Chapter 12
An Energy Data Base for GTAP

Maksym Chepeliev[1]

12.1 Overview

For energy, as with agriculture, we prepare a special data set not just to supplement data from sector-generic sources (as we use the agricultural data set to disaggregate I-O tables) but to override them.[2] This special treatment arises from users' concern about apparent divergences of energy data in earlier GTAP releases from International Energy Agency (IEA) data (see, e.g., Babiker and Rutherford 1997). To address these concerns, we prepare an energy data set from IEA and other energy-oriented data, covering energy flows, prices, and taxes, and use it to modify the I-O, protection, and trade data.

Also special to energy is the inclusion in the GTAP data set not only of money value data but also of volume data. Supplementing the core money value flows, we provide a separate file of I-O and international trade flows measured in millions of tons of oil equivalent (MTOE).

In this chapter, we describe the preparation of the GTAP energy data set (EDS). In the course of constructing the trade component of that data set, we modify the GTAP trade data set. In the energy data set, we include tax data, that later procedures use instead of the main protection data set in handling energy sectors. We also include targets for money values of energy flows, used by later procedures to modify the I-O data.

Although, in the main, the energy data set overrides the I-O data, in a few cases we find it necessary to adjust energy data for compatibility with the I-O data, as explained in chapter 17. Furthermore, since the source data specific to the energy data set are incomplete, we draw on the GTAP protection, I-O, and trade data sets as supplementary data inputs into the construction of the energy data set.

Since our main source data do not support all the distinctions required for the GTAP sectoral classification, the energy data set uses its own special-purpose sectoral classifications. The

[1]Center for Global Trade Analysis, Dept. of Agricultural Economics, Purdue University. Email:mchepeli@purdue.edu
[2]This Chapter incorporates materials from McDougall and Chepeliev 2021, McDougall and Lee 2006, Burniaux and Truong 2002.

EDS industry classification (EDSIC) is a 22-sector aggregation of the GTAP sectoral classification, as listed in table 12.1. The EDS use classes include, besides intermediate usage by each of the EDS industries, just private consumption and exports; implicitly, we assume that government consumption and investment usage are zero. The EDS energy commodity classification (EDSECC) is a 5-sector aggregation of the 6 GTAP energy sectors: the GTAP sectors gas (primary gas production) and gdt (gas manufacture and distribution) map to the single EDSECC sector gas.

Table 12.1: Concordance between EDS Industries and GTAP Sectors

EDS industry		GTAP sector	
Code	Description	Code	Description
agr	Agriculture, forestry, fishing	pdr	Paddy rice
		wht	Wheat
		gro	Cereal grains n.e.c.
		v_f	Vegetables, fruit, nuts
		osd	Oil seeds
		c_b	Sugar cane, sugar beet
		pfb	Plant-based fibers
		ocr	Crops n.e.c.
		ctl	Bovine cattle, sheep and goats, horses
		oap	Animal products n.e.c.
		rmk	Raw milk
		wol	Wool, silk-worm cocoons
		frs	Forestry
		fsh	Fishing
coa	Coal	coa	Coal
oil	Oil	oil	Oil
gas	Gas	gas	Gas
		gdt	Gas manufacture, distribution
omn	Minerals n.e.c.	oxt	Other extraction
fpr	Food products, beverages	cmt	Bovine meat products
		omt	Meat products n.e.c.
		vol	Vegetable oils and fats
		mil	Dairy products
		pcr	Processed rice
		sgr	Sugar
		ofd	Food products n.e.c.
		b_t	Beverages and tobacco products
twl	Textiles, wearing apparel, leather	tex	Textiles
		wap	Wearing apparel
		lea	Leather products
lum	Wood products	lum	Wood products
ppp	Paper products, publishing	ppp	Paper products, publishing
p_c	Petroleum, coal products	p_c	Petroleum, coal products
crp	Chemical, rubber, plastic products	chm	Chemical products
		bph	Basic pharmaceutical products

Continued ...

...Continued

EDS industry		GTAP sector	
Code	Description	Code	Description
		rpp	Rubber and plastic products
nmm	Mineral products n.e.c.	nmm	Mineral products n.e.c.
i_s	Ferrous metals	i_s	Ferrous metals
nfm	Non-ferrous metals	nfm	Metals n.e.c.
ome	Machinery and equipment n.e.c.	fmp	Metal products
		ele	Computer, electronic and optical products
		eeq	Electrical equipment
		ome	Machinery and equipment n.e.c.
teq	Transport equipment	mvh	Motor vehicles and parts
		otn	Transport equipment n.e.c.
omf	Manufactures n.e.c.	omf	Manufactures n.e.c.
ely	Electricity	ely	Electricity
cns	Construction	cns	Construction
tpt	Transport	otp	Transport n.e.c.
		wtp	Water transport
		atp	Air transport
ser		wtr	Water
		trd	Trade
		afs	Accommodation, Food and service activities
		cmn	Communication
		ofi	Financial services n.e.c.
		ins	Insurance
		rsa	Real estate activities
		obs	Business services n.e.c.
		ros	Recreational and other services
		osg	Public administration and defense
		edu	Education
		hht	Human health and social work activities
		dwe	Dwellings

The energy data set generated by the energy module includes data for:

— quantity of energy usage, by energy commodity and energy use class,

— money value of energy usage, by energy commodity and energy use class,

— FOB and CIF values of trade in energy commodities, by GTAP commodity, source region, and destination region,

— energy output subsidy rates by GTAP industry, and

— powers of taxes on intermediate usage and private consumption of energy commodities, by GTAP commodity and industry.

In broad outline, we proceed by constructing from separate sources an energy volumes data set (section 12.2) and an energy price and tax rate data set (section 12.3). We then integrate these two data sets with each other and with separately constructed data for trade and protection (section 12.4).

12.2 Energy Volumes: Initial Construction

The main data source for energy volumes is the International Energy Agency's "Extended Energy Balances" (EEB) (IEA 2021b). This provides a wide country coverage (151 countries), rich sectoral detail, and a high degree of internal consistency and completeness. The energy balance framework, though different in externals from the input-output framework, translates readily into it. The industry sector classification relies heavily on the International Standard Industry Classification (ISIC), the same basis as the GTAP sectoral classification, facilitating matching to GTAP sectors. The main difficulties arise from mismatches between IEA energy sectors and GTAP sectors, joint production, and special categories such as transfers, distribution losses, and international marine bunkers.

The energy balances constitute a large array of energy flows, measured in kilotons of oil equivalent, indexed by year, country, flow, and product. The flow index includes both basic elements such as primary production, imports, and exports, and elements such as "public electricity plant" or "iron and steel" corresponding to I-O industries. The products correspond to I-O commodities. We extract data for all countries, flows, and products; for years we extract just the GTAP 11 Data Base reference years – 2004, 2007, 2011, 2014 and 2017.

As noted above, the EEB data set generally maintains a high level of internal consistency. Nevertheless, it does exhibit some inconsistencies. Based as it is on individual country reports, it does not attempt to enforce cross-country balance, so there is no match between global exports and global imports. It also contains some small logical inconsistencies; for example, it reports both that some natural gas is used in oil and gas extraction in Estonia, and that Estonia produces no oil or gas (it may be that these reports are consistent under EEB conventions, but under I-O conventions we cannot allow non-zero inputs with zero output; we treat industries with zero output as non-existent).

The EEB and GTAP take different approaches to recording trade. The EEB includes re-exports in its trade flows, GTAP excludes them.

To prepare the energy volume data for GTAP use we take the following steps:

1. map from the 151 non-overlapping EEB countries to the 251 standard GTAP countries;
2. map from EEB flows and products to EDS industries and commodities;
3. aggregate from the 251 standard GTAP countries to the 160 GTAP release 11 regions;
4. balance the energy flows for internal consistency;
5. remove re-exports; and
6. eliminate energy usage by non-existent industries.

The resultant energy volume data set records imports, production, intermediate usage, private consumption, and exports of energy, for 2004, 2007, 2011, 2014 and 2017, in millions of tons of oil equivalent.

It uses the standard GTAP regional classification, the EDS commodity classification, and, for intermediate usage, the EDS industry classification.

12.2.1 Mapping from EEB to GTAP countries

IEA energy balances report data for 151 individual countries out of 251 countries distinguished in the GTAP country list, therefore not explicitly covering 100 individual countries (some of these countries are not reported by the IEA at all, while others are reported aggregately within composite regions, e.g. Other Africa or Other Asia). Most of the individual countries not explicitly covered by the IEA are small island economies, however, there are some exceptions including Chad, Mali, Uganda and Afghanistan (Table 12.2). To improve the representation of the energy data in GTAP, we further complement the data construction process with information from the United Nations (UN) energy balances (UN 2020). The latter provides information for 214 individual countries but has more aggregate commodity and use coverage compared to the IEA energy balances (UN data reports 11 commodities and 39 transaction types). UN energy balances are used to gap-fill data for countries that are not represented in the IEA accounts as further discussed below.

