The Process of Incorporating Energy Data into GTAP

Gerard Malcolm and Truong Phuoc Truong

DRAFT GTAP Technical Paper

DRAFT FOR COMMENT ONLY

November 1, 1999

The work in this paper was carried out by Gerard Malcolm while a visiting Researcher at the Center for Global Trade Analysis. Truong Phuoc Truong subsequently modified and completed the work and wrote this document based on the technical programs and input files written by Gerard Malcolm.

The Process of Incorporating Energy Data into GTAP

by Gerard Malcolm
and
Truong Phuoc Truong

GTAP Technical Paper No. xxx

Abstract

It was observed that data relating to certain of the energy commodities in version 3 of the GTAP data base were of questionable reliability (Babiker and Rutherford, 1997). In particular the data on energy commodities seemed to be at variance with the energy statistics of the International Energy Agency (IEA). A project was undertaken at the Center for Global Trade Analysis with the financial support from the US Department of Energy to look into this issue. As part of the Project, Christophe Complainville of the OECD Development Center collected information on energy volume (i.e. quantity) flows, and price information were collected by Gerard Malcolm at Purdue University. Gerard Malcolm then went on to 'incorporate' the energy volume and price information into the newly released Version 4 of the GTAP data base. The result was an interim release of a special data base called GTAP-4E in September and October of 1998. This Technical Paper documents the process of integrating energy volume and prices into the standard GTAP data base.

1. Introduction

It was observed that data relating to the energy flows in version 3 of the GTAP data base were of questionable reliability[1]. In particular the data on energy commodities seemed to be at variance with the energy statistics of the International Energy Agency (IEA). A project was undertaken at the Center for Global Trade Analysis (CGTA) with the financial support from the US Department of Energy to look into this issue. As part of the Project, information on energy volume (i.e. quantity) flows were collected by Christophe Complainville of the OECD Development Center, and price information were collected by Gerard Malcolm of CGTA. Gerard Malcolm then went on to 'incorporate' the energy volume and price information into the newly released Version 4 of the GTAP data base. The result was an interim release of a special data base called GTAP-4E in September and October of 1998. This Technical Paper documents the process of integrating energy volume and price information into the GTAP data base. The original design was by Gerard Malcolm. Subsequent modifications were made by Truong Truong.

2. Overview

The process of incorporating the energy *volume* and *price* information[2] into a standard GTAP data base can be rather complicated. Firstly, we note that GTAP is a data base expressed in terms of *value* units (i.e. dollars), and since value = volume * price, only *two* of the three sources of information (value, price, quantity or volume) can be regarded as independent of each other at any one time. Therefore, we cannot hope to incorporate *both* the energy volume *and* the price information into the GTAP data base without running into the risk of some internal inconsistencies arising from the problem of over-identification in data analysis. Secondly, the information in the volume and price data bases may not be complete at times. Therefore, it may become necessary to 'fill the gaps', by estimating the missing information, based on the known prices, volumes, or even the GTAP value data, and this can be seen as 'circular' adjustments in some cases.

Despite these difficulties, it seems there is potentially much to be gained from incorporating the additional information contained in the volume and price data bases into the GTAP data base. Firstly, we note that the price information contained in the price data base is the best currently available and documented. It has been meticulously collected from various sources[3],

---

1 Babiker and Rutherford (1997).

2 Babiker, Maskus, and Rutherford (1997).

3 The International Energy Agency (IEA), the World Bank (Survey of Asia's Energy Prices), the Organizacion Latino Americana de Energia (OLADE), the Asian Development Bank (ADB), the US Department of Energy (DOE), China Energy Databook (CED), and the Tata Energy Research Institute (TERI).

reconciled and checked for inaccuracies and inconsistencies. To complement this valuable information, there is also a volume data base to go with it, which has also been collected mainly from the International Energy Agency (IEA). As noted by Babiker and Rutherford (1977, p. 4): 'the GTAP dataset alone is not suitable by itself for assessing embodied carbon and modeling carbon and trade policies' because it is denominated only in financial terms. What is needed also is 'a dataset in which the physical units of fuel use by sector and region are accessible with associated prices which are jointly consistent with a balanced set of economic accounts'.

GTAP-4E is in fact a balanced set of economic accounts (in value terms) which are to be made 'jointly consistent' with the energy price, tax, and volume data set. This section provides an overview of the process which tries to fulfill this objective.

Figure 1 describes the whole process in simple graphical terms. The process is broken down into several 'steps', each performing one or more distinct tasks. What follows is a brief description of the tasks involved in each step. More detailed descriptions will be given later in the technical section and in the Appendix.

- Step 1: checks for consistency in the volume data base as collected by Christophe Complainville and make some minor adjustments.
- Step 2: gathers the price information collected from various and 'fills in' the missing information where necessary.
- Step 3: aggregates the price and volume information (*country* by *product* by *flow*) into a level which is consistent with the GTAP data base classification (*region* by *commodity* by *sector*); converts the (energy part of the) GTAP bilateral trade flow matrix from a *value* base to a *volume* base, using the energy *prices* produced from step 2; reconciles the derived energy volume trade flows with the total volume targets as contained in the volume data base; reconverts the (reconciled) volume trade flows into value trade flows using the energy price information and the margin information from GTAP; puts this energy trade flow back into the GTAP trade flow matrix.

  In addition to the modification of the energy part of the bilateral-trade flow in GTAP, Step 3 also estimates the values of final consumption, intermediate usage, imports and exports of, and taxes on, the energy commodities, based on the energy prices, taxes, and volume information. These are to be used as 'targets' in a later stage (Step 5) which will try to 'FIT'[4] the newly estimated energy values into the existing GTAP input-output data structure.

[4] See McDougall, R., Yu, W., and Malcolm, G. (1999), "Fitting", Chapter 20 in McDougall, R., Elbehri, A., and Truong, T., *Global Trade, Assistance, and Protection: the GTAP 4 Data Base*, Center for Global Trade Analysis, Purdue University.

- Step 4 prepares the single-region input-output (I/O) tables for use in the 'FIT' stage; calculates the 'targets' for the various macroeconomic variables (including taxes, subsidies, and protection information); calculates also the targets for energy variables (intermediate and final consumption levels, imports, exports, taxes).
- Step 5 carries out the 'FIT' process. This is a data updating process which involves adjusting the input-output tables to make them consistent with the aggregate base year macroeconomic information. Since the process now also involves updating the energy information, it is therefore called the FIT-E process.

3. Strengths and limitations of the existing process

The strengths of the current GTAP-4E data base and the incorporation process can be described as follows.

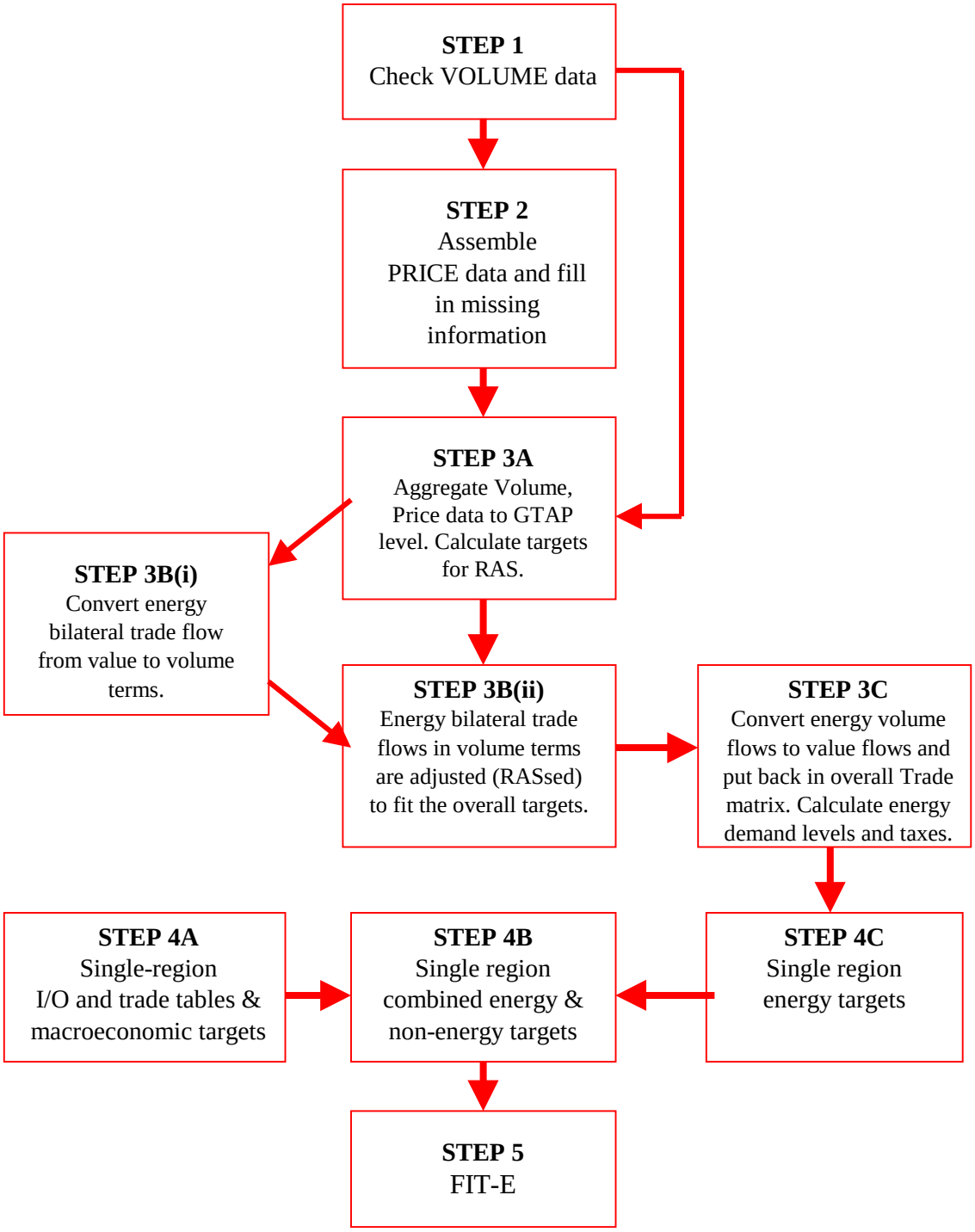
First, as mentioned at the beginning, it was the a priori belief that certain of the information on the energy commodities in the version 3 of the GTAP data base were of questionable reliability that led to the collection of additional information on energy volume, prices, and taxes. Although the GTAP version 3 data base is now replaced by version 4, there is no reason to doubt that by continuing to adjust the energy information in the GTAP data base with independent information collected from the IEA/OECD data sources, this will help to enrich rather than decrease the quality and content of the GTAP information A particular example may be given with respect to the case of energy taxation. Information on energy commodity taxes are missing or inaccurate for some important countries (such as Japan, many of the European countries) in the GTAP version 4 data base. When this information is further supplemented with the independently collected energy tax information, this resulted in a more complete and accurate picture of the actual situation. This is of importance for studies (such as those looking at energy or carbon taxation) which rely on a good description of the current situation for reliable projections into the future.

The GTAP-4E data base can be described as a version of the GTAP data base which is made to be more "consistent" with the energy prices and volume data bases. As a result, in studies which need to use the latter data bases for certain calculations (such as the CO_2 emission levels, which require the information on the volume data) it is essential that we use the GTAP 4E rather than the standard GTAP 4 data base as input data, to sure that some elements of consistency is maintained between the different types of data inputs into the model.

