

RESEARCH

Open Access



Improving the European input–output database for global trade analysis

José M. Rueda-Cantuche¹, Tamas Revesz^{1,2}, Antonio F. Amores¹, Agustín Velázquez¹, Marian Mraz¹, Emanuele Ferrari¹, Alfredo J. Mainar-Causapé^{1*}, Letizia Montinari¹ and Bert Saveyn¹

*Correspondence:
amainar@us.es

¹ European Commission,
Joint Research Centre, Seville,
Spain

Full list of author information
is available at the end of the
article

Abstract

There are increasing numbers of published articles in the field of input–output analysis and modelling that use the GTAP input–output database; particularly, in relation to the estimation of carbon, energy and water footprints and the analysis of global value chains and international trade. The policy relevance of those topics is also increasing, thus calling for consistently linking these databases with official statistics. Although, so far, GTAP has been using their own classification and reconciliation methods, this paper develops a new conversion method for the EU that guarantees that the EU-GTAP database respects the new statistical standards and Eurostat official statistics. We recommend for future updates, a shift of the current GTAP classification of industries to the new official standard classifications to which countries are progressively moving to. Otherwise, the lack of matching official data would jeopardize the usefulness of such database. This method can be extended to other similar input–output databases with different classification schemes from the original input data sources.

Keywords: Input–output tables, European Union, GTAP, Supply and use tables

1 Introduction

During the last 25 years, an increasing number of academic articles and policy reports applying input–output analysis and multisectoral modelling have used the GTAP databases. GTAP (Global Trade Analysis Project) is a global network of researchers and policy-makers conducting quantitative analysis of international policy issues. GTAP's goal is to improve the quality of quantitative analysis of global economic issues within an economy-wide framework (<https://www.gtap.agecon.purdue.edu/>).¹ Recent applications² span areas ranging from carbon emission and climate change (Weber and Matthews 2007; Lenzen et al. 2010; Wiedmann et al. 2010; Saveyn et al. 2011; Peters et al. 2011; Arto et al. 2014; Steen-Olsen et al. 2014; Edens et al. 2015; Labat et al. 2015; Vanduyck et al. 2016), energy (Wiebe et al. 2012; Peters and Hertel 2016), air quality (Vrontisi et al. 2016; OECD 2016, Kitous et al. 2017) and water footprints (Feng et al. 2011; Holland et al. 2015; Cazcarro et al. 2016); to agricultural economics and policy (Banse

¹ The GTAP Project is coordinated by the Center for Global Trade Analysis, which was founded in 1992 by Professor Thomas W. Hertel and is housed in the Department of Agricultural Economics at Purdue University.

² The literature review is organised by topic even though in some cases the database used may not be, strictly speaking, the GTAP database of national IOTs but rather multiregional IOTs derived from such database.

et al. 2008; Hertel et al. 2010), trade policies and free trade agreements (European Commission 2013, 2016; Pelkmans et al. 2014; Boyer and Schuschny 2010; Kutlina-Dimitrova 2015; USITC 2006, 2007), and the analysis of global value chains, productivity and international trade (Lejour et al. 2012; Roson and Sartori 2016; Owen et al. 2016; Steen-Olsen et al. 2016). Users of the GTAP database can be found in universities, academic institutions and intergovernmental organizations³ alike. Given the global reach of these publications and their policy relevance, as well as the differences found in the literature with respect to other similar input–output databases (Andrew and Peters 2013; Owen et al. 2014; Inomata and Owen 2014; Peters et al. 2011; Tukker and Dietzenbacher 2013; Jones et al. 2016), there is a pressing need for consistently matching the GTAP database with official national statistics from countries and making sure that the GTAP database has sufficient statistical quality to address such policy analyses. The GTAP database requires input–output tables at basic prices with a distinction between domestic and imported uses and input–output tables at producer prices, including separated matrices of taxes less subsidies on products (Huff et al. 2000).

With such purpose, this paper develops the EU-GTAP conversion method, a new conversion method for the whole EU that guarantees that the EU data supplied to the GTAP database respect the new statistical standards (European System of Accounts—ESA10⁴) and Eurostat (ESTAT) official statistics. The resulting input–output tables (IOTs) in GTAP for the EU (for each Member State, 28 sets of tables) include the most recent updated supply, use and input–output tables (SUIOTs) and methods from Eurostat, while they are in line with the GTAP requirements. Further, the work follows Eurostat’s recommendations for the estimation of missing IOTs (Rueda-Cantuche et al. 2017). Eurostat has been consulted throughout the different stages of the work. Even though this conversion method has been developed for the GTAP database, it can be easily extended and applied to other similar IO databases with different classifications schemes in relation to the original data sources.

Figure 1 schematically shows the EU-GTAP conversion method that is further explained in detail with Additional files 1, 2, 3, 4 and a numerical example. Section 2 gives an overview of the GTAP requirement, the data sources and estimation methods for missing input data. Section 3 states the problem of conversion and describes the main challenges while Sect. 4 provides some insights about the possible causes of the differences found between the current and former estimates of the EU-GTAP IOTs. Section 5 concludes with some lessons learnt and recommendations for future updates. Additional files 1, 2, 3, 4 are also provided online including: (1) a step-wise detailed description of the EU-GTAP conversion method with a numerical example in a separate file; (2) the correspondence list of the Eurostat and GTAP sectorial classifications, and

³ For a list of the GTAP consortium, please look at <https://www.gtap.agecon.purdue.edu/about/consortium.asp>.

⁴ The new SNA08/ESA10 brings in new data challenges that lead to make trade statistics and national accounts trade values more different than ever before (e.g. goods sent abroad for processing and merchanting). The reader should be aware that other countries in the GTAP database may have not shifted to the new system yet or that the corresponding satellite accounts are not consistent with the new statistical regulations. This is also valid for other international database projects such as those compiling global multiregional IOTs. Progressively, all countries will be producing official statistics under the new regulations, but in the meantime there is indeed a period where countries may differ in their statistical production of national IOTs and we can do very little about it, i.e. just be cautious in our analyses.