In addition to 151 individual countries, the IEA EEBs report data for three composite groups. "Other Africa" groups 24 GTAP countries, "Other Asia" 15, and "Other Latin America" 25. Thus, the IEA covers a total of 215 countries (151 reported individually and an additional 61 reported within composite regions). There remain 40 GTAP countries that are not reported by the IEA (Table 12.2). When combined with data available in the UN, we find that out of 100 countries not explicitly reported by the IEA (absent or reported within composite regions), 64 are explicitly reported in the UN. Still, there are 31 countries that are not reported by neither IEA nor UN (Table 12.2).

IEA energy balances of the 151 individual countries are complemented with UN energy balances. For the case of countries reported within IEA composite regions, UN energy balances are scaled to match the aggregate energy flows of each corresponding composite region. IEA energy balances of the composite regions are then used to provide a disaggregated commodity and supply categories representation of the UN individual countries that are mapped to the IEA composite regions. To disaggregate the energy flows of individual countries reported by the UN but not covered by the IEA composite regions, we use the world aggregate energy balances from the IEA data. To the 31 GTAP countries not covered by the IEA or UN, we assign flows equal to the EEB's world energy flows, multiplied by their share in world GDP.

12.2.2 Mapping from EEB Flows and Products to EDS Sectors

Although the EEB classification of energy flows and products is much more detailed than GTAP's, its classification of non-energy sectors is less detailed. Furthermore, unlike GTAP, it does not recognize gas distribution as a separate activity. Hence arises the need for the special energy data set sectoral classifications, less detailed than the standard GTAP sectoral classification, described in section 12.1. For the most part, we treat the IEA EEB sectoral classifications as disaggregations of the EDS classifications. The exceptions fall into three classes. First, we discard some of the EEB sectors. These include sectors such as "statistical differences" that represent nothing in the real world, but are items of accounting convenience. We also choose to discard some EEB flows and products relating to energy production by predominantly non-energy sectors in the EDSIC. For instance, rather than treat the chemical industry as energy-producing, we discard the EEB flow "charcoal production plants" and the EEB product "charcoal". Second, some of the EEB flows belong

Table 12.2: Mapping between EEB and GTAP Countries: Countries Grouped in or Absent from the EEB

EEB	GTAP
Other Africa (IEA)	Burkina Faso, Burundi, Cabo Verde, Central African Republic, Chad, Comoros, Djibouti, Eswatini, Gambia, Guinea, Guinea-Bissau, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Rwanda, Sao Tome and Principe, Seychelles, Sierra Leone, Somalia, Uganda, Réunion
Other Asia (IEA)	Afghanistan, Bhutan, Cook Islands, Fiji, French Polynesia, Kiribati, China, Macao Special Administrative Region, Maldives, New Caledonia, Papua New Guinea, Samoa, Solomon Islands, Timor-Leste, Tonga, Vanuatu
Other Latin America (IEA)	Anguilla, Antigua and Barbuda, Aruba, Bahamas, Barbados, Belize, Bermuda, British Virgin Islands, Cayman Islands, Dominica, Falkland Islands (Malvinas), Grenada, Guyana, Montserrat, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Pierre and Miquelon, Saint Vincent and the Grenadines, Turks and Caicos Islands, Bonaire, Sint Eustatius and Saba, French Guiana, Guadeloupe, Martinique, Sint Maarten (Dutch part)
Absent in IEA	Åland Islands, Andorra, American Samoa, Antarctica, French Southern Territories, Saint Barthélemy, Bouvet Island, Cocos (Keeling) Islands, Christmas Island, Western Sahara, Faroe Islands, Micronesia (Federated States of), Guernsey, Guam, Heard Island and McDonald Islands, Isle of Man, British Indian Ocean Territory, Jersey, Liechtenstein, Saint Martin (French Part), Monaco, Marshall Islands, Northern Mariana Islands, Mayotte, Norfolk Island, Niue, Nauru, Pitcairn, Palau, Sark, South Georgia and the South Sandwich Islands, Saint Helena, Svalbard and Jan Mayen Islands, San Marino, Tokelau, Tuvalu, United States Minor Outlying Islands, Holy See, United States Virgin Islands, Wallis and Futuna Islands
Absent or reported within composite regions in IEA but explicitly reported in UN	Aruba, Afghanistan, Anguilla, Andorra, Antigua and Barbuda, Burundi, Bonaire, Sint Eustatius and Saba, Burkina Faso, Bahamas, Belize, Bermuda, Barbados, Bhutan, Central African Republic, Cook Islands, Comoros, Cabo Verde, Cayman Islands, Djibouti, Dominica, Fiji, Falkland Islands (Malvinas), Faroe Islands, Micronesia (Federated States of), Guinea, Gambia, Guinea-Bissau, Grenada, Kiribati, Saint Kitts and Nevis, Liberia, Saint Lucia, Lesotho, China, Macao Special Administrative Region, Maldives, Mauritania, Montserrat, Malawi, New Caledonia, Niue, Nauru, Palau, Papua New Guinea, French Polynesia, Rwanda, Saint Helena, Solomon Islands, Sierra Leone, Somalia, Saint Pierre and Miquelon, Sao Tome and Principe, Eswatini, Sint Maarten (Dutch part), Seychelles, Turks and Caicos Islands, Chad, Timor-Leste, Tonga, Tuvalu, Saint Vincent and the Grenadines, British Virgin Islands, Vanuatu, Wallis and Futuna Islands, Samoa
Absent in both IEA and UN	Åland Islands, American Samoa, Antarctica, French Southern Territories, Saint Barthélemy, Bouvet Island, Cocos (Keeling) Islands, Christmas Island, Western Sahara, Guernsey, Guam, Heard Island and McDonald Islands, Isle of Man, British Indian Ocean Territory, Jersey, Liechtenstein, Saint Martin (French Part), Monaco, Marshall Islands, Northern Mariana Islands, Mayotte, Norfolk Island, Pitcairn, Sark, South Georgia and the South Sandwich Islands, Svalbard and Jan Mayen Islands, San Marino, Tokelau, United States Minor Outlying Islands, Holy See, United States Virgin Islands

in the EDS but not in the intermediate usage block. These include flows such as production, exports, and imports. Third, some EEB flows combine uses that the EDS must separate, for example, the gas and crude oil industries, or the transport industry and private consumption. In particular:

— The EEB flow "oil and gas extraction" corresponds to the EDS industries "oil" and "gas". We assign all inflows of the EEB product "crude oil" into "oil and gas extraction" to the EDS industry "oil"; we assign all inflows of "natural gas" to "gas"; and we split all other inflows between "oil" and "gas" in proportion to production of "crude oil" and "natural gas".

— We apportion the EEB flow "non-specified energy sector" between all EDS energy-using sectors in proportion to their own energy use.

— The EEB flow "non-specified industry" formally corresponds to the ISIC rev. 4 divisions 22 (rubber and plastics products), 26 (computer electronic and optical products), 31 (furniture), 32 (other manufacturing) and 38 (waste collection and treatment). In the EDSIC, ISIC rev. 4 division 22 maps to crp (chemical, rubber, plastic products), division 26 to ome (machinery and equipment n.e.c.), 32 to omf (manufactures n.e.c.), and 38 to ser (services n.e.c.). In practice, as the IEA notes warn, countries are liable to use the "non-specified industry" category for any flows they find difficult to assign to particular industries. From the size of the flow, we guess that most of it does in fact represent unclassified flows rather than its four proper ISIC divisions. We apportion 75 per cent of its inputs between the other EDS manufacturing industries and construction. The remaining 25 per cent we assign to the four properly matching EDS industries. For want of better knowledge, we apportion the remainder between these four industries in equal shares.

— The EEB flow "road" includes the operation of road vehicles by the transport industry, as an ancillary activity by other industries, and by private households. We ignore the problem of ancillary activity by other industries. We assign one-half of the road sector's usage of motor gasoline to private consumption, and one-quarter of its usage of gas/diesel oil. Its remaining usage of motor gasoline and gas/diesel oil, and all usage of other fuels by "road"

— liquefied petroleum gases, liquid biomass, etc.

— we assign to the EDS transport industry.

— The EEB flow "commercial and public services" comprises 24 ISIC divisions, most of which map to EDS industry ser, "Services n.e.c.". One, however, division 52, "warehousing and support activities for transportation", maps to EDS industry tpt, "transport". Not knowing how energy use varies across these 24 divisions, we apportion "commercial public and services flows" between the two EDS industries according to the division count, in the proportions 1:23.

— We apportion the EEB flows "non-energy use ind/transf/energy", "non-energy use in transport", and "non-energy use in other sectors" between all energy-using sectors in proportion to their total energy use of EEB commodities.

Aside from these special cases, and basic elements such as production, exports, and imports, we treat the IEA flow classification as a disaggregation of the EDS industry classification. Similarly, we treat the IEA products as subclasses of EDS energy commodities. Table 12.3 shows the correspondence between IEA flows and EDS industries, and table 12.4 the correspondence between IEA products and EDS energy commodities.

