

Chapter 1

Introduction

Angel Aguiar and Thomas W. Hertel

The goal of this chapter is to introduce readers to the Global Trade Analysis Project (GTAP) and provide a bit of history on the data base and software which accompany this documentation. Those readers who are already familiar with GTAP may wish to briefly scan Section 1.2 to get a feel for new developments in the GTAP 12 Data Base, before moving on to Chapter 2. However, people who are not familiar with GTAP will likely find that this material provides a useful lens through which to view the GTAP 12 Data Base and associated software. By reading about the historical evolution of this unique database, one acquires not only an appreciation of how far we have come, but also a sense of what might be possible in the future.

1.1 Overview of GTAP

1.1.1 What is GTAP?

GTAP was established in 1992, with the objective of lowering the cost of entry for those seeking to conduct quantitative analyses of international economic issues in an economy-wide framework. The Project consists of several components:

- a fully documented, publicly available, global data base,
- a standard general equilibrium modeling framework,
- software for manipulating the data and implementing the standard model,
- a collection of courses (GTAP-U) offering network members training, starting from the basics of GTAP modeling (GTAP-101) to courses on advanced topics.
- a global network of more than 33,000 researchers in more than 181 countries with a common interest in global economic analysis of trade, resources and the environment,
- a consortium of national and international agencies and institutions providing leadership and a base level of support for the Project, and — a website for dissemination of data, software and project-related information (www.gtap.org).

1.1.2 Motivation for GTAP

As the world economy becomes more interdependent, there is an increasing demand for quantitative analyses of policy issues on a global basis. Due to its coverage across the economy, GTAP is particularly useful for analyzing issues that span many diverse sectors. This data base is popular with researchers analyzing the potential impact of: (a) global trade issues, (b) regional trade agreements or disagreements, (c) economic consequences of attempts to reduce carbon dioxide (CO₂) emissions via carbon taxes, (d) energy and agricultural policies, (e) food, nutrition and sustainability challenges, (f) the economic impacts of climate change, and (g) the distributional consequences of domestic and international policy interventions. Sector-by-sector analyses of these questions can provide valuable input to studies of these issues. However, by their very nature, these types of policy interventions and external shocks generally affect all sectors and many regions of the world, so there is no way to avoid employing an analytical framework which is exhaustive in its coverage of commodities and countries. The Global Trade Analysis Project aims to facilitate such multi-country, economy-wide analyses.

1.1.3 GTAP Data Base

The central ingredient in GTAP's success has been the global data base. It combines detailed bilateral trade, transport and protection data characterizing economic linkages among regions, together with individual country input-output (I-O) data bases which account for inter-sectoral linkages within regions.¹ Construction and maintenance of this data base adhere to the following principles:

Public Availability. The data base is made available to anyone requesting it, for a modest fee. This avoids unnecessary duplication of effort in creating this public good. By charging for the data base, we can also cover a portion of the costs incurred in constructing it. However, those who contribute in a significant way to the construction of this data base receive it for free, and are given access to pre-releases as well, alongside the GTAP Consortium members who provide the bulk of the financial support through their annual contributions to the Project.

Regular Updates. The current release is the twelfth (GTAP 12 Data Base) since 1993. Prior to GTAP 12, the average time between releases of a GTAP release was about three years. Beginning with GTAP 12, we expect to shorten the periodicity to two years. In addition, each new release now includes a time series of data bases reaching back to 2004 and covering all prior release years.

Broad Participation. The network of GTAP users represents an excellent resource for the extension of the data base. Another benefit of broad participation is the extensive scrutiny to which the data base is subjected. Those who identify areas for improvement or extension of the data base are free to make this available to GTAP staff to have it considered for incorporation into the next release of the data base. The operational concept is "if you don't like it, help fix it!"

¹See the glossary for a complete list of regions and commodities for the GTAP 12 Data Base.

Comparative Advantage. By making the data base publicly available and offering to incorporate improvements provided by members of the network, each individual is able to work to his/her own comparative advantage, while capitalizing on the contributions of others.

Documentation and Replicability. One requirement for new contributions to the GTAP Data Base is that the sources and procedures used to create them be provided along with the data. This publication represents a summary of the documentation for the GTAP Data Base. Additional detail may be found on the GTAP website or in the documentation of previous versions of GTAP. However, in spite of this extensive documentation, other questions about data sources and methods may arise. Therefore, we refer interested users to the chapter authors themselves in those cases where this document is insufficiently detailed. Often it is only through bilateral correspondence of this sort that data base limitations can be properly identified and remedied. This takes us to the final ingredient of the GTAP Data Base, namely quality control.