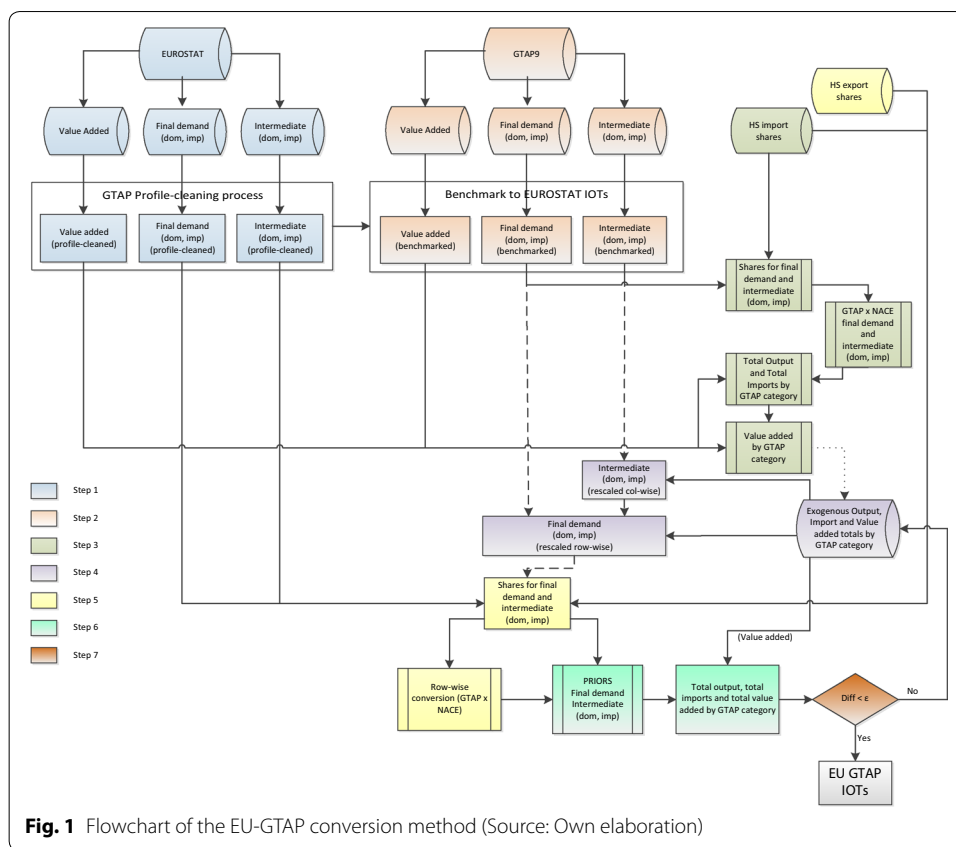


Fig. 1 Flowchart of the EU-GTAP conversion method (Source: Own elaboration)

(3) more detailed insights into some of the element-wise differences between the current and former versions of the EU-GTAP IOTs.

2 Preparing the official EU input–output tables

This section briefly introduces the GTAP requirements for the IOT submissions, describes the data sources used, and reviews the main features of the construction of the missing EU input–output tables. This section basically describes how we made sure that the most recent updated Eurostat data and methods were incorporated in the results of the GTAP database.

2.1 GTAP requirements

The main data sources of the EU-GTAP Project⁵ are the supply, use and input–output tables (SUIOTs) and the valuation matrices (matrices of Taxes less Subsidies on Products and matrices of Trade and Transport Margins) of the 28 Member States. Besides, other datasets were also useful for transforming the resulting tables from the CPA/NACE⁶ format into the GTAP classification. Huff et al. (2000) describes the following requirements of the input–output databases contributions to the GTAP database:

⁵ EU-GTAP project refers to European Commission Project JRC N°33705-2014-11/DG TRADE 2014/G2/G10, realized in collaboration with the GTAP consortium and Eurostat.

⁶ Classification of Products by Activity, <https://ec.europa.eu/eurostat/web/cpa-2008>; NACE is the acronym for “Nomenclature statistique des activités économiques dans la Communauté européenne” or the Statistical classification of economic activities in the European Community (NACE Rev.2, 2008): <https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF>.

- a. The construction of product-by-product input–output tables;
- b. The product breakdown of the original IOT (CPA/NACE) must match the GTAP sectorial classification and the resulting IOTs must have the required GTAP's format;
- c. Treatment of imports (e.g. proportional split of total uses into domestic uses and imports);
- d. Checking accounting identities and non-negativity;
- e. Reporting data sources and problems encountered must be documented.

In line with these requirements, the final dataset consists of a set of IOTs for the 28 EU Member States for 2010⁷ in the new ESA10/SNA08⁸ and the GTAP classification. In particular, the EU submission to GTAP corresponds to IOTs at basic prices (known as 'UF tables' in Huff et al. 2000), with a distinction between domestic and import use matrices and IOTs at producer prices (known as 'UP tables' in Huff et al. 2000)⁹—i.e. with the taxes less subsidies on products included.

2.2 Estimating the missing EU input–output tables

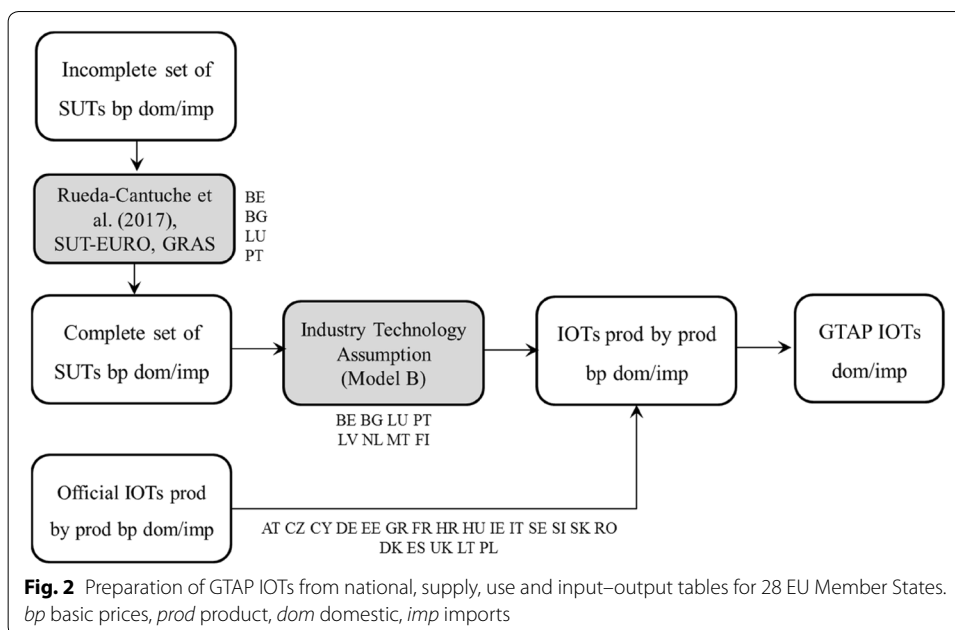
Data released by National Statistical Institutes (NSIs) do not always cover all the elements of a full supply, use and input–output framework, i.e. use tables at basic prices with a distinction between domestic and import uses are not usually available on an annual basis. As of July 31, 2016 (see Fig. 2) Eurostat provided 20 official IOTs (Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, France, Germany, Greece, Hungary, Ireland, Italy, Lithuania, Poland, Romania, Slovakia, Slovenia, Spain, Sweden and the United Kingdom). Other four countries (Finland, Latvia, Malta and the Netherlands) provided supply and use tables (SUTs) at basic prices. Another four countries (Bulgaria, Belgium, Luxembourg and Portugal) did not provide any of the supply, use and input–output tables at basic prices. All SUIOTs have been used in their national currency (the conversion to Euro was made on a later stage using Eurostat's annual exchange rates and once all the countries were available, and the conversion to USD was made by the GTAP consortium after they received our contribution).