The limitations of the existing GTAP-4E data base are to some extent always inherent in the difficulty of trying to integrate three different data bases (*price*, *volume*, and input-output *value* data bases) which have only two degrees of freedom (since *value = price * volume*) into one system. There will always be some degree of inconsistencies in an 'over-identified' data base system, and therefore, unless the inconsistencies are gradually removed by resort to the

better-quality data which are availability over time, they will necessarily be removed by resort to some arbitrary and ad hoc processes in the mean time, and these also represent the current weaknesses of the data base.

Figure 1 – Overview



4. Technical aspects of GTAP-4E

This section describes the technical details underpinning the process of assembling and incorporating energy data (volume, prices, and taxes) into the GTAP data base. For the purpose of the description, the resultant data base is called GTAP-4E. The overall process is divided into steps. In each of these steps, information are read in (from on or more of the sub-directories) and then processed by one or more 'programs'. The results are then written onto some intermediate or final output files. Some of these programs may be just data manipulation routines (e.g. MODHAR), but they can also represent important theoretical 'models' which contain important assumptions underlying the process of data creation (e.g. EPA, EPB). All of these programs are written in the TABLO language, and compiled and executed by the GEMPACK software package[5]. Since some of the steps are not strictly sequential, care has to be taken to ensure that the most up-to-date file(s) are being used in a particular step. This is done automatically by a 'makefile' which calls on the Lahey MAKE utility program to perform on the managing job.

In Steps 1-2, the objective is to assemble the volume and price information into some format consistent with the GTAP classification. The volume data is provided independently by Christophe Complainville[6] of the OECD in a form which is more disaggregate than the GTAP data base: 135 countries, 7 energy "products"[7], and 32 energy 'flows"[8]. The price data are assembled from various sources: IEA (International Energy Agency), OLADE (Organizacion Latino Americana de Energia), ADB (Asian Development Bank), DOE (The US Department of Energy), CED (China Energy Databook), and WB (The World Bank Survey of Asia's Energy Prices). The price information are assembled at the 135-country level, but aggregated

---

[5] See Harrison and Pearson (1998)

[6] See Complainville (1997).

[7] The term 'product' is used in the price and volume data bases to distinguish it from the term 'commodity' which is used in GTAP. In the volume data base, there are 7 different types of energy products: {COL, ELY, GAS, HEAT, OIL, P_C, RENEW} standing for {Coal, Electricity, Natural gas, Heat, Crude oil, Petroleum and coal products, Renewable}. In the price and tax data base, there are 9 different energy products: {S, C, O, N, H, L, G, D, E} standing for {Steaming coal, Coking coal, Crude Oil, Natural gas, Heavy fuel oil (HFO), Light fuel oil (LFO), Gasoline, Diesel, Electricity}. The GTAP energy commodity set consists of the following: {COL, OIL, GAS, P_C, ELY}, standing for {coal, crude oil, natural and manufactured gas, petroleum and coal products, electricity}.

[8] For a list of the countries and energy flows, see Appendix.

from the 9-product level into five (GTAP-based) energy commodities and specified for four different 'user types'[9].

In assembling the volume and price information, certain key assumptions and adjustments are made. These are necessary to either 'correct' the inconsistencies in the original data base, or to fill in the missing values. A list of the key assumptions / adjustments is given in Table 1.

Step 3 is broken down into some minor steps that are not strictly sequential. The overall objective is to use the energy volume and price information to adjust the bilateral trade matrix (in value terms). To do this, first, the energy price information is used to convert the GTAP bilateral trade *value* flow into *volume* flow for the energy commodities. Next, the derived volume flows are then made consistent (through a procedure known as 'RAS'[10]) with the 'target' energy import and export totals as collected in Step 1. Finally, after the RAS process, the adjusted volume flows are then re-converted back into value flows using the price information and also trade margin information contained in GTAP.

Step 4 produces a set of 'target files' and 'shock files' for use in the so-called 'FIT' stage of the data creation process. FIT is a procedure for 'adjusting' the single region input-output (I/O) coefficients to be consistent with the latest information on various macroeconomic and trade variables which can be obtained separately for a particular reference year[11]. In the usual FIT stage of the GTAP data base creation process, important macroeconomic variables[12] are 'targeted' to a particular reference year level. When incorporating energy volume and price information into the GTAP data base, there are additional targets to be imposed for the energy variables. Some of these targets may come into conflict with the normal FIT targets and therefore some procedures for resolving these conflicts have to be devised. Step 5 carries out the usual FIT procedure, but with the additional set of energy targets (hence the name FIT-E).

In what follows, the above five Steps are to be described in greater details.

---

[9] {I, H, U, X} standing for {industry, household, utility, export}.

[10] United Nations (1973), Australian Bureau of Statistics (1990).

[11] McDougall, Yu, and Malcolm (1999).

[12] These variables are: exports, imports, household and government consumption expenditure, aggregate investment level, import tariffs, export subsidies, production taxes and production subsidies.

Table 1 – Important Assumptions / Adjustments made in various Steps

Step	Assumptions / Adjustments
1	Export and import volumes are reduced by an amount necessary to ensure domestic production is greater than export, i.e. having a positive 'export coverage'. (Table 2 lists the countries that have negative export coverage).
2A	Uniform physical conversion factors are used to convert all energy physical units into 'mtoe' (million tons of oil equivalent. A single conversion factor is used for all countries. Conversion formulae are used for both import, export, domestic prices, and taxes. Taxes are assumed to be specific, therefore, not subject to change over the years. Exchange rates versus the US dollar are used to convert all monetary units to USD. US domestic price indices for the years 1990-1995 are used to convert all price information in countries that do not have energy price indices to the 1995 base year.
2B	Missing (*non-export*) *user prices* (UPRC) are set equal to the average user price in 'similar' countries. Countries are split into four groups, according to their net trade position and general energy price level, and averages are taken for these four groups (HPEX, HPIM, LPEX, LPIM). For missing *export* user prices: (i) if export prices are available for some countries and for a particular product, assume that missing prices are equal to the global average, (ii) where no export prices are available, assume that export prices are equivalent to industry prices net of any user taxes. Missing basic prices are estimated using user taxes, user prices, and estimated margins: BPE (estimated basic price) = (user price – user tax) / (1 + margin). Missing taxes are treated as zero taxes in the calculation of BPE. Estimated margin is calculated as a weighted average of the known margins from certain countries (US, Japan) and the average of the margins on imported products. To calculate the latter, use the formula: margin on import = {(user price – user tax) / (basic import price)} – 1; where basic import price = import price + import duty. The overall *weighted*-average margin is then set equal to: 0.6* average of margins on imports + 0.2**Japanese* margin + 0.2**US* margin.

Table 1 – Important Assumptions / Adjustments made in various Steps (ctd.)

Step	Assumptions / Adjustments
2B (ctd.)	Prices are aggregated from 'product level' to 'commodity' level, using the mapping: {S, C}={COL}, {O}={OIL}, {N}={GAS}, {H, L, G, D}={P_C}, {E}={ELY}, and volume weights from DEB data base. User-specific basic price (CBPU) is aggregated over all users to get a 'unified' basic price (CBP). The margins are then re-estimated, using the relation: (implied dollar) margin = commodity user-price (CUP) – commodity tax (CTAX) – unified basic price (CBP), or (implied percentage) margin = [(CUP – CTAX)/CBP] –1.
3	Price and tax information of 44 out of the 135 countries are 'missing'. Use the prices and taxes of 'similar' countries for these countries (see Table 4). Volume weights are used to aggregate price and tax information from country level to region level. Regional energy prices are now used to convert part of the GTAP bilateral trade matrix (TRADE.har) to energy volume terms. They are then adjusted (RASsed) to fit the volume targets, and then reconverted back to value terms, using energy prices and margin information from GTAP.

Table 2: Countries with negative export coverage. (‘000 tons of oil equivalent (toe)).

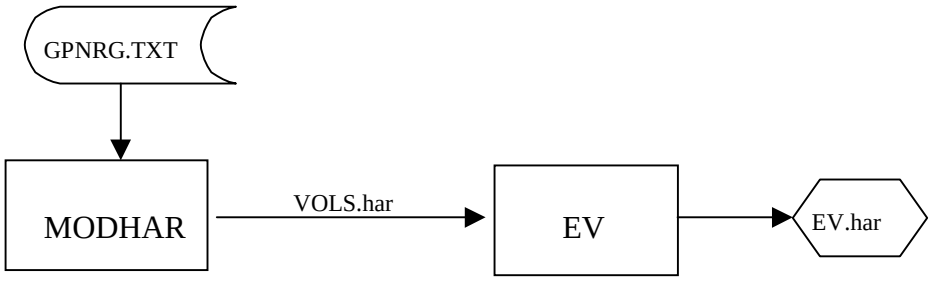
Country Code[(1)]	COL	OIL	GAS	P_C	ELY
7 PHL		-40.83			
8 SGP	-1.29	-62.77			
12 HKG				-7934.57	
17 NPL				-13.14	
22 ANT		-518.13			
39 ARG	-158.77				
46 DNK	-4.05				
47 SWE		-679.45			
48 FIN		-1.82			
50 BEL	-699.48	-2598.94			
53 GIB				-7.38	
54 GRC		-203.79			
56 ITA	-5.89				
57 LUX				-1.99	-21.03
58 NLD	-2559.06				
59 PRT	-49.31				
60 CHE	-5.33				
61 ISL					
62 NOR					
63 BGR					
64 CZE					
73 EST				-163.67	
76 KGZ				-0.99	
78 LVA				-64.41	
100 DZA	-4.53				
101 EGY	-292.93				
110 BEN				-6.94	
134 OTHERASIA				-141.66	
135 OTHERLATIN				-1158.63	

(1) For the full names of these country codes, see the Appendix.

Step 1: Checking and modifying Volume Data

In this step, the energy volume data created by Christophe Complainville in a text file (GPNRG.TXT) is read in using the routine MODHAR in GEMPACK, and re-written into a GEMPACK-consistent 'header array", or "har" file (VOLS.har). Export volumes are then set to positive values. Total production and usage (flows) are then checked for any imbalances. Production coverage of export is then defined as the difference between domestic production and the volume of exports. If this difference is negative (i.e. exports exceed production), then both export and import volumes will be reduced by the amount required to make export coverage a non-negative number. Results are then written onto a file (EV.har). A schematic representation of Step 1 is given below.

Step 1: Checking and modifying the Volume Data



The dimensions of the data base at this stage are in terms of:

- the set of countries (135) [13];
- the set of energy products (7):
  `COL, ELY, GAS, HEAT, OIL, P_C, RENEW;`
- and, the set of energy flows (32):
  `AGR, COL, CNS, CRP, DWE, ELY, EXPORTS, FPR, GAS, HEAT, I_S, IMPORTS, INDPROD, LUM, NEINTREN, NEOTHER, NETRANS, NFM, NMM, NONROAD, OIL, OME, OMF, OMN, OWNUSE, P_C, PPP, RENEW, ROAD, SER, TRN, TWL.`

[13] See Appendix for a list of these countries, and the description of the energy products and energy flows.