Quality Assurance. As the GTAP Data Base has become more widely used and the policy analyses based on this platform have become more influential, the demands for improved quality control have also increased. Over the past decade, this has emerged as the top priority for consortium members and the Center has therefore devoted increasing resources and attention to this issue. There are a variety of ways in which the Center seeks to provide quality assurance. This begins with documentation and replication. If users do not know what is in the data base and how it has been constructed, they cannot intelligently critique it.

Our approach to quality control differs somewhat between the international data bases (trade data, tariffs, energy volumes, etc.) and the domestic data bases (most notably the national input-output tables).

Quality control for the international data base inputs is in some ways easier. Here we often have alternative data bases against which to compare the source data supplied to GTAP. For example, before the decision was made on where to source the tariff data for the GTAP 8 Data Base, we compared two different international data bases, each of which was maintained by a different consortium member. We also have a policy of placing a premium on continuity of suppliers. Recognizing that data bases only improve through use, revision, and further use, we do not change sources lightly. For example, the international trade data base in the GTAP 12 Data Base is supplied by the same individual (Mark Gehlhar, see Gehlhar 1996) who has supplied trade data for versions 1 – 11 of the data base. Over time, through extensive feedback from users, the methodology used for trade data reconciliation and construction has evolved considerably.

Quality assurance for the domestic data bases is more difficult. This begins with Center staff working closely with potential suppliers of national data and checking their submissions for basic accounting consistency. This process of iteration between the Center and individual contributors often takes several months and, in the case of new regions/contributors, may involve a certain amount of informal training in methods for sector disaggregation, adjustment and re-balancing. Once the data base satisfies the basic GTAP requirements, we move to the next stage of quality assurance. Here, we look for anomalies. The easiest way to identify these is to compare the new data base with the earlier version for that country and look for dramatic changes in basic economic relationships. This is quite effective, but it is not possible when a new region is being supplied.

Furthermore, if the new I-O table is supplied by the same individual, this is just a check on consistency of procedures, not an independent check on content. Therefore, we also compare the data base to a “representative” I-O table that comprises those national data bases with full sectoral coverage. This often turns up errors in data processing and disaggregation. We also seek outside peer reviews of new data base contributions. All of these efforts notwithstanding, there is only so much quality assurance that can be provided at the Center. In the end, we rely heavily on national users to evaluate the quality of the data base. Indeed, before each public release, the data base goes through a series of “pre-releases” in which consortium members and contributors evaluate the data base.

In summary, the Center for Global Trade Analysis places a high priority on quality assurance. However, ultimately, we rely on members of the network, with specific areas of expertise, to scrutinize the data base. In this sense, quality assurance in the global economic data base is an ongoing process.

1.1.4 Model and Software

In order to operationalize this large data base, a standard modeling framework has been developed. The components of this multi-region, applied general equilibrium model are relatively standard. For a complete description of the GTAP modeling standard framework, see Corong et al. 2017. A copy of the current version of the model and software for implementing it are freely available on the GTAP website.²

The standard model is designed to be easy to modify and extend, and there are a variety of model extensions available on the GTAP website, under the technical paper series and in the Journal for Global Economic Analysis (www.jgea.org). These extensions include features such as nutrition and food waste, imperfect competition, firm heterogeneity, technology spillovers, detailed treatment of energy demands and emissions (CO₂ and non-CO₂), air quality, land use, agricultural commodity markets, as well as dynamics.

The model is implemented using the GEMPACK software suite, maintained at the Centre of Policy Studies (CoPS), Victoria University, under the direction of Michael Jerie. This software permits the user to conduct simulations of the standard model in which changes in policy, technology, population, and factor endowments are examined. The user specifies the split between exogenous and endogenous variables (i.e., model closure). Behavioral parameters may also be altered. Outputs include a complete matrix of bilateral trade, activity flows (and percentage changes) by sector and region, private and government consumption, regional welfare, and a variety of summary variables. Users with access to GEMPACK may also modify the theory of the model. The RunGTAP software environment greatly facilitates use of the standard model, as well as replication of the GTAP Technical Papers. There are also other models designed to run on top of the GTAP Data Base, many of these use GAMS. For more information on these, visit the ‘Models’ section of the GTAP website.³

²See: GTAP model web page.

³Other models using the GTAP Data Base.

Additional programs have been developed by members of the GTAP Network for use with the GTAP Data Base and model. Among these programs, there is a facility for disaggregating sectors within the GTAP Data Base (SplitCom) and a facility for aggregating and analyzing tariffs at the HS-6 level (TASTE), both produced by Mark Horridge from the Centre of Policy Studies, Victoria University, Australia.