Following the approach of Rueda-Cantuche et al. (2017), the missing supply and use tables (SUTs) at basic prices were estimated using the SUT-EURO method as in Valderas-Jaramillo et al. (2018) (Fig. 2). Compared with other methods like those using cross-entropy functions or the minimum information loss principle, the SUT-EURO method assumes the Leontief input–output model to make the estimations, rather than minimizing the distance between the resulting table and the initial one. The Bulgarian SUTs at basic prices from 2011 was rescaled back to 2010. For Belgium and Luxembourg, we used an alternative bi-proportional adjustment method (generalized RAS—Temurshoev et al. 2011) given the available supply and use tables at basic prices in ESA95 and

⁷ According to the ESA10 Transmission Programme of data, EU Member States must provide Eurostat with supply, use and input–output tables at basic prices, including trade and transport margins matrices and taxes less subsidies on products matrices every year ending in 0 and 5. Use tables at purchaser's prices are also compulsory on an annual basis together with the supply tables at basic prices. Accordingly, we preferred to make the conversion of the ESTAT IOTs of the year 2010 (instead of 2011, the base year of the GTAP database), because we had much more official data available. The projections to 2011 were done by the GTAP consortium and falls beyond the scope of this article.

⁸ http://ec.europa.eu/eurostat/statistics-explained/index.php/European_system_of_national_and_regional_accounts_-_ESA_2010.

⁹ The construction of the UP tables falls beyond the scope of this paper.



in ESA10. For Portugal, we used the row structures of the (available) input–output table of imports (from a previous year) to estimate use tables of imports and, by difference, use table of domestic uses.

There are various methods to derive IOTs from SUTs (Eurostat 2008). The product-by-product IOTs are based on technology assumptions while the industry-by-industry IOTs are derived from sales structure assumptions. A “technology” assumption is a strong assumption in the sense that it is based on production theory that cannot be underpinned by observed statistical data. The sales structure assumptions are weaker assumptions, as in general, they only utilize observed sales structures for the actual year. From a statistical perspective, the two types of IOTs thus reflect quite different approaches (see Rueda-Cantuche (2011) for a detailed discussion about the pros and cons of the types of IOTs and the choice of assumptions). In this paper, we have constructed product-by-product IOTs using the industry technology assumption (Model B, in Eurostat 2008) as Eurostat does for the estimation of the EU and Euro Area consolidated IO tables. Although the product technology assumption might be preferable, it can lead to implausible solutions such as negative coefficients. All other remaining 24 countries had either SUTs (i.e. Finland, Latvia, Malta and the Netherlands) or product-by-product IOTs (remaining 20 countries) at basic prices available.

The IOTs were eventually transformed into IOTs at basic prices with a distinction between domestic and import uses (equivalent to the so-called UF tables, according to the GTAP requirements in Huff et al. 2000) complying with the GTAP classification. If the distinction between domestic and import uses was missing, we followed Rueda-Cantuche et al. (2017) to estimate separately domestic and import SUTs before making the conversion to the GTAP classification.

The confidential, missing values of the SUIOTs of Ireland, Lithuania, Poland and Sweden were estimated by applying the column/row structures of the EU consolidated IOTs published by Eurostat. The adjustments were absorbed by the highest value of the final

demand category in each case and by the gross operating surplus in order to keep the balance of total supply and demand unchanged. However, the GDP of the affected countries were very slightly modified as a result of these adjustments (Ireland: 0.006%; Lithuania: 0.014%; Poland: 0.0004%; Sweden: 0.007%).

3 The EU-GTAP conversion method for input–output tables

This section addresses the main challenges of the EU-GTAP conversion method. The conversion of the 2010 ESTAT IOTs into the 57 GTAP sectors turned out to be highly time and resource consuming, mainly due to the fact that the GTAP classification has a clear correspondence only to the NACE Rev.1.1/ISIC Rev.3 classification, but not to the new NACE Rev.2/ISIC Rev.4 classification which is used in the current ESTAT SUIOTs.

3.1 GTAP vs. NACE Rev.2/ISIC Rev.4 classification

The conversion implies (dis)aggregations of four different types: one-to-one cases, many-to-one cases, one-to-many cases, and many-to-many cases (Table 1). The ‘One-to-One’ and ‘Many-to-One’ cases are relatively straightforward.

The ‘One-to-Many’ case requires certain allocation shares that need to be collected from detailed (national) statistics. Regarding allocation shares, we must be aware that there is no single fits-all matrix of allocation shares because their values depend on the country, variable (imports, exports, domestic goods, value-added components and indirect tax) and user sector (for intermediate use). Therefore, country-, category- and user-specific allocation shares are computed from the relevant detailed statistics of imports, exports, gross outputs, labour costs, etc., and in the case of the user sectors, also from the GTAP 9.1 database (which is based on the data available from the most recent year). We should also note that allocation shares (transformation coefficient matrices) are recalculated in various steps of the matrix adjustment process.

The ‘Many-to-Many’ cases are much more complicated. For instance, “man-made fibres” (C20.6) are considered chemical products (C20) in the ESTAT IOTs, but they are considered instead textile products (tex) in the GTAP classification. This implies that a part of the ESTAT sector C20 (i.e. C20.6) has to be reallocated to the ESTAT sector C13 (Textiles) because the GTAP sector (tex) includes “man-made fibres”. As a result, the adjusted (or modified) new ESTAT sector C13 should now include all of the same (textile) commodities as the GTAP sector “tex”, leading to a one-to-one correspondence (i.e. ESTAT adjusted sector C13 vs. GTAP sector “tex”). Ultimately, the rest (remaining part) of the ESTAT sector C20 (along with the C21 and C22 coded sectors) would fully correspond to the GTAP sector of chemical products (crp).