Table 12.3: Concordance between EEB flows and EDS industries

EDS industry		EEB flows
Code	Description	Description
agr	Agriculture, forestry, and fishing	Agriculture
coa	Coal	Coal mines
oil	Oil	Oil and gas extraction (part)
gas	Gas	Oil and gas extraction (part), Gas works
omn	Minerals n.e.c.	Mining and quarrying
fpr	Food products, beverages	Food and tobacco
twl	Textiles, wearing apparel, leather	Textile and leather
lum	Wood products	Wood and wood products
ppp	Paper products, publishing	Paper, pulp and printing
p_c	Petroleum, coal products	Patent fuel plants, Coke ovens, Blast furnaces, Petrochemical industry, Brown coal briquettes (BKB) plants, Petroleum refineries, Liquefaction plants, LNG plants
crp	Chemical, rubber, plastic products	Chemical and petrochemical, Non-specified industry (1/16)
nmm	Mineral products n.e.c.	Non-metallic minerals
i_s	Ferrous metals	Iron and steel
nfm	Metals n.e.c.	Non-ferrous metals
teq	Transport equipment	Transport equipment
ome	Machinery and equipment n.e.c.	Machinery, Non-specified industry (1/16)
omf	Manufactures n.e.c.	Non-specified industry (1/16)
ely	Electricity	Public electricity plant, Autoproducer electricity plant, Public Combined heat and power (CHP) plant, Autoproducer CHP plant, Public heat plant, Autoproducer heat plant, Heat pumps, Electric boilers, Own use in electricity, CHP and heat plants, Pumped storage (electricity)
cns	Construction	Construction
tpt	Transport	International civil aviation, Domestic air transport, Road, Rail, Pipeline transport, Internal navigation, Non-specified transport, Commercial and public services (1/24)
ser	Services n.e.c.	Non-specified industry (1/16), Commercial and public services (23/24), Non-specified other
...	Production	Production, From other sources — primary energy
...	Private consumption	Residential
...	Exports	Exports, International marine bunkers
...	Imports	Imports
...	Special treatment	Transfers, Non-specified energy sector, Distribution losses, Non-specified industry (3/4), Non-energy use in ind/transf/energy, Non-energy use in transport, Non-energy use in other sectors

Table 12.4: Concordance between EEB Products and EDS Commodities

Code	EDS commodity Description	EEB products Description
coa	Coal	Coking coal, Other bituminous coal and anthracite, Sub-bituminous coal, Lignite/brown coal, Peat, Patent fuel, BKB/peat briquettes
oil	Oil	Crude oil
gas	Gas	Gas works gas, Natural gas, Natural gas liquids
p_c	Petroleum, coal products	Coke oven coke and lignite coke, Gas coke, Coke oven gas, Blast furnace gas, Oxygen steel furnace gas, Refinery feedstocks, Refinery gas, Ethane, Liquefied petroleum gases (LPG), Motor gasoline, Aviation gasoline, Gasoline type jet fuel, Kerosene type jet fuel, Other kerosene, Gas/diesel oil, Heavy fuel oil, Naphtha, White spirit and SBP, Lubricants, Bitumen, Paraffin waxes, Petroleum coke, Other petroleum products
ely	Electricity	Electricity, Heat
...	Discarded	Combustible renewables and waste, Industrial wastes, Municipal wastes renewables, Municipal wastes non-renewables, Primary solid biomass, Biogas, Liquid biomass, Non-specified combust. renewables and waste, Charcoal, Additives/blending compounds, Inputs other than crude or NGL, Nuclear, Hydro, Geothermal, Solar photovoltaics, Solar thermal, Tide, wave and ocean, Wind, Other fuel sources of electricity

Some of the sectoral mappings are doubtful of necessity or anomalous by inadvertence or choice. The main motive for the deliberate anomalies is to avoid recognizing energy production by EDS industries that predominantly produce non-energy goods, for fear that this might complicate later energy data processing.

— We record inflows into the EEB flow "international marine bunkers" as exports, but record no corresponding imports. International marine bunker usage is assigned to the right countries, on the basis of international services trade data. The assumption is that international marine bunker usage is proportional to water transport services exports (McDougall and Leeuwen 2010). The ideal handling of bunkers requires much more than sectoral remapping. In principle, we should balance the trade flows by recording international marine bunker usage as imports into the country of residence of the ship operator, and as use by that country's transport industry. This being a sizable task, and data on ship operators' nationality being lacking, we set it aside as an area for future work.

— We class inflows of the EEB flow "transfers" as inputs into the petroleum and coal products industry, and outflows as production of petroleum and coal products. In fact, if we understand the EEB correctly, this flow is an accounting convenience, allowing the IEA to record production and consumption of the same substance under two different descriptions (for example, production as "natural gas liquids" but consumption as "liquefied petroleum gas"). The accounts are balanced by entries to the "transfers" flow (for example, inflow of natural gas liquids into "transfers", and outflow of liquefied petroleum gas from "transfers"). Our representing these fictitious flows as intra-industry usage of petroleum products is mostly harmless (but see section 12.4.3), for EEB products mapped to the ESD commodity "petroleum and coal products", but creates imbalances for EEB products mapped to other ESD commodities or discarded.

— We class the flow “heat” and the product “heat plant” with electricity. A more ISIC-compliant treatment would class them with gas, since the EDS sector gas includes the standard GTAP sector gdt which includes the ISIC group 353, “steam and hot water supply”. We do not however feel bound by our own definition of gdt, since it was adopted without a view to energy data set construction, and since its I-O data are in any case to be overwritten with EDS data. Since a great part of heat output is produced jointly with electricity in combined heat and power plant (Chepeliev 2020), and since in cost structure heat production is more similar to electricity than to gas, we find it expedient to class it with electricity.

— Since we class heat with electricity, we class the operation of combined heat and power plant, heat pumps, and electric boilers with electricity generation.

— Our treatment of patent fuel is inconsistent. We class the product with coal, but the flow “patent fuel plants” with petroleum and coal processing. ISIC compliance would class both flow and product with coal; in future we may adopt that treatment.

— Our treatment of the EEB flow “petrochemical industry” and the EEB product “refinery feedstocks” is not ISIC-compliant. Following the ISIC, we would class these in the EDS “chemical, rubber, plastic products”; in fact, we class them in “petroleum, coal products”. Again, this allows us to avoid treating the chemical industry as an energy industry.

— We discard the EEB flow “gasification plants for biogas” and the EEB product “biogas”. An ISIC-compliant treatment would map the flow to ISIC group 352 as “manufacture of gas”, and therefore map both the flow and the product to EDS sector gas.

— We discard the EEB flow “nuclear industry”. In the EEB this flow serves mainly to record consumption of electricity by nuclear power plants; in future, we may record this as sales of “electricity” to “electricity”.

— We discard the EEB product “charcoal”, which maps under the ISIC to basic chemicals, to avoid treating the chemicals industry as energy-producing.

— Some part of the EEB flow “residential”, corresponding to such activities as heating of apartment blocks, should likely map to the EDS services sector. Lacking the data to split it, we assign it all to private consumption.

— What the EEB flow “non-specified other” includes is not altogether clear, but the IEA notes give one example, military fuel consumption. Military fuel consumption maps to ISIC division 84, “public administration and defense; compulsory social security”, and to EDS industry ser, “services n.e.c.”. Accordingly, we map “non-specified other” to ser.

— The EEB products “industrial wastes”, “municipal wastes renewables”, and “municipal wastes non-renewables” are difficult to locate in the ISIC. Part may be output of ISIC class 4669, “wholesale of other intermediate products, waste and scrap” (included in EDS industry “services n.e.c.”), part output of ISIC divisions 37-39 (sewerage, waste collection and treatment, etc.), (also included in “services n.e.c.), part output of the waste producing industries themselves, part sales by final buyers (i.e., households). We simply discard them.

— We discard the EEB product “primary solid biomass”, rather than recognize it as output of agriculture and forestry; likewise, we discard “liquid biomass” (largely ethanol) and “additives/blending compounds”, rather than recognize them as output of the chemicals industry.

— We discard the EEB products “non-specified combust. renewables and waste” and “other fuel sources of electricity”, which we don’t know how to class.

— We discard the EEB products “nuclear”, “hydro”, “geothermal”, “solar photovoltaics”, “solar thermal”, “tide, wind, and ocean”, and “wind”, as items of convenience in the EEB system, but not required in the I-O accounts.

12.2.3 Mapping from GTAP Countries to GTAP Regions

After performing the sectoral mapping, we aggregate from GTAP countries to GTAP regions, a routine operation requiring no special comment.

12.2.4 Balancing for Internal Consistency

The IEA’s EEB data set is balanced within countries, in that total supply and total consumption of each product are equal. It is not balanced across countries, in that world exports and world imports of each commodity differ. We create further inconsistencies, in adding new regions, mapping some sectors inconsistently, and discarding some flows.