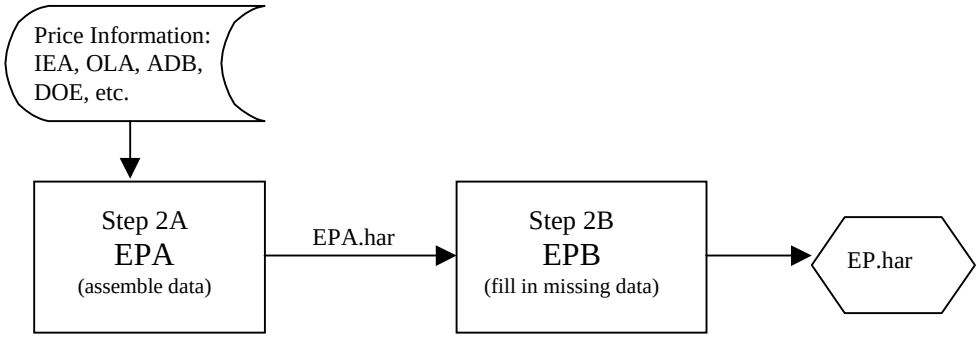
Comment on Step 1:

The process of reducing the export and import volumes by an amount to ensure domestic production is sufficient to cover all export seems ad hoc and unsatisfactory. It also results in the total volume of world export and import of energy commodities being smaller than what they actually are. Future revisions of this step should attempt to address this issue.

Step 2: Assembling Price Data and Filling in Missing Observations

This step consists of two parts: step 2A assembles and converts data into standard units, Step 2B then estimates and fills in the missing values. A schematic representation of this step is given below.

Step 2: Assembling Price Data and Filling in Missing values



Step 2A: Assembling and Conversion of Price Data

In this step, energy price data from various sources are assembled and converted into standard units, using information about exchange rate, time series price indices, and import duties. The program used in this step is EPA. The set of *.har* files used as inputs in this step are:

- IEA : The file containing price information from IEA sources;
- IEASET : The file containing set information on IEA data;
- OLA : The file containing price information from OLADE sources;
- OLASET : The file containing set information on OLADE data;
- ADB : Information from ADB's 1994 'Energy indicators from DC member countries';
- ADBSET : The file containing set information on ADB data;
- DOE : The file containing price information from DOE sources;
- DOESET : The file containing set information on DOE data;

- CED : The file containing price information from China Energy Databook;
- WBK : The file containing price information from World Bank's 'A survey of Asia's energy prices';
- IDUTY : The file containing information on import duty rates;
- CONFACS : The file containing information on conversion factors;
- XRATE : The file (IFSXR.har)[14] containing information on exchange rates;
- PINDEX : The file (INDEX.har) containing information on product-specific price indices for 1990-1995 for some countries[15];
- CTYSET : The file (WBSET.har) information on the set of countries with import duty data;
- DEB : The file (DB.har) containing detailed energy balances data;

The procedures involved in this step consist of the following:

- Price adjustment: conversion factors are used to convert collected prices to a standard monetary unit (US$), physical unit (toe), and time (1995). The input information which are partly stored in the EPA program, and partly read in from the input *.har* files to be used in the program are:
  - Exchange rates vs USD for all countries
  - Price indexes for 1990-1995
  - Intra-dimension conversion factors (e.g. lb-kg, gal-bbl, kWh-toe)
  - Inter-dimension product specific conversion factors (eg tonnes-toe for coal)
- Conversion formulae can be used for both import and export prices, and domestic prices. For taxes, conversion factors can be used in the same way, but they are assumed to be specific taxes, and therefore, not subject to conversion to a reference year by the use of price indices.
- A single conversion factor is used for all countries, and for import and domestic prices. This is because of the scarcity of up-to-date country specific information, and because price differences seem to be less than 10% between countries.
- Detailed Energy Balances (DEB) volume data from the IEA are used to decide when import prices are important, to categorize countries as net importers or exporters, to weight basic prices, and to aggregate price information from product level to commodity

---

[14] File names are changed between the coding stage (i.e. within the TAB file) and the processing stage (i.e. when the program is executed).

[15] See Malcolm, G. and Babiker, M. (1998) "Documentation for Energy Price/Tax Database", Center for Global Trade Analysis, April (http://www.agecon.purdue.edu/gtap/database/energy/status.htm) for more details on these price indexes.

level. Hopefully in future iterations, the commodity detail of the price database and the volume database will be equivalent, so that this process will become redundant.

- Special treatment is required for diesel, as both light fuel oil (LFO) and diesel are included in the DEB product 'GASDIES'. We regard the GASDIES flow to ROAD as diesel use by industry, and other flows as LFO use. We also remove the flow of electricity to utilities, as this is an outflow. Flows in Barbados are assumed to be the same as in Jamaica.

- China Energy Databook (CED) gives minemouth average cost for raw and washed coal. Units are yuan/metric ton for 1992 - this is case [k][16]: For ex-refinery oil products, prices are in US$/ton for 1994 - this is a special case because the standard unit for Gasoline (G) and Diesel (D) is bbl (a volume unit), but tonne is a mass unit. However, according to Table 11 of ADB (1994), the TOE/ton conversion factor for both G and D is 1.

- TERI yearbook gives average pithead value of coal, Rs/tonne - case [e]

- It would be useful to add another index to user prices to identify the data sources. Then we could explicitly rank the different sources or even take some combination of the different available prices. This has not been done to date. At the moment, whichever data source comes last is the preferred source by default.

- For IEA export price of crude oil, a number of spot prices are available. These are treated as export prices. Brent crude is used for Great Britain, West Texas Intermediate (WTI) for the US, Dubai for Saudi Arabia, Iranian Light for Iran, Ural for Russia, Minas for Indonesia, and Tapis for Malaysia. These are in US$/bbl in 1995 - case [b].

- For steam and coking coal, export prices are available for Australia, the USA and Canada. These are in US$/tonne - case [b]. For natural gas, prices of exports by pipeline from the Netherlands (NLD) and Norway (NOR) are available. These are in US$/toe for 1993 (NLD) and 1994 (NOR) - case [g]

- For petroleum products, spot prices at Rotterdam, New York, and Singapore are available for gasoline, gasoil, jet/kero, naphtha, low sulphur fuel oil (LSFO) and high sulphur fuel oil (HSFO).

- In terms of the products used here, the following mapping is used: Gasoline -> Gasoline, gasoil -> LFO, average of LSFO and HSFO -> HFO. These prices are all US$/bbl fob. For gasoline and LFO, this is case [b]. For HFO, we need to convert from bbl to tonnes.

---

[16] See Malcolm, G. and Babiker, M. (1998) for more details on these 'conversion cases'.

Conversion factors are given in IEA (1997, p.2)[17] : 6.49 bbl/tonne for LSFO at Rotterdam, 6.31 bbl/tonne for HSFO at Rotterdam and 6.46 bbl/tonne for HSFO at Singapore. A value of 6.4 bbl/tonne is applied.

- For CED data, Tables VI-11 to VI-33 give provincial retail electricity prices. An average of these is taken. Data is yuan/kWh for 1993, case [k]. Table VI-2 also provides an ex-Qinhuangdao port cost for Datong Lump Coal. This is considered to be product S - steam coal. This is derived from the minemouth coal cost 77.93 yuan/tonne plus all shipping and handling costs except (4) ocean shipping cost to Shanghai, which come to 112.6 yuan/tonne, giving a total fob price of 190.53 yuan/tonne. This is case [k] - yuan/tonne in 1992.

- For World Bank data (Survey of Asia's Energy Prices), there are user prices for petroleum products. However, these are in a variety of different units and to a large extent overlap with other data sources. Thus they have not been used.

- For TERI Indian Energy Yearbook[18]: see TERI table 7. Crude oil price is also given by ADB, but only for 1992. TERI source is preferred, denominated in Rs / tonne. Note that tonne is not the standard unit for crude oil, but assume tonne of crude oil = toe. This is case [d].

- Domestic User Product Taxes: OLADE[19] taxes are not collected from SIEE[20] data base, but from a Fax sent from OLADE. SIEE does not contain information on tax. For TERI taxes, TERI gives detailed information on taxes on coal in India. These include both ad valorem (central or state sales tax) and fixed (royalties and duty) elements. See pp20, 28. The simple average royalty is 103 Rs / tonne. Duty is 3.5 Rs / tonne. Either state (rate varies, some 9%) or central (rate of 4%) sales tax is levied - an average of 6% is assumed. Given the average pithead value of 442 Rs / unit (assumed to be tonne) in 1995, a 6% rate is equivalent to an amount of 28.7 Rs / tonne. The sum of these three therefore is applied: 135 Rs / tonne. This is case [e]. Note: IEA also gives tax information for Indian coal. It is superseded. TERI table 7 gives rates for royalty and cess, 528 and 900 Rs respectively. Denominated in Rs per tonne - this is case [d].

---

[17] IEA *Energy Prices and Taxes*, First Quarter, 1997, OECD, Paris, p. 2.

[18] TERI Energy Data Directory & Yearbook 1997/98, Tata Energy Research Institute, New Dehli, India, 1997.

[19] Organizacion Latino Americana de Energia.

[20] Systema de Informacion Economica-Energetica.

- To convert taxes to 1995 values, the question is: what is the best assumption to make about how tax level changes? Some taxes are ad valorem (notably the VAT components), but many energy taxes are specific taxes which do not necessarily change with price changes. So the assumption taken here is that tax levels do not change over time.

The output generated from the program EPA is a *.har* file (EPA.har) which displays the following information:

- ABP : actual basic price;
- VOL : Detailed Energy Balance (DEB) volume data;
- BPI : import-imputed basic price;
- UPRC: user price;
- UTAX : user tax;
- ucnt : count of user prices/taxes available (of 33)[21];
- icnt : count of import prices available (of 33);
- uct6 : count of 'Big Six' user prices/taxes available;
- ict6: count of 'Big Six' import prices available".
- USXR : US exchange rate;

Note that BPI is calculated as follows:

$$BPI = IPRC * (1 + IDUT / 100)$$

where IPRC is import price converted to 1995 US$/toe, and IDUT is import duty rate.

User prices and user taxes are read in from various sources and converted to standard units.

EPA.har is then used as input into another program EPB, which fills in the 'missing observations'. The procedures and assumptions for doing this are given below.

Step 2B: Estimating and filling in missing values

The objective of program EPB is to calculate the following:

CBP: Commodity Basic Price; (1995 $US/toe) (commodity[22] by country)

---

[21] The 33 single-country regions in the GTAP version 4 data base are: aus, nzl, jpn, kor, idn, mys, phl, sgp, tha, vnm, chn, hkg, twn, ind, lka, can, usa, mex, ven, col, arg, bra, chl, ury, gbr, deu, dnk, swe, fin, rus, tur, mar, zaf.

[22] 'commodity' is a term used to refer to the GTAP aggregation level, as distinguished from the original data sources 'product' level. The GTAP energy commodity set consist of the following: {COL, OIL, GAS, P_C, ELY}, standing for {coal, crude oil, natural and manufactured gas, petroleum and coal products, electricity}. The 'products' set varies for volume data and price data (see the Appendix). For volume data, the product set is: {COL, ELY, GAS, HEAT, OIL, P_C, RENEW}. For price data, the product set consists of: {S, C, O, N, H, L,

CBPU: Commodity Basic Price – user specific; (1995 $US/toe) (user by commodity by country)
CUP: Commodity User Price (1995 $US/toe); (user by commodity by country)
CTAX: Commodity User Tax (1995 $US/toe); (user by commodity by country)

To aggregate from the product level to the commodity level, the following mapping is used: {S, C}={COL}, {O}={OIL}, {N}={GAS}, {H, L, G, D}={P_C}, {E}={ELY}.