TASTE for the 2004 data was co-developed with David Laborde (currently at the UN Food and Agriculture Organization–FAO, then from the International Food Policy Research Institute–IFPRI). Subsequently, Janine Pelikan from the Thünen Institute, Germany, has taken the lead in updating TASTE since GTAP 9 using International Trade Centre (ITC) MAcMap tariff data. The Social Accounting Matrix (SAM) extraction facility first developed by Scott McDonald, at the time from Oxford Brookes University, the United Kingdom, and Karen Thierfelder, from the US Naval Academy, USA has been recoded by Erwin Corong from GTAP and is now part of the standard GTAP distributions. These and other utilities are made available to users via the GTAP website (see Data Base — GTAP website).

1.1.5 Courses in Global Trade Analysis

A natural point of entry into the GTAP community is through GTAP 101 – a fully online course offered two times a year. For a more in-depth introduction to the GTAP data base and modeling framework, we offer annual courses during the May - August period – involving a mix of online and in-person instruction – with the latter typically on the campus of Purdue University. The goal of the annual course is to introduce newcomers to the GTAP Model, GEMPACK software and GTAP Data Base. This course begins with an eight-week, web-based introduction to the material, followed by an intensive, on-site component lasting one week. Emphasis on hands-on training has proven to be an excellent way for interested researchers to become fully operational with the GTAP framework.

More specialized online courses are also offered through the educational arm of GTAP (nick-named GTAP-U), including a course focusing on preferential trade agreements (GTAP-PTA) and a new offering on firm heterogeneity within the GTAP framework (GTAP-HET). A short course introducing participants to dynamic economic analysis is also offered, every other year, at Purdue University. More information about these offerings is available on the GTAP-U webpage.⁴

1.2 Historical Development of the Data Base

The GTAP 12 Data Base builds heavily on earlier work at Purdue, as well as research and data base development efforts at a number of national and international agencies. The earliest versions of the GTAP Data Base built very heavily on the SALTER Project which was undertaken at the Australian Industry Commission during the 1980s and early 90s. Version 1 of the GTAP Data

⁴The GTAP University web page.

Base used the same thirteen input-output (I-O) tables as SALTER (and much of the software for processing them), while GTAP developed new bilateral trade and protection data. Versions 2 and 3 added new regions to the data base, while gradually updating the original SALTER I-O tables. Versions 4 and 5 replaced all the remaining original SALTER data bases and added many more regions, reaching a total of 66 in GTAP 5. In version 6 the number of regions increased further to 87. In the GTAP 7 Data Base the number of regions increased again to 113 regions, while one region (Myanmar) was removed due to quality issues in version 7.1, bringing down the number to 112. The GTAP 8 Data Base had 129 regions and two reference years: 2004 and 2007. The GTAP 9 Data Base had 2004, 2007 and 2011 as reference years for 140 regions. With GTAP 10, the sector classification increased from 57 sectors to 65 for all 141 regions and it added 2014 to the previous reference years. Since version 10, each economy in GTAP has been described by 65 products and services. While in version 11, we covered 160 regions, version 12 considered one less aggregate region because of the addition of 4 new countries. In total, GTAP 12 reports 145 countries and 18 aggregate regions to complete the world economy. For GTAP 12 we are adding 2 new reference years 2023 and 2019. Therefore, for the latest GTAP release, the total number of reference years is now seven (i.e., 2004, 2007, 2011, 2014, 2017, 2019, and 2023).

Virtually all of these country additions to the data base have been provided by members of the GTAP Network, often resident in the countries for which they are supplying data. With the expansion of the number of countries in the data base, the relative importance of the remaining, “aggregate regions” has become continually smaller, thereby greatly improving the accuracy with which the global economy is represented.

While it has become relatively routine to add new regions to the data base, the disaggregation of sectors remains a costly venture, as it involves revisiting all of the existing national and international data bases. Versions 4, 5 and 10 of the data base added sectoral detail. In GTAP 4, we disaggregated the 37 sectors previously used to the level of 50 sectors. Much of the additional detail was provided in food and agriculture. While there was already a great deal of detail in this area, it was not particularly useful detail, from the viewpoint of many GTAP Consortium members. In addition, version 4 broke out autos and parts and electronic equipment, in light of their dominance in world trade. In addition, electricity, gas and water were disaggregated in order to better serve the interests of those working on energy-environment issues. Version 5 supported improved analysis of services by disaggregating an additional seven service sectors, including communications, financial services, and insurance, as well as three modes of transportation: air, sea and land. With the advent of the transportation splits, the GTAP 5 and 6 data bases disaggregated international transport margins by mode.