After mapping to GTAP categories, therefore, we have imbalances both between supply and use within countries, and between exports and imports globally (table 12.5). As table 12.5 shows, the export-import imbalances are all moderate.

Table 12.5: Balancing Imports and Exports, World Totals (MTOE), 2017

Commodity	Before balancing Exports	Before balancing Imports	After balancing
Coal	839	808	824
Oil	2246	2350	2298
Gas	1094	1013	1050
Petroleum products	1929	1885	1907
Electricity	63	64	63

To remove the imbalances, in an I-O quantity model, we rescale use and supply in opposite senses to remove supply-use imbalances, and simultaneously rescale exports and imports in opposite senses to remove export-import imbalances. The re-scalings are interrelated, since imports are part of supply, and exports part of use.

12.2.5 Removing Reexports

The Extended Energy Balances record in each country’s imports or exports all goods that cross the country’s border. Thus, countries can export not only their own products but goods imported from elsewhere; these are called reexports. In GTAP, however, we do not record reexports in the transit countries trade; we record the trade as exports from the country of origin to the final destination. We need therefore to remove reexports from the trade data.

We do this in the simplest way, assuming that imports and domestic products have the same propensity to be exported. To the extent that trade is driven by product differentiation, this treatment may tend to over-adjust for reexports. The adjustment reduces global energy trade by 13 per cent in 2017, from 6.1 to 5.3 GTOE (Table 12.6). Among commodities, the adjustment is greatest for petroleum and coal products, trade in which falls by 28.9 per cent, from 1.9 to 1.4 GTOE (Table 12.6).

Table 12.6: Trade before and after Removing Reexports, World Totals (MTOE), 2017

Commodity	Trade (Before)	Trade (After)	Percentage change
Coal	824	803	-2.5
Oil	2298	2186	-4.8
Gas	1050	942	-10.2
Petroleum products	1907	1356	-28.9
Electricity	63	58	-8.4
Total	6141	5346	-13

12.2.6 Eliminating Energy Usage by Non-Existent Industries

We adjust the volume data to eliminate certain internal inconsistencies. Specifically, we ensure that energy industries with zero output also have zero usage of energy commodities. We make the adjustment in two steps. In the first step, we zero out selected data cells to impose the condition just described. In the second step, we rebalance the data to restore certain balance conditions: that global imports and global exports are equal for each energy commodity, that domestic production is equal to usage of domestic product for each energy commodity in each country, and that domestic absorption is equal to net supply to the domestic market for each energy commodity in each country.

12.3 Energy Prices and Tax Rates: Initial Construction

Since the GTAP Data Base is essentially a collection of money value flows, we must convert the energy volume data described in section 12.2 into money values in order to use it in the database. And to do that, we require data on energy prices.

For energy prices, as for volumes, the principle data source is the IEA. But in every other way we face a far different situation. Whereas the EEB provides a unified, comprehensive, and consistent representation of energy volumes, for prices the data series are multiple, partial, and diverse in every way. To fill in some critical gaps in the IEA data, we employ several other data sources, including some non-serial publications, and some subsequently discontinued series.

In the initial implementation of the price data construction procedure, since uniformly up-to-date data could not be obtained, data were extracted for several reference years (G and M 1999). These were updated to the GTAP Data Base reference year using time series data for energy price indices and exchange rates. For the GTAP 11 implementation, we follow a similar approach to match different base years. What has been said about prices applies equally to tax rates. It remains only to add that since the tax data comprise both tax-inclusive and tax-exclusive prices, we prepare price and tax data together.

For the more heterogeneous EDS energy commodities, especially "petroleum and coal products", we cannot expect to find prices at the EDS commodity level. At the same time, the available price data presents much less commodity detail than the EEB product level. The initial energy price data set (EPDS) therefore uses an intermediate (and incomplete) energy commodity classification, which we may call the energy price data commodity classification (EPDCC). This comprises nine commodities, mapped to EDS commodities as shown in table 12.7.

Table 12.7: EPD commodities, with Concordance to EDS Commodities

EPD commodity		EDS commodity	
Code	Description	Code	Description
S	Steaming coal	coa	Coal
C	Coking coal	coa	Coal
O	Crude oil	oil	Oil
N	Natural gas	gas	Gas
H	Heavy fuel oil	p_c	Petroleum, coal products
L	Light fuel oil	p_c	Petroleum, coal products
G	Gasoline	p_c	Petroleum, coal products
D	Diesel oil	p_c	Petroleum, coal products
E	Electricity	ely	Electricity

One feature of the available price and tax data is that they often present several prices for the same commodity and country, differentiated by user. Naturally, the user detail is at much less than the EDS industry level, let alone the EEB flow level. To accommodate the available data we define an EPD use classification, with just four classes, as shown in table 12.8.

Table 12.8: EPD Use Classes

Code	Description	Corresponding EDS use classes
I	Industrial	Intermediate usage by industries other than electricity
H	Household	Private consumption
U	Utilities	Intermediate usage by the electricity industry
X	Exports	Exports

To accommodate the available data, we need much less than the GTAP standard country set. A set of 92 EPD countries is sufficient, the EPD country set. This is just a subset of the GTAP standard country set.

In constructing the energy price data set we deal with prices at several levels. Purchasers' prices include not only the cost of production but also taxes and trade and transport margins. Purchasers' prices are differentiated by use class. For the use class "exports", they cover margins incurred in bringing the goods to the point of export, and any export taxes or subsidies by the exporting country, but exclude international trade margins and import duties. Thus, they are equivalent to free-on-board (FOB) prices. Basic prices are prices at the point of production, for domestic product, or landed duty-paid prices, for imports.

We construct an initial energy price data set that includes arrays of basic prices, FOB export prices, and tax rates. All prices and tax rates are measured in US dollars per TOE, or, equivalently, in millions of US dollars per MTOE. All are indexed by EPD commodity and country. The tax rates are also indexed by the EPD use class. Production subsidy rates are not included in this data set.

We obtain basic price estimates in three different ways. In a few cases, they are available from source data. More often, we must calculate them, either from import prices or from purchasers' prices. For import-price-based estimates, we calculate basic (landed duty-paid) prices of imports from duty-free (CIF) prices and import duty rates. For purchasers'-price-based estimates, we take use-specific data for purchasers' prices, and calculate use-specific basic price estimates by removing taxes and margins. We then average across uses to obtain use-generic estimates of basic prices.

We then map these data set to EDS categories. First, we aggregate from EPD to EDS commodities, then extend from EPD to standard GTAP countries, then aggregate from GTAP countries to GTAP regions. Key data sources for energy prices and taxes include:

— Energy Prices and Taxes (IEA 2021a),

— the Organización Latinoamericana de Energía's (OLADE's) SIEE database (OLADE 2021),

— the United States Department of Energy's Petroleum Marketing Monthly (DOE 2011) and International Energy Annual (DOE 2002),

— the Lawrence Berkeley National Laboratory's China Energy Databook (Fridley and Aden 2008),

— United Nations (UN) COMTRADE database (UN 2021) for export prices of natural gas.

Because the prices and tax measures involve a variety of physical units –barrels, gallons, kilowatt hours, and so forth –we require physical conversion factors to bring all prices and taxes to a TOE basis.

We use value-share-weighted averages at many points, including averaging across uses and aggregating from EPD to EDS commodities; to calculate these, we need energy volume data. For reasons relating to difficulties in program maintenance, we use not the energy volume data set described in section 12.2, but an older data set acquired for the original implementation. To pass from purchasers' prices to basic prices, we need margin rates; for these, we rely in part on the comparison of import and purchasers' prices, in part on outside margin rate estimates.

The source data for the energy price data set therefore include physical conversion factors, exchange rates, price indices, import duty rates, import prices, margin rates, purchasers' prices, and tax rates.

12.3.1 Physical Conversion Factors

Prices for different energy products are expressed in a variety of units. Physical unit conversion factors are needed in order to convert these prices into tons of oil equivalent (TOE). A single conversion factor is used for all countries, though, in reality, net calorific contents differ somewhat from one country to another. Table 12.9 reports the conversion factors used in this version of the database.

Table 12.9: Conversion Factors

Commodity	Unit	Tons of oil equivalent (TOE)	Source
Steam coal	ton	0.63	US coefficients in IEA 1997 p.43.
Coking coal	ton	0.709	US coefficients in IEA 1997 p.43.
Crude oil	barrel (bbl)	0.1339	Coefficient for Philippine from ADB (1994) table 11, p. XXIX.
Natural gas	cubic meter (m3)	0.0008645	Gross calorific value for Germany from IEA 1997 p. 36 converted from kcal to TOE
Gasoline	barrel (bbl)	0.1324762	Conversion factor for ”all other countries” from IEA 1997 p. 41, converted from liter to bbl
LFO	barrel (bbl)	0.1363651	Conversion factor for ”all other countries” from IEA 1997 p. 41, converted from liter to bbl
HFO	ton	0.96	Conversion factor for ”all other countries” from IEA 1997 p. 42.
Diesel	barrel (bbl)	0.14127	Conversion factor for ”all other countries” from IEA 1997 p. 41, converted from liter to bbl

12.3.2 Price Indices

For exchange rates, we have data for the national currency price of the US dollar, for all 92 countries in the EPD country set. For price indices, our sole data source is the IEA's Energy Prices and Taxes (IEA 2021a). This provides national currency price indices for coal, natural gas, oil products, and electricity. We use the single coal price index for the two corresponding EPD commodities (steaming coal, coking coal), and the single oil products price index for the four corresponding EPD commodities (heavy fuel oil, light fuel oil, gasoline, diesel oil). For crude oil, we take the index of real crude import costs in national currencies (a questionable procedure).

