

Chapter 6

The GTAP Data Base for GAMS Users

*Dominique van der Mensbrugghe, Erwin Corong,
Maksym Chepeliev, and Angel Aguiar*

This chapter describes the content of the GTAP Data Base, which is being provided in GDX format for GAMS-based GTAP models or other downstream processes that use the GDX format. The new content, initially using GEMPACK's HAR formatted files, was developed for Version 7 of the GTAP model (Corong et al. 2017). It incorporates some key changes to the GTAP model, notably the introduction of a make matrix that allows for explicit differentiation between production activities and the supply of commodities among other changes mentioned in detail below. The new format has been tested with versions 9 (Aguiar, Narayanan, and McDougall 2016), 10 (Aguiar, Chepeliev, Corong, McDougall, et al. 2019) and 11 (Aguiar, Chepeliev, Corong, and Mensbrugghe 2022). And it has also been coupled with the standard GTAP Model in GAMS, Version 7 (van der Mensbrugghe 2018).

The GDX containers are available for the fully disaggregated GTAP Data Base. The following is a list of files that are included and discussed in subsequent sections:

- GSDFDAT.GDX
- GSDFPAR.GDX
- GSDFVOLE.GDX
- GSDFEMISS.GDX
- GSDFNCO2.GDX

Other GTAP extensions like GTAP-Power (Peters 2016 and Chepeliev 2020c) provide the same files except for the additional sectors representing electricity. For GTAP-E however, the distribution only includes the parameters file (GSDFPAR), for which contents are distinct from the standard distribution. The data flows and emissions are not included to avoid redundancy since these are consistent with standard GTAP distribution.

6.1 Core Data file (GSDFDAT)

Table 6.1 provides the full list of sets and value-flows contained in the main GTAP Data Base. The major changes that make this version consistent with GTAP version 7 model are:

- The letter ‘M’ has replaced the letter ‘I’ for import matrices/vectors. The letter ‘I’ is now used for the investment vectors.
- The accounting framework has been converted from ‘market/agent’ (M/A) prices to ‘basic/purchaser’ (B/P) prices. This implies that all tax rates (subsidy rates) are positive (negative). This impacts in particular production and export taxes that were negative in the former accounting system.
- The former ‘CGDS’ *activity* has been extracted from firms’ expenditures on goods and services and a new set of vectors have been created for investment expenditures that mirrors the vectors used for private and government expenditures.
- Income taxes on factor revenues have been made activity-specific. This currently has no impact as the ‘build’ procedure continues to assume uniformity across income source activity.
- The database includes a new set of matrices that represent the make matrix that describes the supply of commodities by source activity. The standard database continues to assume diagonality, i.e., a 1-to-1 correspondence between activities and commodities, but future releases may allow for non-diagonality (for example biofuels, vegetable oils, power, etc.). The production tax is applied to the make matrix, i.e., it can be activity- and commodity-specific.

Table 6.1: Contents of Core Data file (GSDFDAT)

Name	Former name	Description	Note
REG	REG	Set of regions	
ACTS	PROD_COMM	Set of activities	This used to include CGDS—investment expenditures are now represented by their own vectors
COMM	TRAD_COMM	Set of commodities	This used to be a subset of PROD_COMM, excluding CGDS.
ENDW	ENDW_COMM	Set of endowments	
MARG	MARG_COMM	Set of margin commodities (a subset of COMM)	
BOPF	N/A	Set of balance of payments flows	

Continued ...