3.2 Many-to-many conversion

The procedure designed to deal with the many-to-many cases consists of seven steps:

1. *GTAP-profile cleaning* process for the domestic and import flows of the IOTs, both for final and intermediate uses, following an “intermediate classification” (IMC),

Table 1 Examples ESTAT-GTAP sector correspondence

One-to-one: one single ESTAT sector corresponds to one single GTAP sector. GTAP IO values fully match those of ESTAT	E.g. insurance (isr), water transport (wtp), air transport (atp), among others
Many-to-one: many ESTAT sectors correspond to one single GTAP sector. The conversion is a simple aggregation	E.g. trade (trd), which gathers “wholesale and retail trade; repair of motor vehicles and motorcycles” (G45, G46, G47), “accommodation” (I55) and “repair of computers and personal and household goods” (S95)
One-to-many: one single ESTAT sector corresponds to many GTAP sectors. Different allocation shares have been used to make the splits	E.g. “electricity, gas, steam and air conditioning supply” (D35), which has to be split up into “electricity” (ely) and “gas manufacture distribution” (gdt); and the “crop and animal production, hunting and related service activities” (A01), which has to be broken down into twelve different GTAP sectors
Many-to-many: many ESTAT sectors correspond to many GTAP sectors. Different allocation shares have been used to make the splits	E.g. “motion picture, video and television programme production, sound recording and music publishing activities” (J59), of which “sound recording and music publishing activities” (J59.2) must be allocated to the GTAP sector “paper products and publishing” (ppp) and “motion picture, video and television programme activities” (J59.1) that has to be allocated to “recreational and other services” (ros). Besides, the GTAP sector “ppp” is also made up of contributions from ESTAT sectors such as “paper and paper products” (C17), “printing and recording services” (C18) and “publishing activities” (J58); and the GTAP sector “ros” is made up of contributions from “creative, arts and entertainment services; library, archive, museum and other cultural services; gambling and betting services” (R90 to R92), “sporting services and amusement and recreation services” (R93), “other personal services” (S96) and “services of households as employers” (T97)

which consists of ESTAT-adjusted NACE sectors that correspond one-to-one to GTAP sectors;¹⁰

2. *Block-wise adjustment* of the base year GTAP9 IOTs (block-wise add-up consistency) to the ESTAT IO data;
3. *Estimation of total imports, gross output and value added* by GTAP commodities/sectors;
4. *Adjustment of intermediate and final uses of domestic goods* to gross output by sector and adjustment of intermediate and final uses of import products to imports by commodity (row-wise transformation of IMC into GTAP classification);
5. *Recalculation of conversion coefficients* matrices;
6. *Estimation of the preliminary GTAP IOTs* prior to its final balancing process (column-wise transformation from IMC into GTAP classification for intermediate uses, value-added components and taxes less subsidies on products);
7. *Estimation of the final GTAP IOTs* via a two-matrix optimization model fulfilling all required constraints.

Additional files 1, 2, 3, 4 gives a detailed description and provides a numerical example of the above seven-step process. Furthermore, it provides the GTAP-profile cleaned (IMC) ESTAT-adjusted sectors and their correspondence to the GTAP sectors. This correspondence is based on Narayanan et al. (2009) mapping between NACE Rev.1.1 and the list of 57 GTAP sectors, the ESTAT’s official tables mapping NACE Rev.1.1 to NACE

¹⁰ The GTAP-profile cleaning process aims at elaborating a sort of intermediate classification under which there are no “many-to-many” cases any more.

Rev.2 at 6-digit level and the specific mapping table between NACE Rev.2 (4-digit) and GTAP sectors produced (although more aggregated) by the APRAISE research project (EPU-NTUA 2013).

The EU-GTAP conversion method is very data-intensive and a number of auxiliary datasets are also required. The necessary data to estimate category-, country- and use(r)-specific transformation coefficient/share matrices to disaggregate the elements of the domestic and import ESTAT IOTs and, subsequently, convert them into GTAP IOTs are listed in Table 2.

Producing the domestic and import prior matrices is a complicated, long and sequential process. Given the 28 countries, the roughly 6000 elements of the domestic and import matrices (4000 and 2000, respectively) and the about 53 variables of the same size to be computed to reach the domestic and import prior matrices, one may estimate that about 9 million cells have to be estimated during the process (about half a million per country). The distance-minimizing two-matrix estimating model usually solves Estimating the missing EU input1–2 min for each country.

3.3 Data challenges

The data challenges are plentiful. Here, we list and explain the main issues to be dealt with.

Consistency with ESTAT IOTs is of utmost importance provided the high quality standards followed by official statistics. The estimated GTAP IOTs are therefore consistent (i.e. ‘block-wise’ add-up consistency¹¹) with the ESTAT IOTs. Accordingly, the conversion method from ESTAT IOTs into GTAP IOTs has used official statistics as much as possible to build up first transformation coefficients matrices¹² to convert ESTAT sectors into GTAP sectors and then, benchmark the resulting GTAP IOTs to the ESTAT IOTs. The estimations and the consistency benchmarks were done separately for domestic and import IOTs.

Sometimes, the lack of proper official data called for endogenous estimations of the *gross output by GTAP sectors*. In all those cases (mainly services), all candidate data sources proved to be incomplete. For example, for many EU countries either the data for the tobacco industry were confidential or simply missing. Besides, although the PRODCOM data were sufficiently detailed to work out the correspondence between NACE Rev.2 and GTAP sectors for certain products, they generally proved to be not sufficiently representative to make out reliable estimates of gross output (e.g. the iron and steel products in Hungary).

Similarly, the structural business statistics (SBS), which report gross output (more precisely: ‘production value’) by industry instead of by product, prove to be heterogeneous, which complicated their use in the general transformation process. However, after estimating the missing values from other data sources, the SBS data still were quite useful for most of the sectors and countries.

¹¹ This refers to the consistency between ESTAT-adjusted sectors (IMC) and GTAP IO sectors (see Additional files 1, 2, 3, 4 for more details, i.e. description of Step 2 in the numerical example).

¹² These are coefficients derived from official statistics and auxiliary data to convert ESTAT IO sectors into GTAP IO sectors (more details can be found in Additional files 1, 2, 3, 4 under the description of Step 3 of the numerical example).