For each index, we extract data from 1978 to 2020. We extract all available countries – up to 37, depending on the series. To fill in missing countries, we use the United States price index, adjusted for exchange rate changes. Where exchange rates are lacking, we just use the United States price index. Of the price series we need to update, some are in national currency units and some in US dollars. From the national currency price indices and exchange rates, we calculate combined price update and exchange conversion factors. For national currency price series, we use the formula

$$P_{ED1} = \frac{1}{E_1} \frac{P_{IN1}}{P_{IN0}} P_{AN0},$$

where P_{ED1} denotes the estimated US dollar price in the later year, P_{AN0} the actual national currency price in the earlier year, P_{IN0} and P_{IN1} national currency price indices in the earlier and later year, and E_1 the exchange rate in the later year. For US dollar price series, we use the formula

$$P_{ED1} = \frac{P_{IN1}/E_1}{P_{IN0}/E_0} P_{AD0},$$

where E_0 denotes the exchange rate in the earlier year, and P_{AD0} the actual US dollar price in the earlier year.

12.3.3 Tax Rates

For taxes as for prices we collect use-specific values, which we arrange in the EPD use classification. We use several sources; where sources overlap, we use the IEA data.

— From the IEA we take tax rates for many commodity-use-country triples for 1978 to 2020. The IEA reports both excise and value added tax, but we take just the excise tax rate.

— From OLADE's SIEE data (OLADE 2021), we take tax rates for twenty-seven Latin American countries for household, utility and industrial use of gasoline, diesel oil, and electricity for 1988 to 2017.

After taking all available data, estimates are still lacking for many commodity-use-country triples. We note for future reference which rates are known and which are unknown. For the price and tax calculations, wherever a required rate is unknown, we assume a rate of zero.

12.3.4 Basic Prices Taken Directly from Source Data

In a few cases, we have source data for basic prices. From the China Energy Databook, we take the coking coal and steaming coal prices averaged across months within years in 1996-2006, and basic prices of gasoline and diesel oil in 2005-2008.

We convert and update these to the GTAP 11 reference year US dollar prices per TOE, using the price and physical conversion factors described in sections 12.3.1 and 12.3.2.

12.3.5 Basic Prices Calculated from Import Duty Rates and CIF Import Prices

For 170 commodity-country pairs, we estimate basic prices from import duty rates and CIF import prices. For duty rates, we take ad valorem rates from the GTAP protection data set (chapter 11). This provides estimates for 186 of the 251 GTAP countries. The commodity classification is of course standard GTAP; this we expand to the EPDCC by assigning each EPD commodity the rate for the corresponding GTAP commodity. For the missing countries, we set duty rates to zero. For CIF import prices, we take data from the IEA 2021a and OLADE 2021:

— from the IEA, prices for steaming coal, coking coal, crude oil and natural gas imports, covering 1980-2020 timeframe and a total of 28 countries (data availability differs by years and commodities),

— from OLADE, prices for eleven energy sources imported from 27 countries, covering 1970-2017 timeframe.

We convert these prices from their various original units to US dollars per TOE and rescale from the source year to GTAP 11 reference years. From them and the duty rates we calculate basic prices, according to the formula

$$BPM = (1 + \frac{1}{100}MDR)CIF, \tag{12.1}$$

where BPM denotes the basic price of imports, MDR the import duty rate (ad valorem or ad valorem equivalent), and CIF the CIF price of imports.

12.3.6 Purchasers' Prices

For purchasers' prices we use an especially wide variety of sources. Following the IEA's Energy Prices and Taxes, we record all domestic purchasers' prices as prices for industry, households, or electricity generation, though not all sources use those descriptions. In some cases, we take averages; for example, to obtain prices paid by households and industry for electricity in China, we average prices reported in the China Energy Databook across provinces. The commodity descriptions also do not always match exactly — from the IEA export price data, for instance, we apply the gasoil price to light fuel oil, and an average of the low sulfur fuel oil and high sulfur prices to heavy fuel oil. From the IEA's Energy Prices and Taxes we extract time series data; from the other sources we extract just the latest available year case-by-case. Further detail is available in Malcolm and Babiker 1998. For natural gas, country-specific export prices are calculated based on the UN COMTRADE (UN 2021) dataset. The latter provides data for 56-61 country-cases depending on the reference year.

We do not attempt to fill all combinations of commodities and use classes; rather, we assume that some combinations, for example household use of heavy fuel oil, or utilities use of light fuel oil, do not occur and require no price data. We take:

— from the IEA, prices for industrial, household and utility use of steaming coal, coking coal, natural gas, heavy fuel oil, light fuel oil, diesel oil, petroleum products, and electricity for various overlapping sets of up to 37 countries, from 1978 to 2020,

— also from the IEA, prices of exports of crude oil, steaming and coking coal and various oil products for selected energy exporters, from 1960-2020 (timespan varies for different commodities),

— from OLADE, prices for industrial use of steaming coal, coking coal, natural gas, heavy fuel oil, light fuel oil, diesel, and electricity, household use of natural gas, light fuel oil, gasoline, diesel, and electricity, and exports of steaming coal, coking coal, crude oil, natural gas, heavy fuel oil, light fuel oil, diesel and gasoline, for 27 Latin American and Caribbean countries in 1988-2017,

— from the United States Department of Energy, prices for industrial use of crude oil in the United States in 1990-2008, and for industry use of heavy fuel oil, light fuel oil, and diesel oil, and household use of light fuel oil and gasoline, in up to 92 countries in 1995-2002,

— from the China Energy Databook, prices for industrial use of steaming coal and coking coal in 1996-2006, heavy fuel oil, and diesel oil in 2005-2008 and household and industrial use of diesel and gasoline in 2005-2008, industrial and household use of electricity in 2008.

As with basic prices, we convert and update these purchasers' prices to GTAP 11 reference year US dollar prices per TOE, using the price and physical conversion factors described in sections 12.3.1 and 12.3.2.

We now fill in the missing prices for the three domestic use classes (industrial, household, and utilities). We divide countries into four classes, according as their energy prices are above or below average, and as they are net exporters or importers of energy. We find seven high-price exporters, 51 high-price importers, 24 low-price exporters, and ten low-price importers. For each of the four country classes, and for each product and use class, we calculate an average price as the simple average of the available prices (table 12.10). We fill in missing values with these average prices.

Table 12.10: Cross-Country Average Purchasers' Prices, by EPD Commodity and Use Class (US$ per TOE), 2017

Commodity	Use class	High-price energy exporter	High-price energy importer	Low-price energy exporter	Low-price energy importer
Steaming coal	Industrial	139	222	242	148
Steaming coal	Household	823	785	NA	NA
Steaming coal	Utilities	168	145	85	77
Coking coal	Industrial	206	262	104	247
Crude oil	Industrial	NA	NA	NA	335
Natural gas	Industrial	416	447	140	121
Natural gas	Household	701	952	232	286
Natural gas	Utilities	254	354	207	149
Heavy fuel oil	Industrial	507	505	265	345
Heavy fuel oil	Utilities	389	388	353	462
Light fuel oil	Industrial	835	865	357	517
Light fuel oil	Household	860	959	405	525
Gasoline	Household	1446	1674	697	998
Diesel oil	Industrial	1079	1135	511	722
Diesel oil	Household	1398	1323	667	687
Electricity	Industrial	1190	1471	860	984
Electricity	Household	2053	2296	765	2113

For commodity and use codes, see tables 12.7 and 12.8

In a few cases, we find that the tax rate exceeds the estimated user price. In such cases, we discard the tax rate, and treat the rate as unknown.

For exports, we use a similar procedure, except that, as data are scarcer for export prices than for domestic uses, we do not separate countries into four groups, but calculate a single average price across all available countries. For products other than diesel oil and electricity, we set the missing values equal to the simple average of the available export prices for each commodity (table 12.11). For diesel oil and electricity, no export prices are available; for these commodities, for each country, we set the export price equal to the industrial price.

Table 12.11: Average export prices (US$ per TOE)

Commodity	Price
Steaming coal	129
Coking coal	261
Crude oil	368
Natural gas	280
Heavy fuel oil	298
Light fuel oil	445
Gasoline	475
Diesel oil	NA
Electricity	NA

12.3.7 Basic Prices Calculated from Tax Rates, Margin Rates, and Purchasers' Prices

Our third method of obtaining basic prices is to calculate them from tax rates, margin rates, and purchasers' prices. Sections 12.3.4 and 12.3.7 relate the derivation of the tax rates and purchasers' prices; we now relate the derivation of the margin rates.