...Continued

Name	Former name	Description	Note
VDFB (COMM, ACTS, REG)	VDFM	Purchases of domestic commodities by firms at basic prices	
VDFP (COMM, ACTS, REG)	VDFA	Purchases of domestic commodities by firms at purchasers' prices	
VMFB (COMM, ACTS, REG)	VIFM	Purchases of imported commodities by firms at basic prices	
VMFP (COMM, ACTS, REG)	VIFA	Purchases of imported commodities by firms at purchasers' prices	
VDPB (COMM, REG)	VDPM	Purchases of domestic commodities by private agents at basic prices	
VDPP (COMM, REG)	VDPA	Purchases of domestic commodities by private agents at purchasers' prices	
VMPB (COMM, REG)	VIPM	Purchases of imported commodities by private household agents at basic prices	
VMPP (COMM, REG)	VIPA	Purchases of imported commodities by private household agents at purchasers' prices	
VDGB (COMM, REG)	VDGM	Purchases of domestic commodities by government agents at basic prices	
VDGP (COMM, REG)	VDGA	Purchases of domestic commodities by government agents at purchasers' prices	
VMGB (COMM, REG)	VIGM	Purchases of imported commodities by government agents at basic prices	
VMGP (COMM, REG)	VIGA	Purchases of imported commodities by government agents at purchasers' prices	
VDIB (COMM, REG)	VDFM	Purchases of domestic commodities by investment agents at basic prices	The investment vectors are new. They have been extracted from the 'CGDS' activity from the former database format.
VDIP (COMM, REG)	VDFA	Purchases of domestic commodities by investment agents at purchasers' prices	
VMIB (COMM, REG)	VIFM	Purchases of imported commodities by investment agents at basic prices	
VMIP (COMM, REG)	VIFA	Purchases of imported commodities by investment agents at purchasers' prices	

Continued ...

... Continued

Name	Former name	Description	Note
EVFB (ENDW, ACTS, REG)	VFM	Purchases of endowments by firms at basic prices	
EVFP (ENDW, ACTS, REG)	EVFA	Purchases of endowments by firms at purchasers' prices	
EVOS (ENDW, ACTS, REG)	EVOA	Post income tax factor income	In the former database, the income tax on factor income was assumed uniform across source activities.
VXSB (COMM, REG, REG)	VXMD	The value of bilateral trade at basic prices	
VFOB (COMM, ACTS, REG)	VXWD	The value of bilateral trade at FOB prices	
VCIF (COMM, ACTS, REG)	VIWS	The value of bilateral trade at CIF prices	
VMSB (COMM, ACTS, REG)	VIMS	The value of bilateral trade at basic (tariff-inclusive) prices	
VST (MARG, REG)	VST	The value of regional exports of international trade & transport services	
VTWR (MARG, COMM, REG, REG)	VTWR	Modal distribution of international trade & transport services	
SAVE (REG)	SAVE	Domestic savings	
VBPF (REG, BOPF)	N/A	Balance of payments flows	
VDEP (REG)	VDEP	Value of depreciation	
VKB (REG)	VKB	Value of capital stock	
POP (REG)	POP	Population in millions	
MAKS (COMM, ACTS, REG)		Make matrix at production prices	The new database allows for non-diagonal make matrices, though the default database assumes diagonality.
MAKB (COMM, ACTS, REG)		Make matrix at basic prices	The former production tax is implemented at this stage and is assumed to be activity and commodity specific.
PTAX (COMM, ACTS, REG)		Production tax revenues	
VOSB (COMM, REG)		The value of commodity supply pre-production tax.	

Unless indicated otherwise, all values are in current millions of U.S. dollars using prices and market exchange rates for the relevant reference year.¹

The Appendix in Chapter 2 provides the standard set of commodities in the GTAP Data

¹I-O tables being of different reference years are an exception, because they are being updated from their local currency to U.S. dollars using macroeconomic data, see Chapter 7.

Base. In the standard database, these are the same as the set of activities. The standard commodity subsets are:

- MARG: OTP, WTP, ATP
- ERG : COA, OIL, GAS, P_C, ELY, GDT
- FUEL: COA, OIL, GAS, P_C, GDT

6.2 *Parameters file (GSDFPAR)*

Table 6.2 provides the full list of parameters that are included in the standard GTAP release. All files include any set definitions embedded in the parameters—this makes the files self-contained even if it duplicates information from other files. Users are of course free to ignore the sets if already defined in the code. The major changes in the parameter database include:

- The new GTAP model has added a new CES nest in the production structure—intermediate goods are combined together in a composite bundle. This allows for substitution across intermediate goods.
- A number of former parameters have been given additional dimensions, which allows for greater heterogeneity—for example, the Armington elasticities are now indexed by region as well as by commodity.
- The new make matrix requires additional parameters. A CET specification is allowed to allocate domestic output across commodities (e.g. vegetable oils). A CES specification is used to bundle 'similar' output from various activities (e.g. electricity).
- Government and investment expenditures are specified using a generic CES expenditure function. The same for the bundling of margin services across source regions.
- In the case of perfect transformation, a value of infinity ('+INF') is used rather than using subsets.