The first step in the program EPB is to fill in the missing values.

Missing user taxes (UTAX) are assumed to be zero. Missing (non-export) user prices (UPRC) are set equal to the average of the values in 'similar' countries[23]. Countries are split into 4 groups[24]: (HPEX, HPIM, LPEX, LPIM) and the average user prices for these groups, denoted by (HXAP, LMAP, LXAP, LMAP) are then used to fill in the missing values. Using this process, it was found that the only 'missing' user price left to fill in was the crude oil price for industry in HPEX countries, since none of these countries have an actual price for crude oil. As a result, to fill in this final gap, the average price for HPIM was used instead[25]. Missing *export* user prices can be filled in as follows: (i) for products where some export prices are available, assume that the missing prices are equal to the global average, (ii) where no export prices are available, assume that export prices are equivalent to industry prices net of any user taxes.

Missing basic prices are then estimated using the relation:

$$BPE = (UPRC - UTAX)/(1+ MARG)$$

where MARG is the estimated domestic margins. To estimate MARG, first calculate the margin on *imported* commodities (IMARG) by using the relation: IMARG = {(UPRC – UTAX) / (BPI)} – 1; where BPI is the estimated import basic price. Next, calculate a simple average of the IMARG's, and call this AVIMARG. The value for MARG is then set as a weighted-average of the AVIMARG, the margin for the US (USMARG) and the margin for

---

G, D, E}, standing for {Steaming coal, Coking coal, Crude Oil, Natural gas, HFO, LFO, Gasoline, Diesel, Electricity}.

[23] At first, other methods were attempted. For example, industry and utility prices were checked to see if they were similar. If they were, then missing value for one can be filled in by the availability of the other. It turned out that utility and industry prices were very much different, on average, by about 20-30%. Therefore, this method was not attempted.

[24] For more details about these four groups, see Malcolm and Babiker (1998).

[25] HPEX = {arg, gbr, nor, tun, zaf}, HPIM = {aut, bel, brb, che, deu, dnk, fin, fra, grc, hkg, ita, isl, irl, jam, jpn, lux, mar, nld, nic, pan, per, prt, png, svn, swe, sgp, twn, ury}.

Japan (JPNMRG) as follows: MARG = 0.6*AVIMARG + 0.2*USMARG + 0.2* JPNMARG.

Once the missing values are filled in, program EPB then aggregates the price information from the 'product' level to the 'commodity' level. For the aggregation process, two different sets of weights are used: one for aggregating *user* prices and taxes (UPRODWT), and one for aggregating *basic* prices (TPRODWT). The former is the (DEB) [26] volume weight over the appropriate set of products. The latter is the volume weights over the products and over all users.

Since the commodity basic prices are averaged over all users, the basic relationship between these basic prices, user taxes, user prices, and margins, is now 'destroyed', so we have to re-estimate the (implied, user-specific) margin again:

Implied (dollar, user-specific, commodity) margin = CUP – CTAX – CBP.
Implied (percentage, user-specific, commodity) margin = [(CUP – CTAX)/CBP] – 1.

Using the estimated commodity user price CUP, commodity user-tax CTAX and the averaged ('unified') commodity basic price CBP.

Results from the program EPB is written onto a *.har* file (EP.har) which contains, among other variables, the following sets of information:

- BPE: country-level estimated basic prices;
- CBP: unified commodity-level basic price;
- CUP: commodity-level user price;
- CTAX: commodity-level user tax;
- MRGD: commodity-level implied margin ($);
- MRGP: commodity-level implied margin (%);

The dimensions of the data base at this stage are still in terms of the 135 countries (see Step 1), but with the sets of energy 'products' and energy 'flows' now replaced by the set of GTAP commodities:

- {COL, OIL, GAS, P_C, ELY}

and, the set of users:

- {I, H, U, X.}

---

[26] Note that since zeros in the DEB volume data base do not correspond exactly to zeros in the Christophe Complainville's volume data base, we use the average of the two data base figures in the case of zero values in DEB volume.

Comment on Step 2B:

1. The assumption that all missing user taxes are treated as zero taxes can be modified. For example, an alternative is to treat only missing *export* user-taxes as zero taxes. For non-export users, rather than assuming that missing taxes are zeros, the BPE can first be calculated for those users who do not have missing taxes, then the average of the BPE for these users can then be used to fill in for the rest of the user(s). Only if *all* (non-export) users have missing taxes that missing taxes can be assumed to be zeros.

2. In the calculation of the estimated basic prices BPE's for the US and Japan, actual US and Japanese margins (USMARGIN, JPNMARGIN) should be used, instead of the 'weighted-average' margin (IMARG).

3. The process of averaging over the users to get the 'unified' commodity basic prices (CBP) 'destroys' the basic *user-specific* relationship between user-prices, user-taxes, and basic prices, and hence user-specific margins have to be re-estimated from the new CBP (*implied* margins). Currently, these implied margins are not used for any subsequent calculations or included explicitly as part of the data base. In future extensions of the GTAP model, where domestic margins can be represented, these margins should be utilised as part of the data base, to account for the differences in user-specific (agent's, or market) prices.

Step 3: Combining price and volume data into value data

The objective of Step 3 is to combine prices and volumes data into a value data set. To do this, it is necessary to break the step into four parts (3A, 3B(i), 3B(ii), 3C). In Step 3A, energy prices and volume data bases produced in Steps 1 and 2 are further aggregated into a GTAP-consistent format. Step 3B(i) then calculates the bilateral trade *volume* flow matrices for the energy commodities, (egy_vol.har) using GTAP's *value* flow matrix (TRADE.har) and the energy exporter-specific *fob* prices produced in Step 2. To ensure that the results are consistent with the *volume* data information, it is necessary to impose some constraints or 'targets' on these *derived* volume flows. These targets (egy_trgt.har) are calculated as the total export and import volumes as contained in the volume data base (EV.har). The process of imposing these targets on the volume flow bilateral trade matrices is called 'RAS' and is carried out in Step 3B(ii). To assist in the RAS operation, it is necessary to 'smear'[27] the matrices (i.e. replacing zeros with very small numbers and adjust the rest of the matrix to

[27] Since the program is written in percentage changes, a non-zero starting point is necessary to change the value of the variable, hence the need for 'smearing'. The 'smearing' process is thus necessary to enable the RAS operation to converge, even if the value used in the smearing process is very small (to improve on the accuracy of the RAS process).

preserve the column sum and row sum totals). After RAS is done, the bilateral energy trade matrices in volume term are then converted back into value terms and incorporated back into the GTAP data base. This is done in Step 3C.

Step 3A: Aggregating energy price and volume data, calculate energy targets for RAS

Energy prices, taxes, and volume data bases as produced in Steps 1 and 2 are available at the level of individual *countries*. The first step in making these data bases consistent with the GTAP format is to aggregate them into *regions*. However, before we can do this, note that only 91 out of the total of 135 countries with energy *volume* information have prices and/or taxes information available[28]. So it is necessary to 'fill in' the missing prices and taxes for the remaining 44 countries. This is done by firstly assuming that some countries are 'similar', therefore, energy prices and taxes information from one country can be used for the other. A list of countries with 'similar' energy prices and taxes is given in Table 3.

After the missing prices and taxes information are filled in, the next step is to aggregate the 135 countries into 45 regions. In aggregating the *basic* prices, we use the volumes of domestic usage of the particular energy commodity as weights (usage = domestic production + imports – exports). In aggregating *fob* prices, we use the volume of gross trade (exports + imports) as weights. The situation for taxes is a little bit different. For some countries, price information may be available but tax information may not be. One option is to use the GTAP regional tax information whenever the country tax information is missing. However, we decided against this approach. So the alternatives are: (i) whenever a region contains NO countries for which user tax information is available, that regional tax is set to zero, (ii) but if a subset of the countries within the region have some tax information, then the (user-specific) regional tax for that region will be set equal to the average of the taxes of these countries.

After the aggregation to the regional level is done, the next step is to aggregate from the energy "product" to the energy "commodity" level, and also to assign the various energy "flows" to GTAP industries or sectors.

Mapping from the energy volume[29] products to energy volume commodities is straightforward, since the energy commodity volume set[30] is simply a subset of the energy

---

[28] The list of (91) countries for which price information are available are: {AUS, AUT, ARG, ARE, AZE, BEL, BRA, BRB, BGD, BOL, BRN, CAN, CHE, CHL, CHN, COL, CRI, CZE, CUB, DEU, DNK, DOM, DZA, ECU, ESP, EST, FIN, FRA, GBR, GRC, GTM, GAB, GHA, HTI, HUN, HKG, HND, ITA, ISL, IRL, IND, IDN, IRN, JAM, JPN, KOR, KAZ, KGZ, KWT, LUX, LKA, LTU, LBN, LVA, LBY, MEX, MYS, MAR, NLD, NOR, NZL, NPL, NIC, NGA, PAN, PER, POL, PRT, PAK, PHL, PRY, PNG, QAT, RUS, ROM, SLV, SVK, SVN, SWE, SGP, SAU, TUR, TWN, THA, TTO, TUN, URY, USA, VEN, VNM, ZAF}. Some of these countries may not have tax information available, or that taxes are assumed to be zero.

[29] Note that the energy prices have already been aggregated to a commodity level in Step2.

product set[31]. To map the energy volume 'flows' into the various GTAP 'sectors', however, some assumptions are necessary. First, we note that 21 flows correspond exactly to an existing GTAP sector (or sectors): {AGR, COL, OIL, GAS, P_C, ELY, OMN, FPR, TWL, LUM, PPP, CRP, NMM, I_S, NFM, TRN, OME, OMF, CNS, SER, DWE). The rest, {EXPORTS, HEAT, IMPORTS, INDPROD, NEINTREN, NEOTHER, NETRANS, NONROAD, OWNUSE, RENEW, ROAD} need to be re-allocated.

First, we need to create a new 'flow' category called T_T which also corresponds to the GTAP sector T_T. Then, we need to allocate some of the original[32] flows to the new flow as follows:

- We assume that flows of all fuels into ROAD, NONROAD, NETRANS will now go into the T_T flow, with the exception that flow of P_C into ROAD is split 50% to households consumption[33] and 50% to the (T_T) flow.

- For the allocation of non-energy use to specific industries, NEINTREN is allocated to CRP (as per Annex-4 of the Proposed Design document) and NEOTHER is allocated to AGR.

- Originally the category OWNUSE was assumed to be intra-industry use for all energy commodities but this caused a contradiction when a commodity was predominantly imported (e.g. coal in Japan). Now, 'own-use' is an empty category, and we treat COL into COL as though COL into ELY, and OIL into OIL as though OIL into P_C . This left the two categories COL into COL and OIL into OIL as empty categories.

- GAS into GAS is split 50% as GAS into GAS (which includes gas distribution) and 50% as GAS into ELY.

- All the energy volume flows into DWE is treated as final consumption (FVOL)[34].

---

30 {COL, OIL, GAS, P_C, ELY}.

31 {COL, ELY, GAS, HEAT, OIL, P_C, RENEW}.