Table 2 Auxiliary datasets

Name of dataset	Description
Final report of the APRAISE project	Maps 64 NACE Rev.2 sectors into 26 aggregated GTAP sectors (instead of 57 sectors). As a result, the APRAISE's correspondence matrix could not be used to split the values of agriculture and energy utilities into the requested 12 agriculture and 8 food-industry sectors and into the electricity, supply and gas distribution sectors, respectively
COMEXT/RAMON	COMEXT: Eurostat's 4- and 6-digit-level HS foreign trade statistics for the EU countries RAMON: Eurostat's correspondence tables between the HS 4/6 digits product classification and NACE Rev.2 classification at 4-digit level
PRODCOM	Eurostat database containing the value of output, exports and imports of about 3700 industrial products by NACE Rev.2 classification at 8-digit level
EUROSTAT structural business statistics (SBS)	Dataset containing output, value-added and main cost categories of about 400 industrial and service sectors
US Geological Survey Yearbook	Used to split the IO data on mining
Energy balances of the International Energy Agency	Used to split the electricity, gas, steam and air conditioning supply
Eurostat's Agricultural Economic Accounts	Containing output mix and main costs of the agricultural sector of the EU countries
CAPRI database	http://www.capri-model.org/dokuwiki/doku.php?id=start
GTAP9 dataset	Used mainly as initial shares of domestic and import flows by user
Matrix Insight Ltd (2013)	Report to separate out the tobacco industry and the food and beverage industry
OECD energy price statistics	Used to split the electricity, gas, steam and air conditioning supply and IO data on mining
Sectoral reports	E.g. for rice-processing industries (Global Rice Science Partnership 2013) and sugar manufacturing activities (European Commission 2012)

Other missing data were retrieved from other official sources, or were reasonably estimated by using other related sources as proxies (e.g. detailed employment, energy balance sheet data or agricultural satellite accounts). For the mining sector, the US Geological Survey Yearbook was very useful for the physical output, while the corresponding sales prices were estimated from the OECD energy price statistics (albeit not available by country and user). Sometimes, it was not clear whether the produced quantities were conveyed to the market or even taken into account (imputed) in the IOTs (e.g. in Ireland).

The *taxes less subsidies on products* (TLS) matrices have been estimated for all countries in a product by product format (as in IOTs) and their respective matrix decompositions: Value-added taxes; excises, other taxes on products (excluding excises and import tariffs); import tariffs; and subsidies on products. Subsequently, they were converted into the GTAP format using the results obtained by the EU-GTAP conversion method for IOTs and the benchmarks obtained from Eurostat and European Commission sources.

Sometimes, in the absence of comprehensive and reliable data from official statistics the *value added by GTAP sectors was estimated*. Given the endogenous estimation of gross output, the value added was the difference between the gross output and the total estimated (domestic and imported) intermediate uses. According to Huff et al. (2000), value added has to be split up into three components: labour compensation (lab), capital

compensation (cap) and agricultural land. However, recent versions of these requirements¹³ identified labour, capital and other net taxes on production (ontp) as the main components of the value-added part of the IO table. Accordingly, we used SBS data to estimate the labour compensation components by GTAP sectors and the corresponding estimated gross output, value added or labour compensation to allocate the other net taxes on production. Capital compensation was estimated as the difference between gross output and total intermediate uses (domestic and imported), TLS, labour compensation and other net taxes on production. In the exceptional case when the capital values turned out to be negative, labour cost was estimated as residual instead. Otherwise, trying to estimate labour and capital shares independently might have led to RAS-type infeasibility problems due to the known total value added by category (row) and the estimated (column) value-added totals by GTAP sector. Anyhow, if negative capital returns still persist, we have provided normal (positive) capital shares from other years or other more aggregated product types using mainly the ESTAT's National Accounts and SBS survey data. The same applies to possible negative values in the final demand (i.e. gross fixed capital formation and exports).

HS foreign trade statistics from COMEXT and the Eurostat's RAMON correspondence tables between the HS 4/6 digits product classification and the NACE Rev.2 (4 digits) classification were used for *exports/imports*. Then, by using the APRAISE's correspondence tables between NACE Rev.2 and GTAP sectors, each NACE Rev.2 code was allocated to the appropriate GTAP sector code. In the (rare) cases when not sufficiently detailed information about the correct correspondence to a single GTAP sector was available, the dominant GTAP sector was matched to its correspondent NACE Rev.2 code. In most of the times, "many-to-many" cases were mainly caused by the fact that the natural correspondence of the GTAP sectors is NACE Rev.1.1 instead of NACE Rev.2. For instance, salt recycling activities formerly corresponded to the main related GTAP sector of food industry while for NACE Rev. 2 these would now correspond to mineral products. The APRAISE's correspondence tables were also used for services exports although it would be desirable to use ESTAT's services foreign trade data once these will become more available in the near future.

3.4 Distance-minimizing constrained two-matrix estimation model

Most of the work in the conversion process of ESTAT IOTs was concentrated so far on the estimation of GTAP IOTs fully consistent with ESTAT IO values (block-wise add-up consistency) and with product-wise balanced supply and demand. However, these (prior) tables did not necessarily match the target values of gross output, value added and imports by GTAP sector/product provided by official statistics. Hence, a distance-minimizing constrained two-matrix estimation model was used to find—separately for each country—the final GTAP IOTs (domestic + imports), which was subject to:

- a. Full consistency with ESTAT IOTs (block-wise add-up consistency);
- b. Balanced supply and demand;
- c. Gross output by GTAP sectors (estimated/exogenous);

¹³ <https://www.gtap.agecon.purdue.edu/databases/contribute/altviewformat.asp>.

- d. Value added by GTAP sectors (estimated/exogenous);
- e. Imports by GTAP sectors (estimated/exogenous);
- f. Various constraints on negatives and upper/lower bounds in changes in inventories and export/output ratios that turned out to be necessary.

The objective function was the sum of the squared relative differences of the elements of the final GTAP domestic and import IOTs from the corresponding elements of their initial values (priors) (Friedlander 1961). A detailed description of the two-matrix estimation model is provided in Additional files 1, 2, 3, 4 describing the seventh step of the EU-GTAP conversion method. Needless to say that it is of utmost importance to estimate good prior tables, so that the model can easily find a solution without distorting too much initial values.

The complexity of the distance function used in the optimization model does not render the possibility to visualize the derived distances of the solution. Actually, what matters most would not be the actual distance values, but rather that the difference between the modelled estimated domestic and import matrices and their respective priors is minimum. Therefore, we estimated the average relative (percentage) difference between the modelled estimated coefficients and their priors (initial values in the distance-minimizing model). The results of the above calculations showed an average of 18.8% for all 28 EU Member States, ranging from Luxembourg (46.4%), Malta (45.2%) and Cyprus (36.6%) with the highest scores and Estonia (5.1%), Romania (10.4%) and Slovenia (12.3%) with the lowest ones. However, we must be aware that provided the extensive use of official statistics in our conversion method of the Eurostat input–output tables, a minimal deviation with respect to the prior estimates (or original GTAP 9.1 database) would have been something somewhat unexpected.