Table 6.2: Contents of GTAP Parameter File (GSDFPAR)

Name	Former name	Description	Note
REG	REG	Set of regions	
ACTS	PROD_COMM	Set of activities	This used to include CGDS—investment expenditures are now represented by their own vectors
COMM	TRAD_COMM	Set of commodities	This used to be a subset of PROD_COMM, excluding CGDS.
ENDW	ENDW_COMM	Set of endowments	
MARG	MARG_COMM	Set of margin commodities (a subset of COMM)	
ESUBT (ACTS, REG)	ESUBT	Elasticity of substitution between intermediate inputs and value added bundle	There was no separate aggregate intermediate demand bundle in the previous version. The elasticity has also been made region specific. It defaults to zero.
ESUBC (ACTS, REG)		This is a new elasticity that allows for substitution across intermediate inputs within the intermediate demand bundle. It defaults to zero.	
ESUBVA (ACTS, REG)	ESUBVA	Substitution across endowments in the value added bundle.	It is now region and activity specific.
ETRAQ (ACTS, REG)		Transformation elasticity across commodities for output allocation	This is new to allow for non-diagonal make matrices. It defaults to 5 but is irrelevant in the case of diagonality.
ESUBQ (ACTS, REG)		Substitution across activities in commodity bundle	This is new to allow for non-diagonal make matrices. It defaults to infinity but is irrelevant in the case of diagonality.
INCPAR (COMM, REG)	INCPAR	Expansion parameter for CDE utility function	
SUBPAR (COMM, REG)	SUBPAR	Substitution parameter for CDE utility function	
ESUBG (REG)		Government expenditure substitution elasticity	In the former database/model, government expenditures were by default allocated using a Cobb-Douglas specification. The new model allows for a generic CES function, though defaults to a value of 1.
ESUBI (REG)	ESUBT	Investment expenditure substitution elasticity	In the former database/model, the investment expenditure elasticity was governed by ESUBT and was 0 by default. A generic CES expenditure function is implemented with a default elasticity of 0.
ESUBD (COMM, REG)	ESUBD	Top level Armington elasticity, substitution between domestic goods and import bundle	It is now region-specific as well as commodity-specific. N.B. It could also be made agent-specific as this choice is done at the agent level.
ESUBM (COMM, REG)	ESUBM	Second level Armington elasticity, substitution of imports across source regions	It is now region-specific as well as commodity-specific.
ESUBS (MARG)		Substitution across source regions for trade & transport services	This is new and allows for a generic CES specification. The default elasticity is 1 as in the past implementation.
ETRAE (ENDW, REG)	ETRAE	Transformation (or mobility) elasticity of factors across activities.	The GAMS implementation replaces perfect mobility with a transformation elasticity of infinity. This precludes the need for the ENDWS_COMM that was a subset for the 'sluggish' factors.
RORFLEX (REG)	RORFLEX	Flexibility of expected net rate of return with respect to investment	

6.3 Energy Volumes file (GSDFVOLE)

Table 6.3 provides the full list of value-flows that are included in the energy volumes dataset. The values are in millions of tons of oil equivalent (MTOE). There are two differences with respect to the previous format: (1) energy used in investment expenditures has been extracted from the firms' matrices; and (2) The 'I' identifier for imports has been replaced by 'M'.