In GTAP 10, we moved from 57 to 65 sectors for each economy in GTAP, the expansion added more detail on manufactures and services. In terms of manufactures, we now separately report chemicals, pharmaceuticals, and rubber and plastics products (formerly all 3 combined into sector ‘crp’) and electrical equipment (from the former sector ‘ome’). In terms of services, we now separately report accommodation and restaurants, warehouse, real estate, education and health services.

For GTAP 10 we also improved the estimation of international transport margins. These are based on reported values from a large set of countries with a sizable set of trading partners. In

previous versions of GTAP, international transport margins by commodity and mode of transport were solely based on U.S. international trade margins estimates by mode and commodity. The new estimates are based on data from U.S. imports of merchandise and North American Trans-border Freight Data, EUROSTAT trade data for European countries, as well as the Latin American Integration Association (ALADI) imports data (see Nuno-Ledesma and Villoria 2019).

Services trade data has presented unique challenges due to the limited availability of disaggregated, bilateral trade data for these rapidly growing sectors of the economy. Prior to GTAP 11, trade in services data was based on unilateral services trade information from the IMF, which we had to bilateralize in-house (McDougall 2002; McDougall and Hagemeyer 2006; Lejour, Leeuwen, and McDougall 2010). These data were bilateralized using the RAS method in versions 3, 4 and 5, then improved by using additional sources such as OECD and EUROSTAT, versions 6 to 10. Since GTAP 11, we take advantage of a recently developed dataset provided by the OECD and WTO called the Balanced Trade in Services (BaTiS) (Liberatore and Wettstein 2021), which provides an initial bilateralization that does not need to be reconciled. The starting point for BaTiS is the trade in services dataset developed jointly by the WTO and the United Nations Conference on Trade and Development (WTO-UNCTAD).

The protection data in GTAP has evolved considerably since the project's inception. Compared to version 1, the process of constructing a global protection data base has become infinitely more sophisticated. Most of the work for the first version of the data base was conducted by Bradley McDonald, while he was employed at the Economic Research Service/US Department of Agriculture (ERS/USDA). Tariff data was drawn from the GATT Trade Policy Reviews, while support for agriculture was taken from a combination of OECD and ERS/USDA country studies of Producer Subsidy Equivalents (PSEs). The culmination of the Uruguay Round negotiations provided a rare opportunity to improve GTAP's protection data. With individual countries submitting tariff schedules to the World Trade Organization (WTO), a rich data base emerged. In GTAP 2, we were able to build on disaggregated tariff data provided by the US Trade Representative's office. These data, documented in Chapter 2 of the GTAP book (Hertel 1997), was aggregated up from the tariff line level using import weights. In this way, the GTAP Data Base was able to capture bilateral variation in tariffs for composite products. This variation, due to the composition of trade interacting with varying tariff rates, has been found to be quite significant in some cases. The GTAP 2 Data Base also witnessed the introduction of a variety of non-tariff barriers (NTBs), including anti-dumping duties, countervailing duties and price undertakings. Unfortunately, these proved to be a one-time only contribution from the WTO and have since been dropped as they became severely outdated.

The protection data in GTAP 3 capitalized on work done for the World Bank's 1995 conference on the Uruguay Round and the Developing Countries (Martin and Winters 1996). Pre- and post- Uruguay Round protection data compiled by the World Bank, based on the WTO's Integrated Data Base, as well as other sources, made this a unique data input. Unfortunately, this was also a one-time effort which was not updated. As a result, the tariff data bases for GTAP 4 and 5 were sourced from the UNCTAD TRAINS data base, via the World Bank, courtesy of Will Martin, Jerzy Rozanski and Emiko Fukase. In GTAP 4, this was supplemented on the agriculture side by estimates of market price support contributed by Marinos Tsigas, then at ERS/USDA. However,

up to this point, the treatment of specific tariffs remained a big problem. Since their aggregation requires conversion to ad valorem equivalent form, and since this conversion required additional price data, this was problematic. Fortunately, in GTAP 5 we were able to draw on the Agriculture Market Analysis Database (AMAD) for agricultural tariffs. This data base was supplemented by estimates of the ad valorem equivalent of specific tariffs compiled by Paul Gibson and collaborators at ERS/USDA. This work revealed the great importance of specific tariffs to overall protection in agriculture in the OECD countries.