32 There are originally 32 flows in the volume data base.

33 Final consumption is not included in one of these 22 flows, but defined separately in an array called FVOL (see below).

34 See Annex A, Christophe Complainville and Dominique van der Mensbrugghe (1998) "Construction of an Energy Database for GTAP V4: Concordance with IEA Energy Statistics": here, "DWE" is described as

- HEAT and RENEW flows are not used.

- IMPORTS and EXPORTS flows are now included in separate arrays called MVOL and XVOL respectively.

After all the aggregations are done, the next step is to prepare the energy prices and volume date to be incorporated into the GTAP data base. This is done in several stages, some of which will be done in Step 3A. The overall objective is to use the energy prices and volume data to adjust the GTAP bilateral trade matrix (TRADE.har). To do this, Step 3A calculates the bilateral trade matrices for the energy commodities (egy_vol.har), but in *volume* terms, using the *value* information from TRADE.har and the exporter-specific *fob* prices from EP.har. It then calculates the energy 'target' matrices (egy_trgt.har) to prepare the ground for some adjustment procedure (called RAS).

Comment on Step 3A:

1. Some alternatives to the above scheme of allocation (of energy volume flows to GTAP industries) can be suggested as follows. First, to account for the fact that 'GAS" and "GDT" (gas distribution) are two distinct industries in GTAP, we can create a new flow category called GDT and allocate the existing volume flows to the new category according to the following split: GAS into GAS is divided as Y% GAS into GAS (natural gas) and (1-Y)% as GAS into GDT (gas distribution) with Y being inferred from the GTAP input-ouptut (I/O) value data base. Also, COL into GAS can be assumed as actually COL into GDT[35].

2. The percentage split of P_C into ROAD can be more variable according to the region: X% will go to households consumption and (1-X)% to the (T_T) category, where X is to be inferred from the regional I/O data base[36].

"residential". This means the flow into "DWE" becomes an empty flow (note also that in GTAP classification, the DWE sector is a separate category from the household final consumption).

[35] This is based on an observation of the GTAP I/O data base. For the case of AUS (Australia), for example, the GTAP I/O data shows significant flows of COL (as well as OIL and GAS) into GDT (of US95$406, 301, 146 millions respectively), but only insignificant amounts (US95$4.5, 3.4, 1.8 millions respectively) of GAS into GAS. The energy volume data base, in contrast, shows significant amount of GAS into GAS, but only insignificant amounts of COL or OIL into GAS. This shows that perhaps GAS into GAS in the volume data base (for AUS) is really GAS into GDT, and hence there can be a significant error of classification if the GDT (sector) is not distinguished from the GAS sector.

[36] In the final release of the Version 4E data base, we have in fact adopted this approach.

Table 3: List of ‘similar’ countries from which price/tax information is taken.

No.	Country	Using price/tax of country	No.	Country	Using price/tax of country	No.	Country	Using price/tax of country
1	AUS	AUS	46	DNK	DNK	91	JOR	lbn
2	NZL	NZL	47	SWE	SWE	92	KWT	KWT
3	JPN	JPN	48	FIN	FIN	93	LBN	LBN
4	KOR	KOR	49	AUT	AUT	94	OMN	are
5	IDN	IDN	50	BEL	BEL	95	QAT	QAT
6	MYS	MYS	51	ESP	ESP	96	SAU	SAU
7	PHL	PHL	52	FRA	FRA	97	SYR	lbn
8	SGP	SGP	53	GIB	gbr	98	YEM	are
9	THA	THA	54	GRC	GRC	99	MAR	MAR
10	VNM	VNM	55	IRL	IRL	100	DZA	DZA
11	CHN	CHN	56	ITA	ITA	101	EGY	lby
12	HKG	HKG	57	LUX	LUX	102	LBY	LBY
13	TWN	TWN	58	NLD	NLD	103	TUN	TUN
14	IND	IND	59	PRT	PRT	104	SAF	zaf
15	LKA	LKA	60	CHE	CHE	105	AGO	gha
16	BGD	BGD	61	ISL	ISL	106	MOZ	gha
17	NPL	NPL	62	NOR	NOR	107	TZA	gha
18	PAK	PAK	63	BGR	hun	108	ZMB	gha
19	CAN	CAN	64	CZE	CZE	109	ZWE	gha
20	USA	USA	65	HUN	HUN	110	BEN	lby
21	MEX	MEX	66	POL	POL	111	CIV	gha
22	ANT	brb	67	ROM	ROM	112	CMR	gha
23	CRI	CRI	68	SVK	SVK	113	COG	gha
24	CUB	CUB	69	SVN	SVN	114	ETH	lby
25	DOM	DOM	70	ARM	aze	115	GAB	GAB
26	GTM	GTM	71	AZE	AZE	116	GHA	GHA
27	HND	HND	72	BLR	ltu	117	KEN	gha
28	HTI	HTI	73	EST	EST	118	NGA	NGA
29	JAM	JAM	74	GEO	aze	119	SDN	gha
30	NIC	NIC	75	KAZ	KAZ	120	ZAR	gha
31	PAN	PAN	76	KGZ	KGZ	121	ALB	grc
32	SLV	SLV	77	LTU	LTU	122	BIH	grc
33	TTO	TTO	78	LVA	LVA	123	BRN	BRN
34	VEN	VEN	79	MDA	ltu	124	CYP	grc
35	COL	COL	80	RUS	RUS	125	HRV	gbr
36	BOL	BOL	81	TJK	rus	126	MKD	gbr
37	ECU	ECU	82	TKM	kaz	127	MLT	gbr
38	PER	PER	83	UKR	ltu	128	MMR	ind
39	ARG	ARG	84	UZB	kaz	129	PNG	PNG
40	BRA	BRA	85	TUR	TUR	130	PRK	kor
41	CHL	CHL	86	ARE	ARE	131	SEN	gha
42	URY	URY	87	BHR	are	132	SER	grc
43	PRY	PRY	88	IRN	IRN	133	OTHERAFRIC	gha
44	GBR	GBR	89	IRQ	irn	134	OTHERASIA	tha
45	DEU	DEU	90	ISR	gbr	135	OTHERLATIN	bra

TABLE 4: Concordance between Energy Volume flows and GTAP sectors

No.	Flow Code	GTAP commodity code / description	GTAP code No.(*)
1	AGR	PDR, WHT, GRO, V_F, OSD, C_B, PFB, OCR, CTL, OAP, RMK, WOL, FOR, FSH	1 –14
2	CNS	CNS	46
3	COL	COL	15
4	CRP	CRP	33
5	DWE	final consumption	50
6	ELY	ELY	43
7	EXPORTS	Exports	
8	FPR	CMT, OMT, VOL, MIL, PCR, SGR, OFD, B_T	19-26
9	GAS	GAS, GDT	17, 44
10	HEAT	not used	
11	I_S	I_S	35
12	IMPORTS	Imports	
13	INDPROD	total domestic production	
14	LUM	LUM	30
15	NEINTREN	all goes to CRP	Part 33
16	NEOTHER	all goes to AGR	Part 1-14
17	NETRANS	part to T_T and part to final consumption	Part 47
18	NFM	NFM	36
19	NMM	NMM	34
20	NONROAD	(non-road transport) all goes to T_T	Part 47
21	OIL	OIL	16
22	OME	FMP, ELE, OME	37,40,41
23	OMF	OMF	42
24	OMN	OMN	18
25	OWNUSE	COL into ELY, OIL into P_C, GAS into GAS, GDT and ELY.	
26	P_C	P_C	32
27	PPP	PPP	31
28	RENEW	not used	
29	ROAD	part to T_T and part to final consumption	Part 47
30	SER	WTR, OSP, OSG	45,48,49
31	TRN	MVH, OTN	38,39
32	TWL	TEX, WAP, LEA	27-29

(*) see header array “TTOE” in file map.har.

Note that only 21 flows from this Table correspond to the GTAP industries: {AGR, CNS, COL, CRP, DWE, ELY, FPR, GAS, I_S, LUM, NFM, NMM, OIL, OME, OMF, OMN, P_C, PPP, SER, TRN, TWL}. A new ‘flow’ category T_T is created which corresponds to GTAP industry T_T, and this makes up a total of 22 GTAP industry-based volume flows.

In summary, Step 3A (which uses a program called ECA) requires the following inputs:

- EP: price information from Step2;
- EV: volume information from Step 1;
- VOLS: the original volume data file used in Step 1 as an intermediate input, which also contains the information about country set, energy product set, and energy flows;
- TRADE: Mark Gehlhar's original trade matrix;
- map: file containing some mapping information;
- egy_sets: file containing a list of the Version 4 regions, 188 countries & some mapping information;
- gdp_cnt: file with GDP data for 188 countries;
- gset: file containing commodity set information;

To ensure that the TRADE.har matrix does not show 'infeasible' flows (such as the flow of electricity imports from Africa to Asia), we can 'intervene' at this stage to make some adjustment to the trade matrix before it is being used[37]. Program ECA then produces the following outputs:

- col_trgt.har: coal targets for exports and imports (both *fob*) by region;
- oil_trgt.har: oil targets for exports and imports by GTAP region;
- gas_trgt.har: natural gas targets for exports and imports by GTAP region;
- p_c_trgt.har: petroleum targets for exports and imports by GTAP region;
- ely_trgt.har: electricity targets for exports and imports by GTAP region;
- col_Vol.har: coal trade flow volumes (by GTAP regions);
- oil_Vol.har: oil trade flow volumes (by GTAP regions);
- gas_Vol.har: gas trade flow volumes (by GTAP regions);
- p_c_Vol.har: petroleum product trade flow volumes (by GTAP regions);
- ely_Vol.har: electricity trade flow volumes (by GTAP regions);

Step 3A also writes down the aggregated (5 energy commodities by 45 regions by 23 flows) *volume* information for use in the next few steps. This consists of the following arrays:

- IVOL: Intermediate use volume (mtoe) (region by energy commodities by flows);
- FVOL: Final consumption volume (mtoe) (region by energy commodities);
- XVOL: Export volume (mtoe) (region by energy commodities);
- MVOL: Import volume (mtoe) (region by energy commodities);
- DVOL: Volume of domestic production (mtoe) (region by energy commodities);
- DSHR: Domestic share in local usage;

[37] See the file Adjtrade.sti.

The aggregated price information (5 energy commodities by 45 regions by 4 user types) are written into the following arrays:

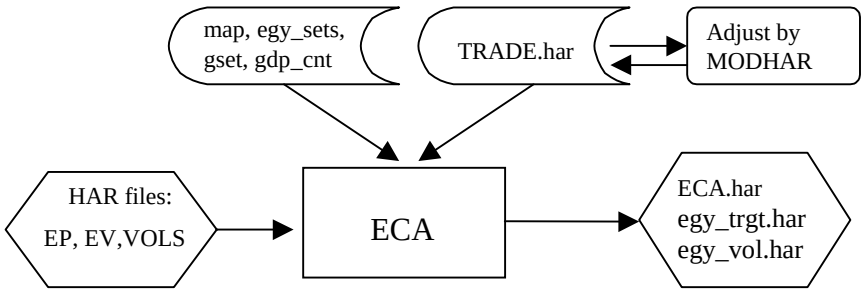
- RGBP: Regional basic price, using countries' domestic usage of energy as weights;
- RGTX: Regional user tax, using countries' domestic usage of energy as weights;

Other useful information are:

- psrb: Primary/secondary[38] basic price ratio for each region;
- avps: Average primary/secondary basic price over all regions;
- FOBP: Regional *fob* price, using countries' volume of gross trade as weights;
- GDP%: Share of composite region GDP;

A schematic representation of Step 3A is given below.