Table 6.3: Contents of Energy Volumes File

Name	Former name	Description	Note
REG	REG	Set of regions	
ACTS	PROD_COMM	Set of activities	This used to include CGDS—investment expenditures are now represented by their own vectors
COMM	TRAD_COMM	Set of commodities	This used to be a subset of PROD_COMM, excluding CGDS.
ERG	ERG_COMM	Set of energy commodities (a subset of COMM)	
EDF (ERG, ACTS, REG)	EDF	Use of domestic energy by firms	
EMF (ERG, ACTS, REG)	EIF	Use of imported energy by firms	
EDP (ERG, REG)	EDP	Use of domestic energy by private household agents	
EMP (ERG, REG)	EIP	Use of imported energy by private household agents	
EDG (ERG, REG)	EDG	Use of domestic energy by government agents	
EMG (ERG, REG)	EIG	Use of imported energy by government agents	
EDI (ERG, REG)		Use of domestic energy by investment agents	These flows used to be incorporated in energy use by firms.
EMI (ERG, REG)		Use of imported energy by investment agents	These flows used to be incorporated in energy use by firms.
EXI (ERG, REG, REG)	EXI	Energy exports	

6.4 CO₂ Emissions file (GSDFEMISS)

Table 6.4 provides the information included in the CO₂ database. The values are in millions of metric tons of CO₂. There are two differences with respect to the previous format: (1) emissions from investment expenditures have been extracted from the firms' matrices; and (2) The 'I' identifier for imports has been replaced by 'M'.

Table 6.4: Contents of CO₂ Emissions File

Name	Former name	Description	Note
REG	REG	Set of regions	
ACTS	PROD_COMM	Set of activities	This used to include CGDS—investment expenditures are now represented by their own vectors
COMM	TRAD_COMM	Set of commodities	This used to be a subset of PROD_COMM, excluding CGDS.
ERG	ERG_COMM	Set of energy commodities (a subset of COMM)	
FUEL	FUEL_COMM	Set of fuel commodities (a subset of ERG)	
MDF (FUEL, ACTS, REG)	MDF	CO ₂ emissions from combustion of domestic energy by firms	
MMF (FUEL, ACTS, REG)	MIF	CO ₂ emissions from combustion of imported energy by firms	
MDP (FUEL, REG)	MDP	CO ₂ emissions from combustion of domestic energy by private household agents	
MMP (FUEL, REG)	MIP	CO ₂ emissions from combustion of imported energy by private household agents	
MDG (FUEL, REG)	MDG	CO ₂ emissions from combustion of domestic energy by government agents	
MMG (FUEL, REG)	MIG	CO ₂ emissions from combustion of imported energy by government agents	
MDI (FUEL, REG)		CO ₂ emissions from combustion of domestic energy by investment agents	These flows used to be incorporated in energy-based emissions by firms.
MMI (FUEL, REG)		CO ₂ emissions from combustion of domestic energy by investment agents	These flows used to be incorporated in energy-based emissions by firms.

6.5 Other emissions (GSDFNCO2)

Table 6.5 provides the full list of matrices in the pollutants auxiliary data set. The GAMS version of the file merges, when available, CO₂ emissions associated with process emissions (i.e., not combustion), the non-CO₂ greenhouse gases and the non-greenhouse gas air emissions.² The two sets of air emissions have largely the same functionality from the point of view of the geometry of the data sets and thus it is practical to merge them. Differences in functionality are noted below.