4 Comparison with the previous GTAP 9.1 version

A detailed comparison between the estimated results and previous versions of the EU-GTAP IOTs (GTAP9) is discussed in Additional files 1, 2, 3, 4. Generally speaking, it is extremely difficult to isolate how much of the differences found between the former and current estimations are due to the methods of conversion alone since there are many factors affecting those observed differences. A non-exhaustive list is provided hereafter:

Some input coefficients from the ESTAT IOTs or other official statistics are very different from those in the GTAP9 database. The ESTAT values remained unchanged and only in limited cases they were changed after consulting the respective national statistical offices.

Some odd values were found in estimates based on export and imports statistics. Foreign trade statistics generally differ from National Accounts and Balance of Payments statistics (Eurostat 2016), while the mapping made between HS codes, NACE Rev.2 and GTAP sectors may also have played a role. Crowding-out and crowding-in effects have been identified when exports and imports have been estimated to be too high or too low. In other words, if exports are overestimated then, there is a “crowding in/out” effect for the domestic output (underestimated) given a fixed gross output total.

Crowding-out effect can be observed, for example, in the allocation of the Czech cattle breeding sector's gross output (Czech cattle imports are negligible). The prior matrix for the domestic commodity flows accounted 70% of the total output to exports. As a result, the distance-minimizing model also allocated almost 60% of the cattle sector output to export activities, leaving the rest to domestic cattle breeding sales to the cattle meat industry. However, the resulting amount of cattle breeding consumed by the cattle meat industry might be considered too low with respect to countries such as France (less than half). This can be easily explained because of the large cattle exports crowded-out cattle supply for the domestic market, notably for the cattle meat industry.

'Crowding in' occurs in the opposite case, i.e. when the domestic production is overestimated to the detriment of the amount of sales exported, leading to too big input coefficients, such as for instance the Romanian paddy rice production into the rice-processing sector.

Some odd coefficients might have been inherited from *values of previous GTAP9 versions* which were used to compute the initial matrices and therefore, the preliminary GTAP IOTs (priors). For instance, in some cases the average user distribution of input flows across the rows of the IOT was not consistent with the knowledge about the nature of the technology of the given sector (intermediate user). Here, ad hoc adjustments were made.

New technologies can appear. For instance, from 2008 onwards (white) sugar was more and more produced in the EU from isoglucose—called *corn syrup* in the United States—(allocated to other food—"ofd"—in the GTAP classification) rather than from sugar beet—"c_b" (Zimmer 2013). This led to lower sugar beet input coefficients in the sugar industry and higher input coefficients from other food products. The model, however, would have allocated significant amounts to the own-consumption of the sugar industry even in absence of such adjustment, as it would have realized that the total supply (use) of sugar beet had decreased.

Sometimes, the remaining few odd values of the estimation process come from the *limitations of distance-minimizing objective functions*. When the constraints are tight enough, the model tends to find extreme solutions with few extremely high coefficients and others very close to zero. This is generally resolved by using exogenous information and by adding the inverse of the squared relative errors to the objective function (thereby preventing the turn of significantly positive values to zeros).

IOTs are in current prices and therefore, *input coefficients may change from 1 year to another just due to price changes*. In the comparison with the GTAP9 IOTs this is particularly relevant for the sugar industry and the energy sector, where the world oil and gas prices were fluctuating significantly from 2008 to 2015.

When transforming the value-added block of the use/TLS/IOTs to GTAP sectors, we also have to deal with the possible *differences between the definition (content) of the similar Eurostat and GTAP categories*. Most notably, we can mention the following:

1. The GTAP data for "Tax on labor" may include (a large part of) the employees' social security contribution (SSC); thus not being included in the GTAP "wages and salaries" category. However, the Eurostat's category of "wages and salaries" indeed includes SSC. On the other hand, the GTAP databases take into account only the

actual employers' SSC received by the government and exclude all imputed SSC. Similarly, GTAP data seem to be defined more broadly (possibly including other labour taxes too, like payroll taxes or contributions to state apprenticeship and rehabilitation funds).

2. The GTAP database is based to a large extent on direct statistics on agriculture, which consider the mixed income of farmers as wages while Eurostat considers them as part of the intermediate costs or the operating surplus.
3. Some of the taxes/subsidies on products might have been classified by GTAP as other taxes/subsidies on production (part of the gross value added).
4. Related to the previous issue, the GTAP9 figures for the (net) output (or production) taxes were much higher than those of Eurostat. This was partly due to the fact that GTAP used this category as a residual for balancing input–output tables.

For all these reasons, the quality assessment of the new estimates really depends on a variety of many inter-dependent factors that prevent us doing a more precise quantification of the improvements made to the GTAP database due to the conversion method alone. Nevertheless, we are confident that the use of more official statistics and the efforts made to gain consistency with them gives enough reasons to think that the new EU-GTAP database has the same quality standards as the official statistics.

Additional files [1](#), [2](#), [3](#), [4](#) provide a detailed step-wise description of the full process of conversion together with a numerical example in a separate file. They also provide a correspondence table between the GTAP and IMC classifications. And last but not least, a detailed description of some elements of the EU-GTAP database compared with the GTAP9 version is also included in Additional files [1](#), [2](#), [3](#), [4](#).

5 Conclusions

This article describes the work carried out to produce a set of input–output tables for the 28 EU Member States for the reference year 2010 under the new European System of Accounts methodology (ESA10, complying with UN SNA08) and in compliance with GTAP submission requirements.

The main novel contribution of this paper is the development of a new conversion method that consists of seven steps and converts the ESTAT IOTs (NACE Rev.2) into GTAP full domestic and imports product by product IOTs (GTAP classification). The resulting EU-GTAP IOTs fully comply with Eurostat aggregates and subtotals at a certain common level of aggregation as well as with other official statistics. This method could be used for other geographical regions in the world and may serve for future updates of the EU database in GTAP and other similar input–output databases with different product and industry classification from the original data sources.