However, despite all of these improvements in the tariff data base for GTAP versions 2 - 5, we still lacked a proper treatment of preferences – particularly non-reciprocal preferences granted by industrial countries to developing country trade partners. This omission was particularly problematic in light of the prominent role played by the erosion of such preferences in the context of the Doha Development Agenda (DDA) negotiations at the WTO. This omission was remedied in the GTAP 6 Data Base. Thanks to the outstanding work of two of the GTAP Consortium members – the Centre d’Etudes Prospectives et d’Information Internationales (CEPII) in Paris, and the International Trade Center in Geneva, we have been able to build on the MACMap data base for tariffs and import protection. This includes a comprehensive treatment of trade preferences as well as the conversion of specific tariffs for both agriculture and non-agriculture commodities. Thus, for the first time, we had a comprehensive market access data base that treats agriculture and non-agriculture symmetrically. This was a great advance, and, coupled with detailed data on bound tariffs collected by CEPII, permitted construction of a wide range of highly relevant policy scenarios for the Doha round of trade talks. Since version 8 the contribution of the tariff dataset has come from the ITC-Geneva directly for years 2007, 2011, 2014, 2017, 2019, and 2023.⁵

Another important topic of debate in the DDA negotiations was the reform of domestic subsidies for agriculture. The fact that these had been included in WTO disciplines was a great advance under the Uruguay Round Agreement of the WTO. As increasing attention has been focused on the trade distortions caused by these subsidies, they have evolved from primarily output-based payments to increasingly decoupled payments based, for example, on historical land area. This required a more sophisticated treatment of the subsidies in the GTAP Model. Therefore, in GTAP 5 we took advantage of the PSE classification of subsidies into more refined categories in order to better reflect the economic impact of farm subsidies, treating some as intermediate input subsidies, and others as subsidies on land and/or capital. This work was undertaken at the Danish Institute of Agriculture and Fisheries Economics (SJFI), the ERS/USDA and most recently at the OECD. The allocation of these subsidies is not without controversy. In the end, a standard formula was agreed upon and the same one has been used for the GTAP 6, 7 and 8 Data Bases. Revisions were made in the allocation of different types of subsidy payments, mainly in terms of the extent of decoupling across sectors, in version 7. Further refinements in the allocation approach were made by Badri Narayanan, with inputs from Hans Jensen, in version 8. Detailed information on payments and rates based on the types of payments were provided in version 8, but they were removed in version 8.1, due to the lack of consistency of these rates with the ones in the data base. GTAP 11 and 12 retains all the improvements made in versions 8 to 10 in terms of domestic support dataset, but the contribution for the EU countries for GTAP 10 and 11 came directly from the European Commission’s Joint Research Center (JRC). For GTAP 12, however, the OECD domestic

⁵For 2004 we still rely on prior contribution by CEPII and ITC.

support information was not made available for EU member countries. Therefore we were not able to update this information are still using 2017 rates for EU member countries.

Agricultural export subsidy data for 2004, calculated from country notifications to the WTO was contributed by Aziz Elbehri, then at ERS/USDA, with inputs from David Laborde of IFPRI and Hans Jensen, then at the Department of Food and Resource Economics/University of Copenhagen. These mainly affect the European Union and the United States. This data was carried forward to version 8 from version 7. For 2007 in version 8, David Laborde contributed more comprehensive export subsidies data for 15 sectors, 36 exporters and 227 importers. For 2004, estimates of the export tax equivalent (ETE) of the export quotas on textiles and clothing (wearing apparel) exports under the Agreement on Textiles and Clothing (ATC) were also updated by Joseph Francois and Julia Wörz. For 2007, these ETEs were not relevant for all countries except China, since the ATC quotas had been completely phased out in 2005. For China, we included fractions of ETE's implicit in the rate of increase in China's quotas through the year 2007. In GTAP 10, for 2011 and 2014, agricultural export subsidy data was updated by Jayson Beckman of ERS/USDA, who collected notifications to the World Trade Organization (WTO) and for EU members, draw on information from the European Agricultural Guarantee Fund (EAGF) provided by Alessandro Antimiani of DG Trade/European Commission. For GTAP 11 and 12, Kajode Ajewole of ERS/USDA updated the notifications to the WTO for 2017 and 2019, respectively. Since 2023 information was not available, Jayson Beckman of ERS/USDA submitted the 2022 for Korea, Morocco, and Türkiye.

One point that needs to be strongly emphasized for users of the GTAP Data Base for trade policy analysis is that the protection data supplied in GTAP is intended to represent a starting point for analysis. Any researcher using GTAP to conduct analysis of a specific policy liberalization scenario must scrutinize these data carefully for the focus countries in their analysis. In many cases, some adjustment will be required to reflect improved information that is often available from country-specific sources. Having ascertained cases where the protection information in GTAP must be altered to better reflect reality, the user can take advantage of the ALTERNAX program to make these changes. This program is documented in GTAP Technical Paper No. 12 by Malcolm 1998, and it has been made into an easy-to-use feature of the RunGTAP software authored by Mark Horridge (2001). Users may also want to update the 2023 policy environment as reflected in the GTAP 12 Data Base to a more recent base year (e.g., 2024). This is most naturally done via a simulation in which the intervening policy reforms are implemented, using the model to predict what the resulting world economy would look like. This can be a useful starting point for forward-looking policy analysis.

