B., Stehrer, R., de Vries, G., 2014a. Slicing up global value chains. *J. Econ. Perspect.* 28, 99–118.
- Timmer, M., Los, B., Stehrer, R., de Vries, G., 2014b. Fragmentation, incomes and jobs: an analysis of European competitiveness. *Econ. Pol.* 28, 613–661.
- Timmer, M., Los, B., Stehrer, R., de Vries, G., 2016. *An Anatomy of the Global Trade Slowdown Based on the WIOD 2016 Release*. GGDC Research Memorandum 162. University of Groningen, Groningen.
- Wang, Z., 2003. The impact of China's WTO accession on patterns of world trade. *J. Pol. Model.* 25, 1–41.
- Weber, C., Peters, G., Guan, D., Hubacek, K., 2008. The contribution of Chinese exports to climate change. *Energy Pol.* 36, 3572–3577.
- Wiedmann, T., Wilting, H., Lenzen, M., Lutter, S., Palm, V., 2011. Quo vadis MRIO? Methodological, data and institutional requirements for multi-region input-output analysis. *Ecol. Econ.* 70, 1937–1945.
- World Development Report, 2020. *Trading for Development in the Age of Global Value Chains*. World Bank, Washington, DC.
- Yang, C., Dietzenbacher, E., Pei, J., Chen, X., Zhu, K., Tang, Z., 2015. Processing trade biases the measurement of vertical specialization in China. *Econ. Syst. Res.* 26, 60–76.
- Yao, S., Ma, H., Pei, J., 2015. Import uses and domestic value added in Chinese exports: what can we learn from Chinese micro data? In: Houseman, S., Mandel, M. (Eds.), *Measuring Globalization: Better Trade Statistics for Better Policy*, vol. 2. W.E. Upjohn Institute for Employment Research, Kalamazoo, MI, pp. 221–243.
- Zheng, H., Fang, Q., Wang, C., Jiang, Y., Ren, R., 2018. Updating China's input-output tables series using MTT method and its comparison. *Econ. Modell.* 74, 186–193 (\*).