Regarding the quality assessment of the results coming from the new EU-GTAP conversion method, there is a variety of reasons why the quality assessment of our new estimates cannot be done in a precise manner; these are mainly related to the fact that it is extremely difficult to isolate the improvements made to the GTAP database just due to the application of the new conversion method. As supporting evidence, we can show instead in Additional files [1](#), [2](#), [3](#), [4](#) that the EU-GTAP conversion method reduced

the number of arguable coefficients by 10% with respect to GTAP9, which cannot be regarded as a negligible achievement (see Additional files 1, 2, 3, 4 providing more detailed information on these results). For some countries the reduction was much greater, in particular for the big economies. For Germany, the reduction was 56% and for Italy and the United Kingdom, the corresponding reductions were 39% and 60%, respectively. Croatia's reduction was around 41%. However, for some specific countries, the EU-GTAP estimates were not so successful.

The development of the EU-GTAP conversion method turned out to be highly time and resource consuming, mainly due to the fact that the GTAP classification has a clear correspondence to the NACE Rev.1.1/ISIC Rev.3 classification but not to the new NACE Rev.2/ISIC Rev.4 classification. In addition, the search for more detailed official statistics for 28 individual EU Member States was cumbersome because of the lack of detailed homogenous information on gross output, value added and foreign statistics by GTAP sector, let alone more detailed IOTs.

In our view, the main conclusion that can be drawn from our work for future GTAP database releases is a very strong recommendation to urgently revise the GTAP classification in line with newer classification systems. We are fully aware that revisions in the classifications create inevitably breaks in time-series published in the past and so it will in the GTAP database as well. Then, the GTAP consortium should also think a way to facilitate correspondence tables between old and new classifications wherever appropriate. In addition, countries across the world are progressively moving into NACE Rev.2/ISIC Rev.4 and it will be very difficult to update future GTAP IOTs still based on previous classification systems. For future releases, our estimated transformation coefficient matrices for 2010 (converting trade values from ESTAT IOT sectors to GTAP sectors) may hopefully serve the GTAP consortium as a good basis for developing NACE Rev.2/GTAP sectors correspondence, and therefore, future revisions of the GTAP sectorial classification. Furthermore, future updates of the EU-GTAP IOTs, may benefit from using more detailed¹⁴ SUIOTs from countries, to the extent that they exist.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s40008-020-00208-2>.

Additional file 1. Supplementary Material: Info Method (a).

Additional file 2. Supplementary Material: Info Method (b).

Additional file 3. Supplementary Material: Info Sectors.

Additional file 4. Supplementary Material: Info Results.

Acknowledgements

The authors gratefully acknowledge Zornitsa Kutlina-Dimitrova and Ángel Aguiar for their valuable comments and suggestions in previous versions of this article.

Disclaimer

The views expressed herein are those of the authors and do not necessarily reflect an official position of the European Commission.

Authors' contributions

All authors have made substantial contributions. All authors read and approved the final manuscript.

¹⁴ So far, we used those of the Czech Republic, Denmark, Germany, Hungary, Portugal and United Kingdom.

Funding

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article (and its additional file).

Competing interests

The authors declare that they have no competing interests.

Author details

¹ European Commission, Joint Research Centre, Seville, Spain. ² Corvinus University, Budapest, Hungary.

Received: 16 May 2019 Revised: 18 February 2020 Accepted: 11 April 2020

Published online: 19 April 2020

References

- Andrew RM, Peters GP (2013) A multi-region input–output table based on the global trade analysis project database (GTAP-MRIO). *Econ Syst Res* 25(1):99–121
- Arto I, Rueda-Cantuche JM, Peters GP (2014) Comparing the GTAP-MRIO and WIOD databases for carbon footprint analysis. *Econ Syst Res* 26(3):327–353
- Banase M, van Meijl H, Tabeau A, Woltjer G (2008) Will EU biofuel policies affect global agricultural markets? *Eur Rev Agric Econ* 35(2):117–141. <https://doi.org/10.1093/erae/jbn023>
- Boyer I, Schuschny A (2010) Quantitative assessment of a free trade agreement between MERCOSUR and the European Union. *Series Estudios Estadísticos y Prospectivos (CEPAL)* 69
- Cazcarro I, López-Morales CA, Duchin F (2016) The global economic costs of the need to treat polluted water. *Econ Syst Res* 28(3):295–314
- Edens B, Hoekstra R, Zult D, Lemmers O, Wilting H, Wu R (2015) A method to create carbon footprint estimates consistent with National Accounts. *Econ Syst Res* 27(4):440–457
- EPU-NTUA (2013) Assessment of Policy Impacts on Sustainability in Europe-Baseline and exploratory scenarios, parameters and validation, Deliverable no.: D 4.1, Grant Agreement no.: 283121, Project Acronym: APRAISE, Theme: ENV.2011.4.2.1-1: Efficiency Assessment of Environmental Policy Tools Related to Sustainability
- European Commission (2012) EU sugar balance sheets: final production 2009/10, final production 2010/11, Forecast 2011/12, Balance sheet 2009/10 to 2011/12—Directorate General for Agriculture and Rural Development, Unit C5, version: 22 January 2012
- European Commission (2013) The economic impact of the EU–Singapore Free Trade Agreement, DG for TRADE, Chief Economist Note—Special Report, September 2013, Brussels, Belgium
- European Commission (2016) Assessing the economic impact of the Trade Agreement between the European Union and Ecuador, DG for TRADE, June 2016, Brussels: Belgium
- Eurostat (2008) European manual of supply, use and input–output tables. Methodologies and Working Papers. Office for Official Publications of the European Communities, Luxembourg
- Eurostat (2016) Consistency between national accounts and balance of payments statistics. Office for Official Publications of the European Communities, Luxembourg
- Feng K, Chapagain A, Suh S, Pfister S, Hubacek K (2011) Comparison of bottom-up and top-down approaches to calculating the water footprints of nations. *Econ Syst Res* 23(4):371–385
- Friedlander D (1961) A technique for estimating contingency tables, given marginal totals and some supplemental data. *J R Stat Soc Ser A* 124:412–420
- Global Rice Science Partnership (2013) Rice almanac, 4th edn. International Rice Research Institute, Los Baños
- Hertel TW, Golub AA, Jones AD, O'Hare M, Plevin RJ, Kammen DM (2010) Effects of US maize ethanol on global land use and greenhouse gas emissions: estimating market-mediated responses. *Bioscience* 60(3):223–231. <https://doi.org/10.1525/bio.2010.60.3.8>
- Holland RA, Scott KA, Flörke M, Brown G, Ewers RM, Farmer E, Kapos V, Muggeridge A, Scharlemann JPW, Taylor G, Barrett J, Eigenbrod F (2015) Global impacts of energy demand on the freshwater resources of nations. *PNAS*. <https://doi.org/10.1073/pnas.1507701112>
- Huff K, McDougall R, Walmsley T (2000) Contributing Input–output Tables to the GTAP Data Base. GTAP Technical Paper 1
- Inomata S, Owen A (2014) Comparative evaluation of MRIO databases. *Econ Syst Res* 26(3):239–244
- Jones L, Wang Z, Degain C, Liet X (2016) The similarities and differences among three major inter-country input–output databases and their implications for trade in value-added estimates. In: Paper presented at the 19th annual conference on global economic analysis, Washington DC, USA
- Kitous A, Keramidis K, Vanduyck T, Saveyn B, Van Dingenen R, Spadaro J, Holland M (2017) Global energy and climate outlook 2017: how climate policies improve air quality. JRC Science for Policy Report, Joint Research Centre
- Kutlina-Dimitrova Z (2015) The economic impact of the Russian import ban: a CGE analysis. DG for TRADE, Chief Economist Note, December 2015, Brussels: Belgium
- Labat A, Kitous A, Perry M, Saveyn B, Vanduyck T, Vrontisi Z (2015) GECO2015 Global energy and climate outlook: road to Paris. Assessment of low emission levels under world action integrating national contributions. JRC-IPTS Working Papers JRC95892, Institute for Prospective and Technological Studies, Joint Research Centre
- Lejour A, Rojas-Romagosá H, Veenendaal P (2012) The origins of value in global production chains, Report to the European Commission. CPB Netherlands Bureau for Economic Policy Analysis, The Hague
- Lenzen M, Wood R, Wiedmann T (2010) Uncertainty analysis for multi-region input–output models—a case study of the UK's carbon footprint. *Econ Syst Res* 22(1):43–63