Table 6.5: Non-CO₂ Emissions Database

Name	Former name	Description	Note
AR		IPCC assessment report identifier for GWPs reporting	
COMM	TRAD_COMM	Set of commodities	This used to be a subset of PROD_COMM, excluding CGDS.
ENDW	ENDW_COMM	Set of endowments	
ACTS	PROD_COMM	Set of activities	This used to include CGDS—investment expenditures are now represented by their own vectors
REG	REG	Set of regions	
EM		Full set of air emissions, including CO ₂	This is not part of the GTAP Data Base but has been added for convenience.
EMN		Full set of air emissions excluding CO ₂	This is not part of the GTAP Data Base but has been added for convenience.
CO2		Singleton set for CO ₂	
GHG		Greenhouse gas emissions, a subset of EM	This is not part of the GTAP Data Base but has been added for convenience.
FGAS		Set of the 25 fluorinated gases	This is not part of the GTAP Data Base but has been added for convenience.
NGHG		Set of non-greenhouse gases, a subset of EM	This is not part of the GTAP Data Base but has been added for convenience.
LU_CAT		Set of land use categories for GHG emissions reporting	
LU_SUBCAT		Set of land use subcategories for GHG emissions reporting	
GWP (GHG, REG, AR)		GHG global warming potential across IPCC reports	
EMI_IO (EMN, COMM, ACTS, REG)	NC_TRAD	Non-CO ₂ emissions linked to fossil fuel combustion in production, mmt	
EMI_ENDW (EMN, ENDW, ACTS, REG)	NC_ENDW	Non-CO ₂ emissions linked to factor use, mmt	
EMI_QO (EMN, ACTS, REG)	NC_QO	Non-CO ₂ emissions linked to output, mmt	
EMI_HH (EMN, COMM, REG)	NC_HH	Non-CO ₂ emissions linked to household consumption, mmt	
EMI_IOP (GHG, COMM, ACTS, REG)	NC_TRAD	GHG process emissions linked to intermediate demand, mmt	
EMI_LU (EM, LU_CAT, LU_SUBCAT, REG)		GHG emissions from land use activities, mmt	

The set of non-CO₂ greenhouse gases includes:³

²Chepeliev 2020b describes the construction of the non-CO₂ greenhouse gas database for V10A of the GTAP Data Base and Chepeliev 2020a describes the construction of the non-greenhouse gas air emissions data set. N.B. The latter has had different content across versions.

³Adapted from <https://www.epa.gov/ghgemissions/overview-greenhouse-gases>.

- Methane (CH_4 or CH_4): Methane is emitted during the production and transport of coal, natural gas, and oil. Methane emissions also result from livestock and other agricultural practices (e.g. rice cultivation) and by the decay of organic waste in municipal solid waste landfills.
- Nitrous oxide (N_2O or N_2O): Nitrous oxide is emitted during agricultural and industrial activities, combustion of fossil fuels and solid waste, as well as during treatment of waste water.
- 25 Fluorinated gases (FGAS): Hydrofluorocarbons, perfluorocarbons, sulfur hexafluoride, and nitrogen trifluoride are synthetic, powerful greenhouse gases that are emitted from a variety of industrial processes. Fluorinated gases are sometimes used as substitutes for stratospheric ozone-depleting substances (e.g., chlorofluorocarbons, hydrochlorofluorocarbons, and halons). These gases are typically emitted in smaller quantities, but because they are potent greenhouse gases, they are sometimes referred to as High Global Warming Potential gases ("High GWP gases").

The set of non-greenhouse gas air emissions include:

- Black carbon (BC): Black carbon is the sooty black material emitted from gas and diesel engines, coal-fired power plants, and other sources. It comprises a significant portion of particulate matter or PM, which is an air pollutant.
- Carbon monoxide (CO): Carbon monoxide is a colorless, practically odorless, and tasteless gas or liquid. It results from incomplete oxidation of carbon in combustion.
- Ammonia (NH_3 or NH_3): About 80% of the ammonia produced in industry is used in agriculture as fertilizer. Ammonia is also used as a refrigerant gas, to purify water supplies, and in the manufacture of plastics, explosives, fabrics, pesticides, dyes and other chemicals.
- Non-methane volatile organic compounds (NMVOC): Non-methane volatile organic compounds are a large variety of chemically different compounds, such as benzene, ethanol, formaldehyde, cyclohexane, 1,1,1-trichloroethane or acetone. Essentially, NMVOCs are identical to volatile organic compounds (VOCs), but with methane excluded.
- Nitrogen oxides (NO_x or NO_x): In atmospheric chemistry, NO_x is a generic term for the nitrogen oxides that are most relevant for air pollution, namely nitric oxide (NO) and nitrogen dioxide (NO_2). These gases contribute to the formation of smog and acid rain, as well as affecting tropospheric ozone.
- Organic carbon (OC): Organic carbon aerosols are particulate aerosols formed by incomplete combustion of typically biomass.
- Particulate matter 10 (PM₁₀): PM₁₀ is particulate matter 10 micrometers or less in diameter. PM₁₀ and PM_{2.5} are produced from a wide range of industrial processes through

bulk material handling, combustion and minerals processing. The industries using these processes include brickworks, refineries, cement works, iron and steel making, quarrying, and fossil fuel power plants. Particulates are released from a wide range of diffuse sources. Examples include lawn mowing, wood stoves, fires, and wind generated dust, though this tends to be coarser. Vehicles will generate particulates either from direct emissions from the burning of fuels (especially diesel-powered vehicles) or from wear of tires.