Another important area of recent development for the GTAP Data Base has been the incorporation of increasingly detailed data on the physical energy flows underlying the GTAP Data Base. These data have been obtained from the International Energy Agency (IEA). The energy volume data were first introduced in an interim version following release of GTAP 4 – nicknamed GTAP 4E. This was widely used by those conducting research on climate change policy. The process of reconciling energy volume and price data with the value flows in the GTAP Data Base has proven to be a challenging task. There remained substantial room for improvement after the release of GTAP 4E. The Center invested considerable effort in the energy data base and associated

programs in GTAP 6, with major contributions by Jean-Marc Burniaux, Huey-Lin Lee, and Robert McDougall. These improvements continued into version 8, with an update of the energy price and volumes data undertaken by Robert McDougall and Jevgenijs Steinbuks and subsequently used in version 9. An updated energy data treatment extends an approach first applied in the construction of the GTAP version 10 (McDougall and Chepeliev 2021) and several important modifications are introduced to the new treatment of energy data since GTAP 11 by Maksym Chepeliev, in part due to changes to the accounting of CO₂ emissions (Chepeliev 2022) but also by complementing the extended International Energy Agency (IEA) energy balances, with energy balances reported by the UN (UN 2021). Furthermore, we complement bilateral energy trade flows using data from British Petroleum (BP 2022), the Energy Institute (EI 2025), and the statistical office of the European Union Eurostat 2022.

In addition, fossil-fuel consumption subsidies based on estimates from the International Monetary Fund (IMF) and IEA data are now integrated in the standard database following an approach developed in Chepeliev, McDougall, and Mensbrugghe 2018.

Labor payments in the GTAP Data Base were disaggregated in version 4 using a methodology developed in Liu et al. 1998. In version 9 we re-examined these skill splits, with inputs from Weingarden and Tsigas 2010 based on the International Labour Organization (ILO), and finalized five different skill levels/categories for the standard version of GTAP Data Base. These skill levels are (with GTAP acronyms employed in the data base files, shown in parentheses):

1. Technically skilled professionals (tech_aspros)
2. Clerks (clerks)
3. Service and Shop Floor workers (service_shop)
4. Officers and managerial Professionals (off_mgr_pros)
5. Agricultural and other low-skilled workers (ag_othlowsk)

Since version 10, agricultural disaggregation has been updated using a dataset developed by Everett Peterson. In addition, several emissions data sets are being produced in the Center for Global Trade Analysis that are becoming more widely used. These are GTAP consistent estimates of process CO₂ emissions, non-CO₂ emissions and air pollutants, which add to the already standard CO₂ emissions from fossil fuel combustion.

Since version 10, the database is available in a format that is consistent with the GTAP 7 Model (Corong et al. 2017).⁶ Among some of the features, the data base is transformed to represent flows at basic and producer prices. The data base also includes a MAKE matrix that allows for industries to produce multiple products and for products to be produced by multiple industries. Following the new model, investment expenditures are separately identified in the database as other final demands along with private and government expenditures. There are header arrays for investment products domestically produced or imported, at basic and/or producer prices.

⁶Previous format, consistent with GTAP 6 Model in Hertel 1997 is also available.

1.3 Future Development of the Data Base

The future developments of the data base are heavily influenced by the GTAP Advisory Board. This group is made up of representatives from each of the agencies in the GTAP Consortium. Their continued guidance ensures that the data base will evolve to meet the changing needs of policy makers concerned about global economic issues. As noted above, quality assurance is a top priority. Nonetheless, there are a few areas where we anticipate significant breakthroughs in the not-too-distant future.

In GTAP 12 we continue to expand the time dimension, producing a time-series dataset with consistent sources and methodologies across years. Currently comprising six benchmark years, from 2004 to 2023, this task is challenging, as it involves revisiting each of the prior years with up to date methods and data sources (since these both change over time). Nevertheless, we will continue to extend the time-series GTAP data base drawing from our experience in producing the current version.

The geographical coverage of the GTAP 12 Data Base has increased to 163 regions—145 individual countries and 18 composite regions—with the addition of 4 new countries from Africa. Nineteen updated IO tables are also included thanks to the contributed datasets from a large network of individuals, GTAP Board member agencies, and institutions from around the world. Increasing the representation of countries and sectors in GTAP depends on data availability. Also, in order to improve the time-series dimension of the Data Base, continuous development and contribution of IOTs is critical in order to capture structural changes over time, since the historical years are matched with the closest IOT to that year. For historical reference years (i.e., 2004, 2007, 2011, 2014, and 2017), we rebuild the GTAP Data Base with the latest methodologies and updated inputs. For example, the 2017 reference year available in both GTAP Data Base version 11 and 12 will show differences because of services trade data and energy data have been updated, or perhaps because the IOT was updated/revised in the latest version.