Purchasers' prices and basic prices are related to each other according to the formula

$$PP = (1 + MR)BP + TR, \tag{12.2}$$

where PP denotes purchasers' price, MR the margin rate, BP basic price, and TR the tax rate (a specific rate or specific rate equivalent). Where basic price estimates are available from source data (section 12.3.5) or import prices (section 12.3.6), we use this equation to calculate the margin rate. If the tax rate is unknown, we assume a rate of zero.

This yields some plausible and some wild estimates. We discard all estimates for which imports represent less than twenty per cent of total supply, or for which the calculation yields a negative margin rate. In all cases where the calculation yields an estimate greater than 5, we substitute a rate of 5. For each product and use class, we then calculate an average rate of margins on imports as the simple cross-country average of the available rates. For household and industrial use of diesel oil, finding no rates available, we use the rate for household use of gasoline. Overriding the average calculation, we fix the rate for household use of natural gas at 2.0 and for industrial use at 1.3.

We also obtain margin rates for the United States and Japan from I-O data sources. Then, for the United States and Japan, we use their own margin rate; for all other countries, we set the margin rate equal to a weighted average of the (cross-country average) import margin rate, the United States rate, and the Japanese rate, with weights of 0.6, 0.2, and 0.2. Thus, all countries except the United States and Japan have the same margin rates; these we report in table 12.12.

At this point all variables in equation (12.1) are estimated except for the basic price and in some cases the tax rate. We now use the formula to obtain basic prices. If the tax rate is unknown, for purposes of

Table 12.12: Margin Rates, by Product and Use Class, for Countries Other than the United States and Japan

Commodity	Industrial	Household	Utilities	Export
Steaming coal	0.37	3.1	0.3	0.18
Coking coal	0.3	0.29	0.07	0.28
Crude oil	0.05	0	0.06	0.18
Natural gas	0.83	2	0.24	0.57
Heavy fuel oil	0.37	0	0.27	0.08
Light fuel oil	0.57	0.6	0.17	0.01
Gasoline	0	0.56	0	0.11
Diesel oil	0.37	0.56	0	0.01
Electricity	0	0	0	0

this calculation we use a rate of zero. In this way we obtain estimates of basic prices, indexed not only by EPD commodity and country but also by EPD use class.

12.3.8 Basic Prices: Final Estimates

For each EPD commodity in each EPD country, we now have up to five estimates of the basic price. We may have an estimate extracted from source data or calculated from import prices (in no case do we have both), and we may have up to four estimates calculated from purchasers' prices for the several EPD use classes. Our preferred estimate of the basic price is the estimate extracted from source data or calculated from import prices, if either is available; otherwise, it is a share-weighted average of the available purchasers'-price-based estimates. We use the special-purpose energy volume data (section 12.3.3) to calculate the weights.

We now fill in remaining missing values in two stages. First, for each commodity-use pair, if we have observations for some countries but not all, we fill in the missing countries with the simple average of the available countries' prices. Because of the previous filling in of the purchasers' prices (section 12.3.8), such cases arise only for commodity-use pairs for which no purchasers' prices are available, for example, gasoline usage by industry. In such cases, the available observations are either import-price-based (section 12.3.6) or taken directly from source data (section 12.3.5). The only missing observations are now for prices of electricity in usage by utilities; we set them equal to the corresponding prices in usage by industry.

12.3.9 Aggregating from EPD to ESD Commodities

At this point, we have in hand estimates of basic prices, export prices, indexed by EPD commodity and country (the export prices are a slice of the purchasers' price array obtained as described in section 12.3.7; the part of the array relating to domestic uses may now be discarded). We also have estimates of tax rates (section 12.3.4), indexed as the prices but also by EPD use class. We now aggregate the price and taxes from the nine EPD commodities to the five ESD commodities, using share-weighted averages, calculated from the special-purpose energy volume data (section 12.3.3).

12.3.10 Mapping from EPD Countries to GTAP Regions

We map energy price and tax data from IEA countries to GTAP regions in two steps. First, we map from IEA countries to standard GTAP countries, then from standard GTAP countries to GTAP regions.

Of the 251 standard GTAP countries, 92 are included in the original price and tax data. For the rest of the countries, we assign proxies, and set their prices and tax rates equal to those for the proxy countries. In the main, the proxy countries match their partners in geographical area and direction of net energy trade (importer or exporter), but some quirks are apparent (for example, the matching of Guinea to Brazil rather than to Ghana, and the matching of Barbados to Brazil rather than to itself). Table 12.13 shows the matching; to save space, we have omitted countries matched to themselves.

Table 12.13: Matching GTAP to EPD Countries: Proxy and Missing Countries

EPD country	GTAP countries
United Arab Emirates	Bahrain, Oman, Yemen
Australia	Cocos (Keeling) Islands, Christmas Island, Heard Island and McDonald Islands, Norfolk Island, Pitcairn
Azerbaijan	Armenia, Georgia
Brazil	Aruba, Anguilla, Antigua and Barbuda, Bahamas, Belize, Barbados, Comoros, Cayman Islands, Dominica, Falkland Islands (Malvinas), Guinea, Guadeloupe, Grenada, French Guiana, Guyana, Saint Kitts and Nevis, Saint Lucia, Montserrat, Martinique, Puerto Rico, Suriname, Turks and Caicos Islands, Saint Vincent and the Grenadines, British Virgin Islands, United States Virgin Islands
Barbados	Bonaire, Sint Eustatius and Saba, Saint Barthélemy, Curaçao, Saint Martin (French Part), Sint Maarten (Dutch part)
Canada	Saint Pierre and Miquelon
Switzerland	Liechtenstein
China	Macau
Denmark	Greenland
Spain	Andorra, Gibraltar
Estonia	Croatia
Finland	Aland Islands
France	Monaco
United Kingdom	Faroe Islands, Guernsey, Isle of Man, British Indian Ocean Territory, Jersey, North Macedonia, Sark, South Georgia and the South Sandwich Islands
Ghana	Angola, Burundi, Burkina Faso, Botswana, Central African Republic, Côte d'Ivoire, Cameroon, Democratic Republic of the Congo, Congo, Cabo Verde, Djibouti, Eritrea, Gambia, Guinea-Bissau, Equatorial Guinea, Kenya, Liberia, Lesotho, Madagascar, Mali, Mozambique, Mauritania, Mauritius, Malawi, Mayotte, Namibia, Niger, Réunion, Rwanda, Sudan, Senegal, Saint Helena, Sierra Leone, Somalia, South Sudan, Sao Tome and Principe, Eswatini, Seychelles, Chad, Togo, United Republic of Tanzania, Uganda, Zambia, Zimbabwe

Continued ...

. . . Continued

EPD country	GTAP countries
Greece	Albania, Bosnia and Herzegovina, Cyprus, Montenegro, Serbia
Indonesia	Timor-Leste
India	Maldives, Myanmar
Iran, Islamic Republic of	Iraq
Israel	State of Palestine
Italy	San Marino, Holy See
Kazakhstan	Turkmenistan, Uzbekistan
Korea, Republic of	Korea, Democratic People's Republic of
Lebanon	Jordan, Syrian Arab Republic
Libyan Arab Jamahiriya	Benin, Egypt, Ethiopia
Lithuania	Belarus, Moldova, Republic of, Ukraine
Morocco	Western Sahara
Norway	Bouvet Island, Svalbard and Jan Mayen
Romania	Bulgaria
Russian Federation	Tajikistan
Slovakia	Kosovo
Slovenia	Malta
Thailand	Afghanistan, American Samoa, Antarctica, French Southern Territories, Bhutan, Cook Islands, Fiji, Micronesia (Federated States of), Guam, Cambodia, Kiribati, Lao People's Democratic Republic, Marshall Islands, Mongolia, Northern Mariana Islands, New Caledonia, Niue, Nauru, Palau, French Polynesia, Solomon Islands, Tokelau, Tonga, Tuvalu, Vanuatu, Wallis and Futuna Islands, Samoa
United States of America	Bermuda, United States Minor Outlying Islands

We then average across countries to obtain prices and tax rates for GTAP regions. For tax rates and basic prices, we use domestic-absorption-weighted averages, where domestic absorption is calculated as imports plus production less exports, taken from the energy volume data set after mapping to EPD commodities and industries but before aggregating to GTAP regions (that is, as at the end of section 12.2.2).

For export (FOB) prices, we use trade shares, defined by adding exports and imports together; unless no trade data are available for countries belonging to the region, in which case we use simple averages; and unless all countries belong to the region are missing, in which case we use a world average export price. Thus, we ensure a full array of non-zero export prices.

12.4 Combining Price and Volume Data

With energy price and volume data in hand, we can now construct money value estimates for international trade and for I-O flows. In so doing, we find it necessary to make some adjustments to both volumes and prices.