- Particulate matter 2.5 (PM_{2.5}): PM_{2.5} is particulate matter 2.5 micrometers or less in diameter.
- Sulfur dioxide (SO₂ or SO₂): The largest source of SO₂ in the atmosphere is the burning of fossil fuels by power plants and other industrial facilities. Short-term exposures to SO₂ can harm the human respiratory system and make breathing difficult. People with asthma, particularly children, are sensitive to these effects of SO₂.

Unlike CO₂ emissions, which currently in the GTAP Data Base result from the combustion of fossil fuels, the non-CO₂ greenhouse gas emissions are driven by consumption (intermediate and final), factor use (e.g. methane emissions from livestock capital use or rice land use), and/or output (process-based emissions such as methane from landfills). N.B. The commodity-driven non-CO₂ greenhouse gas emissions are linked to the level of aggregate absorption and not differentiated by source (domestic vs. imported).

All gases are provided in physical units. Unlike previous versions, the file no longer includes the values in so-called CO₂-equivalent (or CO₂e) units, which takes into account the global warming potential (GWP) of the various GHGs. The GWP adjustment coefficients are provided as a separate parameter (GWP) and are currently available for three IPCC assessment reports (AR2, AR4, and AR5). The default in previous versions had been to use the AR4 values. Also note that the current version has 25 fluorinated gases, unlike earlier versions, which had a single aggregate value.

References

- Aguiar, Angel, Maksym Chepeliev, Erwin Corong, Robert McDougall, et al. (June 2019). “The GTAP Data Base: Version 10”. In: *Journal of Global Economic Analysis* 4.1, pp. 1–27. DOI: 10.21642/JGEA.040101AF.
- Aguiar, Angel, Maksym Chepeliev, Erwin Corong, and Dominique van der Mensbrugghe (Dec. 2022). “The Global Trade Analysis Project (GTAP) Data Base: Version 11”. In: *Journal of Global Economic Analysis* 7.2. URL: <https://jgea.org/ojs/index.php/jgea/article/view/181>.
- Aguiar, Angel, Badri Narayanan, and Robert McDougall (June 2016). “An Overview of the GTAP 9 Data Base”. In: *Journal of Global Economic Analysis* 1.1, pp. 181–208. DOI: 10.21642/JGEA.010103AF.

- Chepeliev, Maksym (2020a). *Development of the Air Pollution Database for the GTAP Data Base Version 10A*. GTAP Research Memorandum 33. 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=6163.
- (2020b). *Development of the Non-CO₂ GHG Emissions Database for the GTAP Data Base Version 10A*. GTAP Research Memorandum 32. 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=5993.
- (2020c). *GTAP-Power 10 Data Base: A Technical Note*. GTAP Research Memorandum 31. Department of Agricultural Economics, Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). URL: <https://doi.org/10.21642/GTAP.RM31>.
- 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>.
- Peters, Jeffrey C. (Dec. 2016). “GTAP-E-Power: An Electricity-detailed Economy-wide Model”. In: *Journal of Global Economic Analysis* 1.2, pp. 156–187. DOI: 10.21642/JGEA.010204AF. URL: <https://jgea.org/ojs/index.php/jgea/article/view/27>.
- van der Mensbrugghe, Dominique (June 2018). “The Standard GTAP Model in GAMS, Version 7”. In: *Journal of Global Economic Analysis* 3.1, pp. 1–83. DOI: 10.21642/JGEA.030101AF. eprint: <https://jgea.org/resources/jgea/ojs/index.php/jgea/article/view/62>.