Further improvements of the GTAP Data Base are also influenced by the quality and availability of international data sources. Our objective is to reconcile available information, with the primary aim of improving initial country data to meet the requirements of global economic modeling. The snapshot of the world economy that we have constructed can and should be extended to better meet the needs of research and policy objectives. Greater emphasis can be placed on a particular country, to then perform sub-regional modeling or reflect more recent trends.

Since GTAP 11, the services trade data now rely on new sources (Liberatore and Wettstein 2021 and Wettstein et al. 2021). If these data sources are not maintained, we would need to rely on a different source for GTAP 13. These and other similar instances highlight the particular importance of consistent maintenance and regular updates of the key data sources developed by statistical agencies and other agencies around the world.

One of the key features of the GTAP Data Base includes the reconciliation and merging multiple datasets in an attempt to provide an internally consistent representation of global economic flows. In this regard, we are constantly exploring new datasets that can be used to com-

plement current procedures. Similar to our targeting of agricultural production, we are considering the use of statistics by the UN Industrial Development Organization to target the production of manufactured goods in GTAP 13. This would allow us to provide a better representation of output across these sectors of the economy. The GTAP Data Base is a dynamic entity which is evolving in response to the needs and support of individual users as well as public agencies with an interest in international trade, natural resources and the environment. We encourage you to become involved in this network, subscribing to our discussion list, possibly attending the short course, and using this data base. We look forward to your feedback!

1.4 Outline of this Document

For GTAP 12 citation purposes, please use the open access article on the GTAP 12 Data Base, available in the Journal of Global Economic Analysis. The detailed documentation for the GTAP 12 Data Base is supplied on the GTAP website. This full documentation is made up of three parts. Part I, comprising Chapters 1 – 3, provides an overview of the GTAP 12 Data Base and highlights new features. Part II, Chapters 4 - 6 present three data distributions, two of these are utilities used to aggregate and/or view the data.

Part III begins with a detailed documentation of the macroeconomic data (Chapter 7) and new input- output tables (Chapter 8) for the GTAP 12 Data Base. Chapter 8 is made up of sub-chapters documenting the methodology used by the I-O table contributors in contributing I-O data. Chapter 9 then discuss other important aspects of the domestic data base construction procedures undertaken by the Center, including: the supplementary food and agricultural data base, methods used for disaggregating source tables, supplementary government consumption and procedures for building the composite regions⁷ which cover the “rest of” regions in the world economy.

The trade data base is documented in Chapter 10, including the implementation of bilateral trade margins by mode and services trade flows. Protection data is another large topic, covered in the multi-part Chapter 11. This includes discussion of the tariff data, as well as protection data for agriculture in particular export subsidies and domestic support.⁸

Chapter 12 covers the sources and procedures used to build the energy data base that accompanies the GTAP 12 Data Base. While only volume data are supplied, price data were also required in order to reconcile these volume data with GTAP’s value-based energy expenditures. In Chapter 13, the emissions associated with GTAP are presented, this include CO₂ from fossil fuel combustion, complementary Greenhouse Gas Emissions Database, which include CO₂ from non-fossil fuel combustion activities, and three types of non-CO₂ greenhouse gas (GHGs) – methane (CH₄), nitrous oxide (N₂O) and the group of fluorinated gases (F-gases). Chapter 14 documents the methodology used to derive labor splits for all sectors, and the primary factor shares for agriculture and for the natural resource-based industries; and Chapter 15 the methodology for estimating income taxes. Chapter 16 describes how the GTAP behavioral parameters were obtained. Chap-

⁷The composite regions aggregate countries for which we do not yet have an I-O table.

⁸Includes export tax equivalents associated with the Agreement on Textiles and Clothing for the 2004 dataset.

ters 17 then summarizes the procedure for updating the regional input-output tables to each of the reference years and assembling the full GTAP Data Base.