Step 3A: Aggregating energy data and calculate targets for RAS



Step 3B: Updating the trade flow matrix with energy commodity flows

From the previous step 3A, we recall that the following matrices were produced: (egy_vol.har) or volume trade flow matrices, derived from GTAP's TRADE.har and using the energy price information contained in the EP.har file; (egy_trgt.har) or the energy export and import volume matrices, calculated from the file EV.har, and is going to be used as targets on (egy_vol.har) in a RAS procedure which is performed in program RAS-E (adapted version of RAS for use with energy data). In Step 3B(ii), this RAS procedure is carried out, but before this is done, Step 3B(i) carries out a 'smearing' operation, to enable the RAS procedure to converge. In addition to these adjustments, at this crucial stage of producing the final bilateral trade flow matrices for the energy commodities, we can also intervene a second time to make final adjustments. This is done via the program MODHAR, and using an input file

[38] these are: coal/electricity, gas/electricity, crude oil/petroleum and coal products.

(adjVOL.sti) to modify the energy matrices (egy-vol.har) before RAS was performed on them[39].

After all these adjustments are made, the output from the RAS program (egy_vol.upd, representing the adjusted energy volume trade flows, and egy_tout.har, representing the scaled imports and exports row sums) are then used by the program CIF to convert the energy *volume* flow information back to *value* flow. The procedures involved are as follows:

a) First, the program CIF calculates the *cif/fob* margins for the energy commodities using the GTAP bilateral trade flow matrix TRADE.har. The margins are defined by the ratio (VIWS/VXWD)[40] of the values of imports and exports at world prices.

b) Next, apply the source-wise *fob* prices of the energy commodities (calculated from the previous step 3A) to the RASsed energy volume matrices to find the *fob* values of trade for each energy commodity (this is now the new values of VXWD, and it is equal to: *fob* price * volume).

c) Using the calculated margins from step a) above to derive the *cif* values for the energy commodities (this is the new VIWS, and it is equal to: *cif/fob* margin * new VXWD).

Note that no allowance for sales to the international transport sector has been made in the calculation of the *cif* values of the energy commodities. Because of this neglect of transport costs in the calculation of the bilateral trade for energy commodities, there can be an imbalance between the (revised) demand for margins (which is calculated as the sum over all regions of the difference between VIWS and VXWD) and the (original) supply of margins. To compensate for this, the supply of margins (VST) is changed pro rata to match new demand.

After this is done, the complete set of (new) VIWS, VXWD, and VST (which now contains the updated trade flow values for ALL commodities, not just the energy commodities) are written into an updated trade file (TRADE.upd)

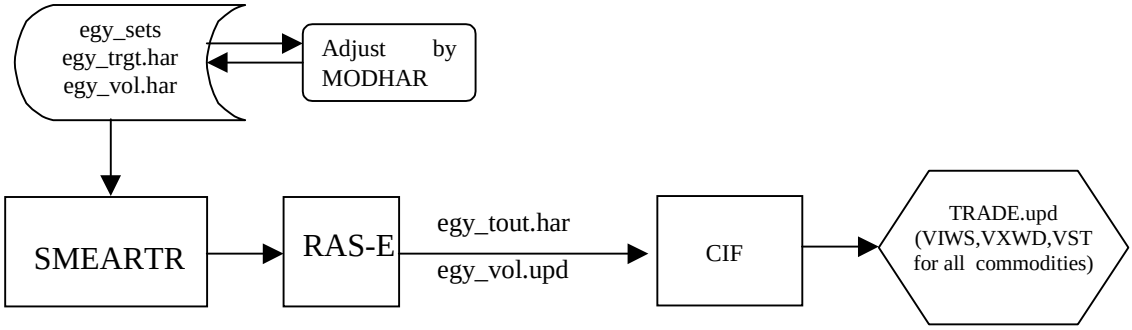
A schematic representation of Steps 3B is given below:

---

[39] This is to supplement the process carried out in Step 3A by MODHAR (using an input file adjTRADE.sti) to modify the TRADE.har matrix directly before it was used to produce the energy trade flow matrices egy_vol.har.

[40] See Hertel, T.W. (1997), pp371-396 for an explanation of these variables.

Steps 3B: Adjusting the bilateral trade (energy) flow matrices



Step 3C: Completion of the energy value matrices

Note that the output of Step 3A from program ECA are still in volume terms. Now it is ready to be converted into value terms. To do this, program ECB first reads in the energy (commodity by region by user) volume, price, and tax information from file ECA.har. It also reads in the (energy) updated portion of the trade flow matrix TRADE.upd from step 3B (which contains the 'new' series of VIWS and VXWD for energy commodities). In addition, it also reads in an import duty matrix (50 commodities by 45 regions by 45 regions) from file (tx04.har). ECB then converts the volume data base into value terms as follows. First, it calculates a few key prices. These are:

Import price:

IMPPRICE(r,i) = sum(s,reg,[VIWS(i,s,r) * (1 + RMT(i,s,r)/100)]) / IMVOL(r,i);

i.e. import price = value of *cif* imports (VIWS) from all regions with duty rate RMT (read in from file tx04.har) applied to these values and then divided by the import volume (IMVOL). The *cif* import value is read in from the file (TRADE.upd), and the import volume is read in from the file (ECA.har).

fob price:

FOBPRICE(r,i) = sum(s,reg,VXWD(i,r,s) / EXVOL(r,i));

i.e. *fob* price = value of exports at world prices (VXWD) to all destinations divided by export volume. The former is read in from the file (TRADE.upd), and the latter from the file (ECA.har).

Average domestic price:

AVPPRICE(r,i) = DOMWT(r,i) * REGBP(r,i) + (1-DOMWT(r,i)) * IMPRICE(r,i);

Where DOMWT(r,i) is the 'domestic weight' calculated as the proportion of energy usage (in volume term) which is from domestic sources, and REGBP is the 'unified' (i.e. averaged over all users) regional basic price calculated from Step 2 at the product level but aggregated into the commodity level in Step 3A.

These key prices are then used to convert the volume information into value information as follows:

VFA(r,i,n) = IIVOL(r,i,n) * AVPRICE(r,i);
VPA(r,i) = FCVOL(R,I) * AVPRICE(r,i);
TP(i,r) = (AVPRICE(r,i) + REGUTAX(r,i,"H")) / AVPRICE(r,i);
TF(i,j,r) = (AVPRICE(r,i) + REGUTAX(r,i,"I")) / AVPRICE(r,i);
TF(i,"ELY",r) = (AVPRICE(r,i) + REGUTAX(r,i,"U")) / AVPRICE(r,i);

Where VFA, VPA are the values of sales of energy commodities to firms (intermediate usage) and private households respectively, at agent's prices, IIVOL and FCVOL are these sales in volume terms; TP is the power of commodity tax on final consumption of energy commodities; TF is the power of commodity tax on intermediate usage of energy commodities; REGUTAX is the regional user-tax on these energy commodities, with "I" standing for the "industrial user", "U" the "utility" user (corresponding to industry sector j="ELY"), and "H" the household user.

Targets for energy commodities:

Program ECB also assesses the feasibility of targets for the energy commodities when they are subject to the FIT stage (see Step 5). Targets for energy commodities are sets equal to the values of domestic production of the energy commodities at producer prices:

VDP(r,i) = DPVOL(r,i) * REGBP(r,i) + EXVOL(r,i) * [FOBPRICE(r,i) - REGBP(r,i)];

If this value happens to be smaller than (or not sufficiently greater than) the value of the total energy costs (ECOSTS) within that industry, where:

ECOSTS (r,i) = sum(j,TF(j,i,r)*VFA(r,j,i));

then it would be infeasible to maintain the targets for the energy production level. To decide whether the targets are to be dropped or not, there are two decision levels, one 'manual', and one decided within the model itself. Manually, we can arbitrarily set a region r and energy commodity i as a 'non-target', by recording this information into a set, DROP(r,i) embedded

in the program ECB itself. This is to be used as a last resort. The other method is to record the value of the 'cost- surplus' defined as the difference:

costSURP(r,i) = VDP(r,i) – ECOSTS(r,i)

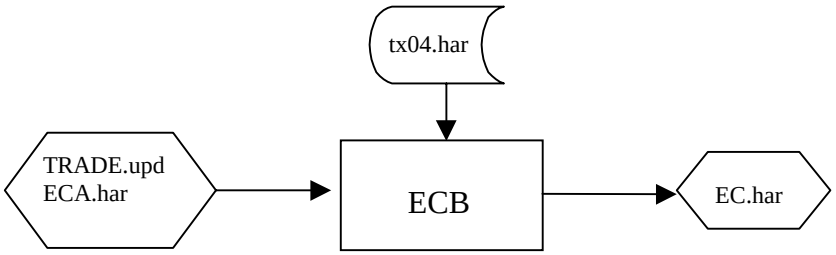
This information is then recorded (in an array called COST) and later read in by Program FITEPRPb or FITEPRPc of Step 6 where it will be used to determine if energy targeting will be dropped (more on this process, see Step 6 below).

Program ECB writes out a set of outputs to the file EC.har, which contains, among others, the following arrays:

VDP: value-of domestic production of energy commodities (at producer's price);
VFA: value of intermediate use;
VPA: value of final consumption;
TP: power of the tax on final consumption;
TF: power of the tax on intermediate use;
COST: Cost surplus = value of energy output ('target') less value of energy inputs.

These are important variables which will be used in later steps to calculate the targets for use in the FIT-E program, hence the accurate determination of these targets are of utmost importance for the successful operation of the FIT-E stage. A schematic representation of Step 3C is given below.

Step 3C: Energy value matrices

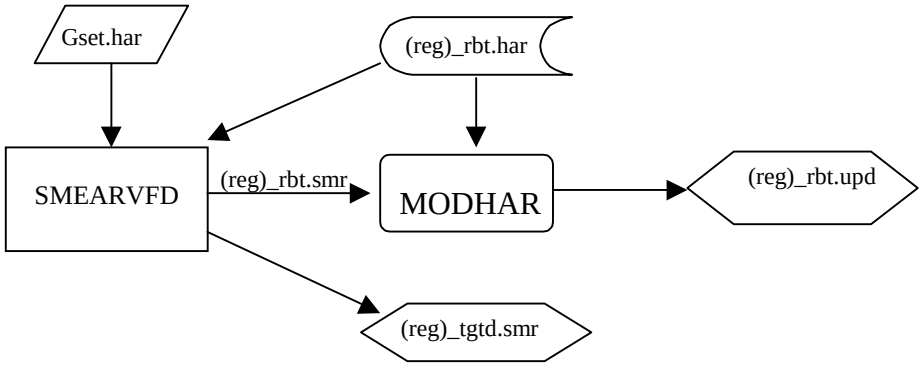


Step 4: Single region I-O data base and targets calculation

The first objective of Step 4 (Step 4A) is to 'smear' part of the original regional data base coefficient matrices, adjust for any imbalances resulting from that operation, and write the results back into an updated I-O matrix file. The smeared array is AI01 (intermediate usage of domestic products). To adjust for imbalances, any change in total intermediate use is compensated for by a change in labor use (array AI13), and any change in total intermediate

sales is compensated for by a change in sales to final consumption (array AI05). The arrays (AI01, AI05, AI13) are then written onto a file (reg)_rbt.smr. The row sums, column sums, and matrix total, are also written out as targets in a file (reg)_tgtd.smr. The program MODHAR then replaces the original arrays (AI01, AI05, AI13) in the regional I/O matrix file (reg)_rbt.har with the new arrays from the file (reg)_rbt.smr, and write the updated version to a new file (reg)_rbt.upd.