As with the GTAP Data Base itself, we give priority to the trade data over the I-O data. The reason is the same: we cannot construct a balanced trade data set from the I-O data, so, where they differ, the I-O data must adjust to the trade data. Although the volume data are balanced internationally (section 12.2.4), the money value data will be unbalanced unless import and export prices are mutually consistent; and nothing in the initial construction of the price data (section 12.3) enforces international consistency of prices.

Strictly speaking, the initial price data set of section 12.3, although it draws on import prices, contains no import price data specifically, since the estimated basic prices apply equally to imports and domestic product. But assuming import parity pricing, they do apply to imports; and yet in general, they are liable to be inconsistent with the export prices and any reasonable estimate of international trade margins.

Our first main task therefore is to construct an internally consistent data set covering volumes, prices, and money values in international energy trade. For this, we draw both on the unilateral trade volumes in the energy volumes data set, and the bilateral trade money values in the trade data set (chapter 10). The second main task is to construct data sets covering volumes, prices, and money values for I-O flows in the individual regions. As it turns out, the I-O money values as initially constructed fail to meet certain necessary inequalities; this we correct by revising the price data.

Some minor tasks also emerge. For technical reasons relating to I-O table balancing (chapter A), it is expedient to replace many zero values in the data with small non-zeros, an operation we call data smearing. For this purpose, we draw on the (unadjusted) I-O tables (chapter 8). To verify the internal consistency of prices and money values in I-O flows (section 12.4.3), we need production subsidy rates; we draw these mainly from the protection data set (chapter 11), but some energy-specific revisions are necessary. And tax rates must be converted from the specific rates recorded in the energy price data set (section 12.3) to the ad valorem rates used in the I-O table balancing. We proceed in this order:

1. Smear the volume data.
2. Construct energy trade data.
3. Revise the production subsidy data.
4. Construct initial estimates of I-O money value flows, and adjust prices to ensure internal consistency of the I-O money value flows.
5. Convert tax rates from specific to ad valorem.

12.4.1 Energy Volumes: Smearing

We adjust the volume data to replace zero levels of domestic production with strictly positive levels. This is obviously undesirable in principle, but since the new non-negative values are small, they should cause little harm in practice. The reason for the adjustment is that the FIT program used to adjust the regional I-O tables to match the energy data (chapter A) cannot cope with zero or tiny shares of domestic product in total usage. This remains an area for further work in later GTAP releases.

We make the adjustment in two stages. First, we make very small adjustments to import, production, and usage volumes, using the regional I-O tables. The resulting energy usage structure is a weighted average of the original structure in the energy data and a structure derived from the I-O tables, with a high weight (31/32) given to the original energy data and a low weight (1/32) to the I-O tables. Since domestic usage and production is non-zero for each commodity in each regional I-O table, this procedure ensures that they are non-zero also in the energy data set.

12.4.2 International Energy Trade

We construct a bilateral international energy trade data set from the export and import slices of the energy volume data (sections 12.2, 12.4.1), the energy block of the international trade data set (chapter 10), and the export price data from the energy price data set (section 12.3).

We revise the trade data set so that for energy commodities it is compatible with the energy data. We take the volume of each region's exports and imports of each energy commodity from the energy volume and price data sets. We take bilateral trade matrices of FOB money values for GTAP energy commodities from the trade data set, aggregate from GTAP to EDS commodities, and divide by export prices from the energy data set to create initial estimates of bilateral trade volumes. We replace zero values in the initial estimates with small non-zeros to facilitate rebalancing, then balance these matrices against the energy volume data set's export and import vectors. Finally, we multiply by export prices to construct bilateral money values, and disaggregate from EDS to standard GTAP commodities using shares from the initial trade data set.

One important modification for the treatment of the bilateral energy data introduced in the GTAP 11 Data Base includes reliance on additional data sources for the specification of bilateral energy trade flows. In particular, one major issue identified in the energy block of the international trade data set (chapter 10) that originates from the underlying UN COMTRADE data (UN 2021) is a significant share of cases with unclassified sources and destinations. For instance, COMTRADE reports almost no imports of natural gas by Germany from Russia, which leads to major misspecification of the resulting bilateral trade flows since we use COMTRADE bilateral shares as starting values. To correct for this, we are now using complementary data sources, including BP 2022 and Eurostat 2022.

For the balancing, we use the bi-proportional adjustment (RAS) method (see, e.g., Schneider and SA 1990): each element A_{ij} in the initial matrix, representing the volume of exports of an energy commodity from region i to region j, is replaced by a new value $R_i A_{ij} S_j$, where R_i and S_j are scaling factors endogenously determined so that the adjusted bilateral matrix agrees with the specified regional export and import totals.

Table 12.14 summarizes the resultant changes in the trade data set. As it shows, there are substantial reductions in the total money value of trade in oil, gas and petroleum products.

We also find some considerable differences between the original trade data and the energy volume data on the pattern of trade. To remove the effect of the country-generic changes in trade values shown in table 12.15, we rescale the original trade arrays to match the commodity sums from the revised arrays. We then identify significant changes between original and revised energy exports (table 12.14) and imports (table 12.16).

In general, the differences between the original and revised trade data appear quite considerable. In several cases, the original trade data show substantial export flows where the energy data set has zero or insignificant exports, for example, gas from Belgium and Switzerland, or petroleum products from Malta.

In one way, the original and revised data are not comparable. The original data include travelers' expenditures (consumption abroad by non-residents), some US$222 billion in total for 2017, which we

Table 12.14: Original and Revised Export Profile, Selected Cases (US$ million CIF), 2017

Country	Commodity	Original exports	Revised exports
Switzerland	gas	4677	18
Malaysia	p_c	18590	3612
Iran (Islamic Republic of)	oil	78877	44074
Belgium	gas	2835	11
Belgium	p_c	24677	10440
Latvia	p_c	2073	10
United Kingdom of Great Britain and Northern Ireland	oil	18826	7639
Germany	gas	4184	339
Netherlands	oil	1773	10
Bahrain	p_c	470	4093
Rest of Caribbean	oil	1471	10
South Africa	oil	1309	5
Russian Federation	gas	59840	41562
Malta	p_c	1171	5
France	gas	1225	6
Canada	oil	65993	47106
Togo	p_c	1133	4
Switzerland	oil	1116	4
Singapore	p_c	33818	21672
Rest of Western Asia	oil	871	3

Table 12.15: Trade in energy (FOB value in US$ billion), 2017

Commodity	Old	New
Coal	129.3	131.7
Oil	936	810.7
Gas	350.1	284.7
Petroleum products	650.8	547.1
Electricity	35	53.4
Gas distribution	1.3	3.5

Table 12.16: Original and Revised Import Profile, Selected Cases (US$ million CIF), 2017

Country	Commodity	Original exports	Revised exports
Netherlands	p_c	41960	13957
Singapore	p_c	40362	13820
Taiwan, China	oil	44236	17467
United Arab Emirates	p_c	13032	1943
United Arab Emirates	oil	3108	13
China	gas	37619	18067
Panama	p_c	428	5760
Canada	oil	15037	4693
Malaysia	p_c	15789	5361
Ukraine	gas	7221	1301
Togo	p_c	3299	291
China, Hong Kong Special Administrative Region	coa	2	1038
Togo	oil	1173	5
Colombia	oil	1167	5
Rest of Europe	p_c	8153	2626
Iraq	ely	1	903
Rest of Western Africa	p_c	954	4561
Spain	gas	21860	12722
Belgium	p_c	13405	6632
Oman	p_c	2480	294

have assigned among GTAP commodities. A small fraction, some $3.3 billion, is allocated to electricity; but though small as a fraction of travelers' expenditures, it is not negligible in relation to global electricity trade. In principle, we should use the energy data to revise just the cross-border energy flows, and add in the travelers' expenditures after the revision. A the same time, in most of the electricity trade discrepancies reported in tables 12.16 and 12.17, travelers' expenditures do little to explain the discrepancy; in many cases, they reduce rather than increase the discrepancy.

The original trade data set contains not only FOB value but also CIF value and trade margin estimates. From these, we can calculate margin rates in international trade. Combining these with our new FOB money value estimates, we obtain new CIF money value estimates. From the new CIF and FOB money value estimates, and the original modal composition of margins, we obtain new estimates of margin usage; summing these over freight commodity, source of freight, and destination of freight, we obtain new estimates of margin supply. We then have a complete new trade data set, which we use in fitting the I-O tables and assembling the main GTAP data file (chapter A).

Also, by dividing the new CIF values by volumes, we calculate source-specific CIF import price estimates. These together with import duty rates (chapter 11) yield source-specific estimates of basic import prices; averaging these over source countries, we obtain basic import prices indexed by EDS commodity and GTAP region. These we use in constructing I-O money value flows (section 12.4.3).

We note that our final energy import price estimates derive not from the source data for energy import prices, but from data for energy export prices, and from the bilateral trade pattern and margin rates in the merchandise trade data set.