References

- BP (2022). “Statistical Review of World Energy”. In: June 2022. British Petroleum. URL: <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>.
- Chepeliev, Maksym (2022). “An Energy Data Base for GTAP”. In: *Global Trade, Assistance, and Production: The GTAP 10 Data Base*. Ed. by Angel Aguiar. Department of Agricultural Economics, Purdue University, West Lafayette, IN. Chap. 11.
- Chepeliev, Maksym, Robert McDougall, and Dominique van der Mensbrugghe (2018). “Including Fossil-fuel Consumption Subsidies in the GTAP Data Base”. In: *Journal of Global Economic Analysis* 3.1, pp. 84–121. DOI: 10.21642/JGEA.030102AF. URL: <https://www.jgea.org/ojs/index.php/jgea/article/view/60>.
- Corong, Erwin et al. (June 2017). “The Standard GTAP Model, Version 7”. In: *Journal of Global Economic Analysis* 2.1, pp. 1–119. DOI: 10.21642/JGEA.020101AF. URL: <https://jgea.org/resources/jgea/ojs/index.php/jgea/article/view/47>.
- EI (2025). “Statistical Review of World Energy 2025”. In: Energy Institute, London. URL: <https://www.energyinst.org/statistical-review/home>.
- Eurostat (2022). “Energy statistics - quantities, annual data”. In: Eurostat. URL: https://ec.europa.eu/eurostat/databrowser/explore/all/envir?lang=en&subtheme=nrg.nrg_quant.nrg_quanta&display=list&sort=category&extractionId=NRG_TI_GAS__custom_2160572.
- Gehlhar, Mark (1996). *Reconciling Bilateral Trade Data for Use in GTAP*. GTAP Technical Paper 10. Department of Agricultural Economics, Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). URL: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=313.
- Hertel, T.W., ed. (1997). *Global Trade Analysis: Modeling and Applications*. United Nations, New York: Cambridge University Press.
- Lejour, Arjan, Nico van Leeuwen, and Robert McDougall (2010). “Services Trade Data”. In: *Global Trade, Assistance, and Production: The GTAP 7 Data Base*. Ed. by Badri Narayanan G. and Terrie L. Walmsley. Department of Agricultural Economics, Purdue University, West Lafayette, IN. Chap. Data. URL: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=3397.
- Liberatore, Antonella and Steen Wettstein (2021). *The OECD-WTO Balanced Trade in Services database (BPM6 edition)*. OECD-WTO. URL: https://www.wto.org/english/res_e/statis_e/daily_update_e/OECD-WTO_Batis_methodology_BPM6.pdf.
- Liu, Jing et al. (1998). *Disaggregating Labor Payments by Skill Level in GTAP*. GTAP Technical Paper 11. Department of Agricultural Economics, Purdue University, West Lafayette, IN:

- Global Trade Analysis Project (GTAP). URL: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=314.
- Malcolm, Gerard (1998). *Adjusting Tax Rates in the GTAP Data Base*. GTAP Technical Paper 12. Department of Agricultural Economics, Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). URL: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=315.
- Martin, Will and L.A. Winters, eds. (1996). *The Uruguay Round and the Developing Countries*. New York: Cambridge University Press.
- McDougall, Robert (2002). “Services Trade Data”. In: *Global Trade, Assistance, and Production: The GTAP 5 Data Base*. Ed. by Betina V. Dimaranan and Robert A. McDougall. Department of Agricultural Economics, Purdue University, West Lafayette, IN. Chap. Data. URL: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=770.
- McDougall, Robert and Maksym Chepeliev (2021). “An Energy Data Base for GTAP”. In: *Global Trade, Assistance, and Production: The GTAP 10 Data Base*. Ed. by Angel Aguiar. Department of Agricultural Economics, Purdue University, West Lafayette, IN. Chap. 11. URL: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=6437.
- McDougall, Robert and Jan Hagemeyer (2006). “Services Trade Data”. In: *Global Trade, Assistance, and Production: The GTAP 6 Data Base*. Ed. by Betina V. Dimaranan. Department of Agricultural Economics, Purdue University, West Lafayette, IN. Chap. Data. URL: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=1945.
- Nuno-Ledesma, Jose and Nelson Villoria (2019). “Estimating International Trade Margins Shares by Mode of Transport for the GTAP Data Base”. In: *Journal of Global Economic Analysis* 4.1, pp. 28–49. DOI: 10.21642/JGEA.040102AF. URL: <https://jgea.org/ojs/index.php/jgea/article/view/73>.
- UN (2021). *2018 Energy Balances*. 2018th ed. United Nations. URL: <https://www.un-ilibrary.org/content/books/9789210055246>.
- Weingarden, Alison and Marinos Tsigas (2010). *Labor Statistics for the GTAP Database*. 13th Annual Conference on Global Economic Analysis, Penang, Malaysia. Center for Global Trade Analysis, Purdue University, West Lafayette, Indiana: Global Trade Analysis Project (GTAP).
- Wettstein, Steen et al. (2021). *A GLOBAL TRADE IN SERVICES DATA SET BY SECTOR AND BY MODE OF SUPPLY*. WTO. URL: https://www.wto.org/english/res_e/statis_e/daily_update_e/Tismos_methodology.pdf.