Step 4A: Smearing regional I-O data base



Step 4B then runs the usual 'fit-shock' module (called FTSH) which calculates the shocks to be used in the FIT-stage (step 5). The shocks are based on the initial 'base' value of, and the final 'target' for, a particular variable. Both of these can be read in from the input files, or calculated within the program. For example, the shocks to the macroeconomic variables: private consumption, private investment, and government consumption are calculated based on the initial values contained in the input file (reg)_rbt.upd, and the target values contained in files (reg)_cig.har. The shocks to the import tariff or the export subsidy are calculated by using the targets values from the input file (reg_ptt.har) and the initial values from the file (reg)_rbt.upd. The shocks to import values are calculated by comparing the target values as read in from the single-region bilateral trade matrix trade (reg_trd.upd) (which is derived from the world trade matrix TRADE.har by the program MODHAR) and the 'initial' values o fimports as read in from (reg_rbt.upd)[41].

The series of shock files produced by the program FTSH will consist of the following, for each region:

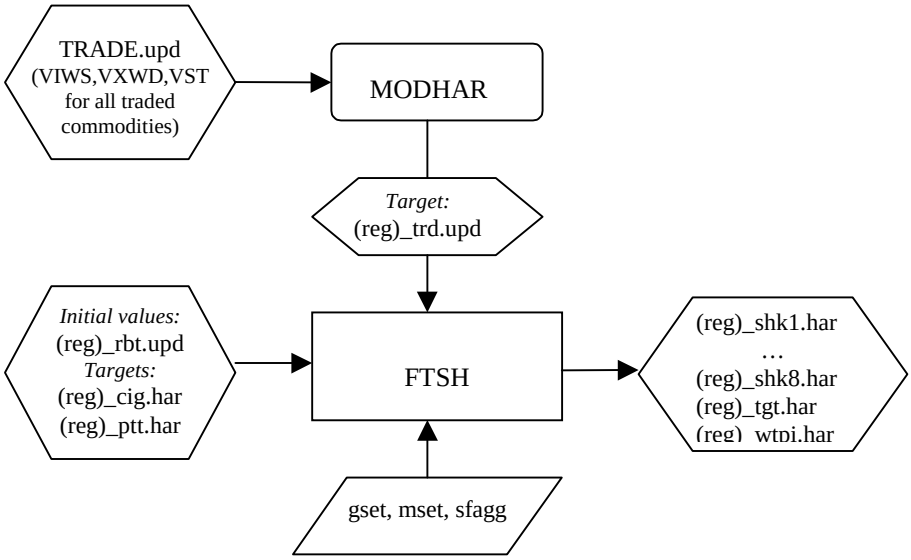
- (shk1.txt) shocks to scalar macroeconomic variables;
- (shk2.txt) shocks to export;
- (shk3.txt) shocks to import;
- (shk4.txt) shocks to domestic inventory;
- (shk5.txt) shocks to imported inventory;
- (shk6.txt) shocks to export tax/subsidy;

[41] Note that in the latest revision to the program FTSH, anti-dumping and price-undertaking data are ignored.

- (shk7.txt) shocks to import tariff;
- (shk8.txt) shocks to 'non commodity indirect taxes net of subsidy' (referred to as TSR);

The shock file (shk8.txt) is also written in another form, (wtpi.har), which is to be used in Step 4C to determine whether an energy commodity is to be targeted or not (see programs FITEPRPb and FITEPRPc below). All the targets and initial values are also written into a set of files with the suffix (_tgt.har). A schematic diagram of Step 4Bis given below.

Step 4B: (Non-energy) shocks and target files for FIT-E



Step 4C then produces the energy commodity targets and shocks. First, the program MODHAR breaks up the energy (price and volume) data base EC.har (produced in Step 3C) into single-region energy files (reg_EC.har). These files contain the following information for each region:

- DROP: information about which commodity will NOT be targeted in FIT-E;
- VDP: value of domestic production of energy commodities (at producers' prices)
- VFA: value of intermediate usage;
- VPA: value of final consumption;
- TP: power of the tax on final consumption;
- TF: power of the tax on intermediate use;
- COST: Cost surplus.

Program FITEPRPb then reads in the single-region energy data files (reg_EC.har), the updated single-region I/O tables (reg_rbt.upd) and the shock files (reg_wtpi.har - containing the shocks to the net non-commodity indirect taxes or TSR produced in Step 4B) to calculate

the shocks to the demands for energy commodities as well as to the taxes on these commodities. The series of shock files consist of the following:

- (shed.txt) shocks to energy intermediate demand;
- (shec.txt) shocks to energy private consumption demand;
- (sdit.txt) shocks to domestic intermediate tax;
- (siit.txt) shocks to imported intermediate tax;
- (sdct.txt) shocks to domestic final consumption tax;
- (sict.txt) shocks to imported final consumption tax;
- (tsr1.har): target for net non-commodity indirect taxes;

The last item, the *target* for net non-commodity indirect taxes (TSR) is calculated as follows. Using the information on the shocks to TSR read in from Step 4B (file reg_wtpi.har), the 'implicit target value' for TSR, called TSR1, is defined as:[42]

TSR1 = [(1 + wtpi/100) * (1+TSR/VOM)]*VOM1 – VOM1.

The values of VOM (value of output at market prices) are read in from the regional I/O tables (reg_rbt_upd). For example, labor cost is taken from the series AI13, net non-commodity indirect taxes TSR is taken from the series AI12, etc. The values of VOM1 (the 'implicit target value' for domestic output of energy commodities) are taken from the series VDP contained in the file EC.har (produced from ECB)[43].

Since there may be potential conflict between the energy commodities targeting and the usual targets in a FIT module, steps are taken to handle the case of so-called 'infeasibility' conditions. For example, an array called DROP is used to register the regions that experience infeasibility in energy targeting. DROP is specified 'manually' in program ECB. If a region is specified as 'dropped' (DROP(r) =1 for region r), then none of the energy industries (corresponding to the energy commodities) in that region will be targeted, i.e. 'partial' targeting is not the option here. However, within the program FITEPRPc[44], partial targeting

---

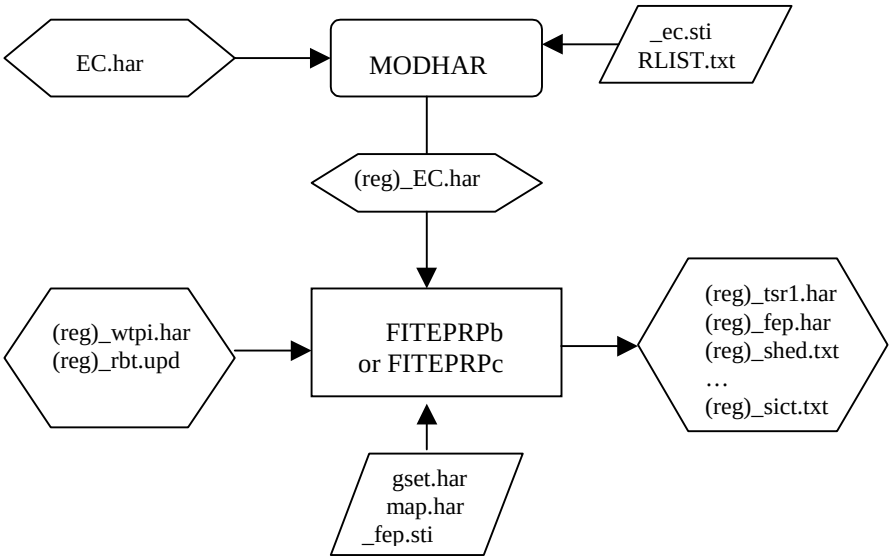
[42] Originally, the formula used was: TSR1*(1 + wtpi/100) = TSR*(VOM1/VOM) + VOM1*(wtpi/100). However, both formulae gave the same results.

[43] Note that the value of wtpi is calculated in Step 4B as follows: wtpi = 100*(target-base)/base, where: base = VOM/VOA using values read in from the regional I/O table, and target = 1/(1+production subsidy rate). Production subsidy rate is read in from the protection file (reg)_ptt.har. A positive production subsidy rate means VOM < VOA. TSR1 is treated as the target value for (VOM-VOA). If VDP is treated as the target value for VOM, then (VDP-TSR1) will imply a target value for VOA, which will then imply a maximum value for the total energy input costs.

[44] The difference between programs (FITEPRPc, FITEcost) on the one hand and programs (FITEPRPb, FITEboth) on the other, is that the latter includes 'manual' dropping of a particular region by reading in the set 'drop' (from program ECB) to drop *all* energy targeting for a particular region, whereas the former only drops a

can be used to drop an energy industry from targeting if it has the value of the 'cost surplus'[45] being less than the implicit target value of the net non-commodity indirect taxes (TSR1). Details on targeting is listed in Table 5. A schematic diagram of Step 4C is given below.

Step 4C: Energy shocks and target files for FIT-E



particular energy industry/commodity which does not satisfy the 'cost surplus' condition, i.e. which has cost surplus being less than TSR1.

[45] Contained in the 'COST' variable read in from Step 3. Cost surplus is defined: VDP - VFA, where VDP is the value of domestic production and VFA is the value of the energy intermediate inputs.

Table 5: Energy Targeting[1].

Variable	Base (I/O table)	Target (EC.har)	FITEPRPc FITEcost	FITEPRPb[2] FITEboth
Private consumption	VDPM+VIPM	VPA(i)	i∈ All_egy_comm	i∈ egy_comm
Intermediate usage	VDFM+VIFM	VFA(i,j)	i∈ All_egy_comm j∈ EV_comm	i∈ egy_comm j∈ EV_comm
Power of the Tax on private consumption	VDPA/VDPM VIPA/VIPM	TP(j) TP(j)	j∈ All_egy_ind[3]	j∈ All_egy_ind [3] or i∈ egy_ind[4]
Power of the Tax on intermediate usage	VDFA/VDFM VIFA/VIFM	TF(j,k) TF(j,k)	j∈ All_egy_ind[3] k∈ TRAD_comm	j∈ All_egy_ind[3] or j∈ egy_comm[4] k∈ TRAD_comm

(1) All_egy_comm: all energy commodities: {COL, OIL, GAS, P_C, ELY}.
Inf_egy_comm: infeasible energy commodities: drop(i) = 1, or cost surplus less than TSR1.
egy_comm = All_egy_comm – Inf_egy_comm.

All_egy_ind: all energy industries: {COL, OIL, GAS, P_C, ELY, GDT}.
Inf_egy_ind: infeasible energy industries corresponding to Inf_egy_comm.
egy_ind = All_egy_ind – Inf_egy_ind.

All_EV_comm: all energy volume flows (22) – see Table 4.
EV_comm = All_EV_comm – Inf_egy_comm.
Inf_EV_comm = All_EV_comm – EV_comm.