- Matrix Insight Ltd (2013) Economic analysis of the EU market of tobacco, nicotine and related products. Revised Final Report for the Executive Agency for Health and Consumers (Specific Request EAH/2011/Health/11 for under EAH/2010/Health/01 Lot 2)
- Narayanan GB, Dimaranan BV, McDougall RA (2009) Chapter 2. In: Guide to the GTAP Data Base. Center for Global Trade Analysis, Purdue University
- OECD (2016) The economic consequences of outdoor air pollution. OECD Publishing, Paris
- Owen A, Steen-Olsen, Barrett J, Wiedmann T, Lenzen M (2014) A structural decomposition approach to comparing MRIO databases. *Econ Syst Res* 26(3):262–283
- Owen A, Wood R, Barrett J, Evans A (2016) Explaining value chain differences in MRIO databases through structural path decomposition. *Econ Syst Res* 28(2):243–272
- Pelkmans J, Lejour A, Schrefler L, Mustilli F, Timini J (2014) The impact of TTIP: the underlying economic model and comparisons. CEPS Special Report 93, TTIP Series No. 1, Brussels
- Peters JC, Hertel TW (2016) Matrix balancing with unknown total costs: preserving economic relationships in the electric power sector. *Econ Syst Res* 28(1):1–20
- Peters GP, Andrew R, Lennox J (2011) Constructing an environmentally extended multiregional input–output table using the GTAP database. *Econ Syst Res* 23(2):131–152
- Roson R, Sartori M (2016) Input–output linkages and the propagation of domestic productivity shocks: assessing alternative theories with stochastic simulation. *Econ Syst Res* 28(1):38–54
- Rueda-Cantuche JM (2011) The choice of type of input–output table revisited: moving towards the use of supply–use tables in impact analysis. *Sort-Stat Oper Res* 35(1):21–38
- Rueda-Cantuche JM, Amores AF, Beutel J, Remond-Tiedrez I (2017) Assessment of European Use tables at basic prices and valuation matrices in the absence of official data. *Econ Syst Res*. <https://doi.org/10.1080/09535314.2017.1372370>
- Saveyn B, Van Regemorter D, Ciscar JC (2011) Economic analysis of the climate pledges of the Copenhagen Accord for the EU and other major countries. *Energy Econ* 33:533–540
- Steen-Olsen K, Owen A, Hertwich EG, Lenzen M (2014) Effects of sector aggregation on CO₂ multipliers in multiregional input–output analyses. *Econ Syst Res* 26(3):284–302
- Steen-Olsen K, Owen A, Barrett J, Guan D, Hertwich EG, Lenzen M, Wiedmann T (2016) Accounting for value added embodied in trade and consumption: an intercomparison of global multiregional input–output databases. *Econ Syst Res* 28(1):78–94
- Temurshoev U, Webb C, Yamano N (2011) Projection of supply and use tables: methods and their empirical assessment. *Econ Syst Res* 23(1):91–123
- Tukker A, Dietzenbacher E (2013) Global multiregional input–output frameworks: an introduction and outlook. *Econ Syst Res* 25(1):1–19
- USITC (2006) US-Colombia free trade agreement. Potential economy-wide and selected sectoral effects. USITC Publication 3896, December 2006. Washington D.C., United States
- USITC (2007) US-Korea free trade agreement. Potential economy-wide and selected sectoral effects. USITC Publication 3949, September 2007. Washington D.C., United States
- Valderas-Jaramillo JM, Rueda-Cantuche JM, Olmedo E, Beutel J (2018) Projecting supply and use tables: new variants and fair comparisons. *Econ Syst Res*. <https://doi.org/10.1080/09535314.2018.1545221>
- Vandyck T, Keramidias K, Saveyn B, Kitous A (2016) A global stocktake of the Paris pledges: implications for energy systems and economy. *Glob Environ Change* 41:46–63
- Vrontisi Z, Abrell J, Neuwahl F, Saveyn B, Wagner F (2016) Economic impacts of EU clean air policies assessed in a CGE framework. *Environ Sci Policy* 55(1):54–64
- Weber CL, Matthews HS (2007) Embodied Environmental Emissions in U.S. International Trade, 1997–2004. *Environ Sci Technol* 41(14):4875–4881
- Wiebe KS, Bruckner M, Giljum S, Lutz C (2012) Calculating energy-related CO₂ emissions embodied in international trade using a global input–output model. *Econ Syst Res* 24(2):113–139
- Wiedmann T, Wood R, Minx JC, Lenzen M, Guan D, Harris R (2010) A carbon footprint time series of the UK—results from a multi-region input–output model. *Econ Syst Res* 22(1):19–42
- Zimmer Y (2013) Isoglucose—how significant is the threat to the EU sugar industry? *Sugar Ind* 138(12):770–777

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.