12.4.3 I-O Money Value Flows; Domestic Price Adjustment

At this point, having obtained internally consistent data sets for energy prices and volumes, on a GTAP-compatible basis, we can combine the two to obtain estimates of money values of energy flows. For imports we use the import price estimates from the energy trade data, and for domestic production, the initial basic price estimates. Where basic price estimates are missing, we use world average prices. From these, we calculate source-generic prices which we apply to all domestic absorption. For exports, we use the export price estimates from the energy price data set. Combining these with the energy volume data set (section 12.2) we obtain money values of energy I-O flows.

We note that, although our source data for energy import prices do not affect these import price estimates (which instead derive from export price data, as explained in section 12.4.2), they are, for many countries and commodities, the main determinant of our domestic product price estimates (see further sections 12.3.6 and 12.3.9).

We find that our initial money value estimates are liable to violate certain necessary inequalities: in some instances, the cost of energy inputs into an energy industry exceeds the industry's total costs, as inferred from the value of its output. Under our accounting assumptions, the producer revenue from sales of energy products must be equal to the total cost of production, including the user cost of capital. However, the cost of energy inputs into the energy industry is larger than the total producer revenue, this would imply negative costs of non-energy inputs, which is obviously impossible.

We deal with this by adjusting producer revenue. Specifically, we require that for each energy industry in each country, the cost of energy inputs must not exceed 90 per cent of producer revenue; where it does exceed 90 per cent, we increase the price of the industry's output so that the 90 per cent rule is satisfied.

There are two ways in which we can make the required adjustment in producer revenue. One is to increase the market value of output, the other to reduce production taxes or increase production subsidies. Of course, we can also use both methods at once, increase the output value and the production subsidy simultaneously. In most cases, we just increase the market value of output, but in a few cases—where we believe that there is a large production subsidy, but the energy tax data set provides no estimate—we also adjust the production subsidy.

But these upward adjustments in output values create sectoral imbalances: the value of output in the adjusted industries exceeds the value of sales of the corresponding commodities.

At the same time, we must address the question of consistency between volumes, prices, and money values. When we increase the money value of output, does this entail an increase in output volume or in output price? Since the source data are better for volumes, we maintain the original volume estimates, and adjust the price estimates. By finding new price estimates we remove the sectoral imbalances and the volume-price-value inconsistencies simultaneously.

We cannot however adjust export or import prices country-by-country, since this would create cross-country inconsistencies between export and import money values. Accordingly, we hold export and import prices fixed, and adjust only prices of domestic products absorbed domestically.

In adjusting prices to remove imbalances in one sector we may of course create problems in other sectors. If for example, we raise the price of petroleum products, to remove a revenue shortfall in the petroleum refining industry, we may create revenue shortfalls in industries that use refined petroleum. Indeed, since the refining industry itself uses refined petroleum, we may find that the revenue shortfall within that industry is not fully removed; intra-industry usage has a multiplier effect. To prevent this we make the adjustments within an IO price model, in which changes in input prices generate changes in output prices. We then solve simultaneously for all prices in a region, so as to remove all imbalances in the region.

We note that treating as intra-industry usage transfers of EEB products mapped to p_c (section 12.2.2), while creating no imbalance in the data, does strengthen the multiplier effect, and increases the price adjustment needed to satisfy the 90 per cent rule.

To remove the imbalance, for example, we may raise the price of corresponding energy commodity, and adjust sales values accordingly. We use the IO price model to find the price of domestically produced petroleum products that removes the revenue shortfall in the refining industry, taking account of the effect of petroleum products price increases on the prices of other energy commodities.

As described so far, the procedure has an upward bias: since the initial imbalances all involve deficits in sales values compared to output values, their removal entails raising the prices of the corresponding commodities. To reduce this bias, we allow the procedure to remove imbalances not only by raising output prices and but also by lowering input prices. So, for instance, in the petroleum refining industry, the imbalance can be addressed not only by raising the price of refined petroleum but also by lowering the price of inputs such as crude oil. But since the price of imported crude oil is fixed, and in cases when the country produces very little crude domestically, lowering the price of domestically produced crude would do little to remove the imbalance. So, under such circumstances, the imbalance is removed mostly by raising output prices, and the tendency to raise the general energy price level, though somewhat weakened, remains strong.

After the removal of imbalances, the aggregate energy output is $8.1 trillion, 4.4 per cent above the initial level of the basic value of output. Table 12.17 shows some of the more significant adjustments.

As table 12.17 shows, the adjustment procedure consistently delivers a final value of output a little higher than the initial cost-based estimate. This reflects the multiplier effect of intra-industry usage. Adjustments to the “petroleum and coal products” and “electricity” sectors are the most frequent.

12.4.4 Ad Valorem Tax Rates

With prices finalized, tax rates are converted from specific-rate (per TOE) to ad valorem rates. These, along with missing value flags, are passed to the I-O table balancing module (chapter A).

Table 12.17: Cost Reconciliation and Price Adjustment, Selected Cases, 2017

		Output value US$ billion		Domestic price (US$/TOE)	
Region	Commodity	Initial	Updated	Initial	Updated
jpn	p_c	70999.9	132310.2	420.6	653.4
usa	ely	383752.8	433539.2	1147.2	1148.7
irq	p_c	7801.9	15094.4	296.8	614.3
usa	p_c	348769.8	389703.9	339	389.7
deu	p_c	51300.7	65715.9	402.1	540.8
ita	p_c	30859.7	41974.6	383.2	589.4
irq	oil	106429.6	89246.1	331.1	331.1
gbr	p_c	27226	35026.1	398.5	552.4
kor	p_c	87669.1	99972.9	499.4	593.2
usa	oil	158070.3	173667.8	359.9	360.1
esp	p_c	32810.7	40059.5	428.6	563.5
bra	p_c	50675.9	59533.7	420.1	500.7
est	p_c	43.9	571	430.4	7754.4
arg	ely	4251.6	6947	340.5	550.4
grc	p_c	13330.9	17839.4	423.3	667
tur	p_c	16759.9	21752.4	427.8	570.2
blr	p_c	7068.7	10095.9	359.7	782.1
usa	gas	106232.5	117010.2	127.6	127.8
jpn	ely	171498.3	184736.2	1968.5	2015.6
dza	gas	41260.3	47560	362	362

References

Babiker, M.H. and T.A. Rutherford (1997). *Input-output and general equilibrium estimates of embodied carbon: A data set and static framework for assessment*. Working Paper 97. University of Colorado.

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.

Burniaux, Jean-Marc and Truong Truong (2002). *GTAP-E: An Energy-Environmental Version of the GTAP Model*. GTAP Technical Paper 16. Department of Agricultural Economics, Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). URL: https://doi.org/10.21642/GTAP.TP16.

Chepeliev, Maksym (Dec. 2020). "GTAP-Power Data Base: Version 10". In: *Journal of Global Economic Analysis* 5.2, pp. 110–137. DOI: 10.21642/JGEA.050203AF. URL: https://jgea.org/ojs/index.php/jgea/article/view/108.

DOE (2002). *United States Department of Energy, International Energy Annual*. Washington, DC.

— (2011). *United States Department of Energy, Petroleum Marketing Monthly*. Washington, DC.

Eurostat (2022). *International Trade*. URL: https://ec.europa.eu/eurostat/web/euro-indicators/information-data/international-trade.

Fridley, D. and N. (ed.) Aden (2008). *China Energy Databook, v7.0*. Lawrence Berkeley National Laboratory, Berkeley, California.

G, Malcolm and Babiker M (1999). "The process of Incorporating Energy data into GTAP". Draft GTAP Technical Paper.

IEA (1997). *International Energy Agency, Energy Prices and Taxes, First Quarter, 1997. Paris*. Paris.

— (2021a). *International Energy Agency, Energy Prices and Taxes, May*. Paris.

— (2021b). *International Energy Agency, World Energy Balances*. Paris, France. (April 2021 edition). URL: https://www.iea.org/data-and-statistics/data-product/world-energy-balances.

Malcolm, Gerard and Mustafa Babiker (1998). *Documentation for Energy Price/Tax Database*. Center for Global Trade Analysis. 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=337.

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 Huey-Lin Lee (2006). "An Energy Data Base for GTAP". 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=1951.

McDougall, Robert and Nico van Leeuwen (2010). *International MariBunkers: An Attempt to Assign its Usage to the Right Countries*. GTAP Research Memorandum 20. Department of Agricultural Economics, Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). URL: https://doi.org/10.21642/GTAP.RM20.

OLADE (2021). *Organización Latino Americana De Energía, Sistema de Información Energética.* URL: https://www.olade.org/en/.

Schneider, MH and Zenios SA (1990). "A comparative study of algorithms for matrix balancing". In: *Operations Research* 38, pp. 439–55.

UN (2020). *United Nations, 2018 Energy Balances*. New York. URL: https://unstats.un.org/unsd/energystats/pubs/balance/documents/2018balance.pdf.

— (2021). *United Nations COMTRADE*. URL: https://comtradeplus.un.org/.