TRAD_comm: GTAP commodities/industries.

(2) Note that in the case of FITEPRPb and FITEboth, 'manual' dropping of target is specified in program ECB.tab which means the 'infeasible' set consists of those industries/commodities for which: (a) cost surplus less than TSR1 *and* (b) drop(i) = 1 (manual dropping). FITEPRPc and FITEcost uses only condition (a). This is the only difference between programs (FITEPRPb, FITEboth) on the one hand, and (FITEPRPc, FITEcost) on the other hand. The word 'cost' means dropping of energy input *costs* only for the infeasible industry, i.e. sales of the infeasible industry are still targeted (hence i∈ All_egy_comm rather than i∈ egy_comm), whereas 'both' means sales targets are also dropped (hence i∈ egy_comm). This applies to usage targets only (i.e. VPA and VFA). For taxes, it is useful to maintain the power of the tax targets, even if usage targets are dropped. Thus, see note 3 below.

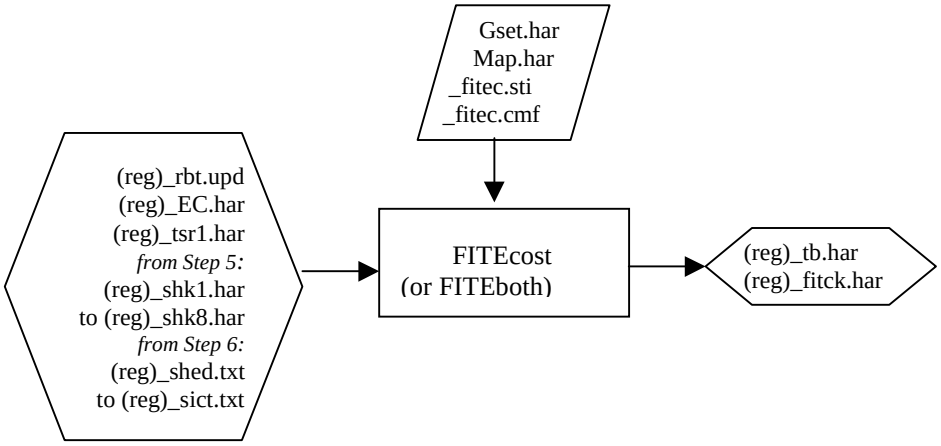
(3) Originally, this was defined in terms of the feasible industry set only (i.e. i∈ egy_ind), but there is no reason why all the tax targets are not maintained even if we drop some of the usage targets, (hence i∈ All_egy_comm) unless non-convergence occurs (see note 4, below).

(4) The smaller set is chosen if convergence is not achieved (for example, the case of ROW).

Step 5: FIT-E

FIT-E is a modified FIT program[46], which, in addition to the usual aspects of FIT, also targets the energy sector variables. The program reads in the energy-updated regional I-O tables (reg_rbt.upd), the regional energy data base (reg_ec.har), the target values for energy industries (reg_tsr1.har), and the various shock files. It is then run individually for each region. Depending on whether a region is classified as of type 'cost' (i.e. only energy input costs variables are to be dropped if they come into conflict with other targets), or of type 'both' (both the energy input costs and sales targets are dropped to help achieving the feasibility condition for the rest of the targets), the different versions of FIT-E: FITEcost, or FITEboth, will be used. The outputs of these programs are then the series of regional I/O tables as contained in the set of (reg)_tb.har files. These regional *.har* files are then used in the process of constructing the GTAP-E data base.

Step 5: FIT-E



[46] See James and McDougall (1993), and also: McDougall, Yu, and Malcolm (1999).

References

ADB (Asian Development Bank) (1994), *Energy Indicators of Developing Member Countries of ADB*, Energy and Industry Department, Asian Development Bank, Manila, The Philippines.

Australian Bureau of Statistics (1990), *Australian National Account, Input-Output Tables, 1986-87*, Cat. No. 5209.0, Canberra, Appendix A.

Babiker, M. H., and T. F. Rutherford (1997) "Input-output and general equilibrium estimates of embodied carbon: A data set and static framework for assessment", Working Paper 97-2, University of Colorado, Boulder.

Babiker, M. H., Maskus, K. E., and T. F. Rutherford, (1997) "Carbon Taxes and the Global Trading System," Working Paper 97-7, University of Colorado,Boulder.

China Energy Databook (CED) (1996), edited by Sinton, J., Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, CA.

Complainville, Christophe, (1997), "Proposed design for the GTAP Energy Data Base", Draft document 9/9/97.

DOE/EIA (US Department of Energy/Energy Information Administration) (1994), *End-use taxes: Current EIA Practices*, August, Washington, DC.

Harrison, W.J., and Pearson, K.D. (1998), *An Introduction to GEMPACK*, Gempack User Documentation, Release 5.2, Melbourne: Centre of Policy Studies and Impact Project, Monash University.

Hertel, T.W. (1997), *Global Trade Analysis – Modeling and Applications*, Cambridge University Press.

IEA (1997b), *Energy Balances of OECD Countries 1994-1995*, OECD, Paris.

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

James, M. , and McDougall, R. (1993), *Fit: An Input-Output Data Update Facility for SALTER*, SALTER Working Paper No. 17, Australian Industry Commission, Canberra.

Malcolm, G. and Babiker, M. (1998), "Documentation for Energy Price/Tax Database", Center for Global Trade Analysis, Purdue University, April.

McDougall, R., Yu, W., and Malcolm, G. (1999), "Fitting", Chapter 20 in McDougall, R., Elbehri, A., and Truong, T. (eds.), *Global Trade, Assistance, and Protection: the GTAP 4 Data Base*, Center for Global Trade Analysis, Purdue University.

McDougall, R., Elbehri, A., and Truong, T. (eds.) (1999), *Global Trade, Assistance, and Protection: the GTAP 4 Data Base*, Center for Global Trade Analysis, Purdue University.

TERI (Tata Energy Research Institute) (1997), *Energy Data Directory & Yearbook 1997/98*, New Dehli, India.

The World Bank (1994), *A Survey of Asia's Energy Prices*, World Bank Technical Paper Number 248, The World Bank, Washington, DC.

United Nations (1973), *Input-Output Tables and Analysis*, Studies in Methods, Series F, No. 4, Reve. 1, New York

APPENDIX

TABLE A1: List of countries with energy volume data and their short name codes.

1	Colombia	COL	46	Denmark	DNK	91	Jordan	JOR
2	Australia	AUS	47	Sweden	SWE	92	Kuwait	KWT
3	New Zealand	NZL	48	Finland	FIN	93	Lebanon	LBN
4	Japan	JPN	49	Austria	AUT	94	Oman	OMN
5	South Korea	KOR	50	Belgium	BEL	95	Qatar	QAT
6	Indonesia	IDN	51	Spain (Espanol)	ESP	96	Saudi Arabia	SAU
7	Malaysia	MYS	52	France	FRA	97	Syria	SYR
8	Philippines	PHL	53	Gibraltar	GIB	98	Yemen	YEM
9	Singapore	SGP	54	Greece	GRC	99	Morocco	MAR
10	Thailand	THA	55	Ireland	IRL	100	Algeria	DZA
11	Viet Nam	VNM	56	Italy	ITA	101	Egypt	EGY
12	China	CHN	57	Luxembourg	LUX	102	Libya	LBY
13	Hong Kong	HKG	58	Netherlands	NLD	103	Tunisia	TUN
14	Taiwan	TWN	59	Portugal	PRT	104	South Africa CU	SAF
15	India	IND	60	Switzerland	CHE	105	Angola	AGO
16	Sri Lanka	LKA	61	Iceland	ISL	106	Mozambique	MOZ
17	Bangladesh	BGD	62	Norway	NOR	107	Tanzania	TZA
18	Nepal	NPL	63	Bulgaria	BGR	108	Zambia	ZMB
19	Pakistan	PAK	64	Czech Republic	CZE	109	Zimbabwe	ZWE
20	Canada	CAN	65	Hungary	HUN	110	Benin	BEN
21	USA	USA	66	Poland	POL	111	Cote d'Ivoire	CIV
22	Mexico	MEX	67	Romania	ROM	112	Cameroon	CMR
23	Netherlands Antilles	ANT	68	Slovakia	SVK	113	Congo	COG
24	Costa Rica	CRI	69	Slovenia	SVN	114	Ethiopia	ETH
25	Cuba	CUB	70	Armenia	ARM	115	Gabon	GAB
26	Dominican Republic	DOM	71	Azerbaijan	AZE	116	Ghana	GHA
27	Guatemala	GTM	72	Belarus	BLR	117	Kenya	KEN
28	Honduras	HND	73	Estonia	EST	118	Nigeria	NGA
29	Haiti	HTI	74	Georgia	GEO	119	Sudan	SDN
30	Jamaica	JAM	75	Kazakhstan	KAZ	120	Zaire	ZAR
31	Nicaragua	NIC	76	Kyrgyzstan	KGZ	121	Albania	ALB
32	Panama	PAN	77	Lithuania	LTU	122	Bosnia Herzegovina	BIH
33	El Salvador	SLV	78	Latvia	LVA	123	Brunei	BRN
34	Trinidad & Tobago	TTO	79	Moldova	MDA	124	Cyprus	CYP
35	Venezuela	VEN	80	Russian Federation	RUS	125	Croatia	HRV
36	Bolivia	BOL	81	Tajikistan	TJK	126	Macedonia	MKD
37	Ecuador	ECU	82	Turkmenistan	TKM	127	Malta	MLT
38	Peru	PER	83	Ukraine	UKR	128	Myanmar	MMR
39	Argentina	ARG	84	Uzbekistan	UZB	129	Papua New Guinea	PNG
40	Brazil	BRA	85	Turkey	TUR	130	North Korea	PRK
41	Chile	CHL	86	United Arab Emirates	ARE	131	Senegal	SEN
42	Uruguay	URY	87	Bahrain	BHR	132	Serbia	SER
43	Paraguay	PRY	88	Iran	IRN	133	(a)	OTHERAFRIC
44	Great Britain	GBR	89	Irag	IRQ	134	(b)	OTHERASIA
45	Germany	DEU	90	Israel	ISR	135	(c)	OTHERLATIN

(a) (b) (c) : See Table A3 for details.

TABLE A2: List of countries with energy volume data – in alphabetical order of codes.

105	Angola	AGO	116	Ghana	GHA	94	Oman	OMN
121	Albania	ALB	53	Gibraltar	GIB	19	Pakistan	PAK
23	Netherlands Antilles	ANT	54	Greece	GRC	32	Panama	PAN
86	United Arab Emirates	ARE	27	Guatemala	GTM	38	Peru	PER
39	Argentina	ARG	13	Hong Kong	HKG	8	Philippines	PHL
70	Armenia	ARM	28	Honduras	HND	129	Papua New Guinea	PNG
2	Australia	AUS	125	Croatia	HRV	66	Poland	POL
49	Austria	AUT	29	Haiti	HTI	130	North Korea	PRK
71	Azerbaijan	AZE	65	Hungary	HUN	59	Portugal	PRT
50	Belgium	BEL	6	Indonesia	IDN	43	Paraguay	PRY
110	Benin	BEN	15	India	IND	95	Qatar	QAT
17	Bangladesh	BGD	55	Ireland	IRL	67	Romania	ROM
63	Bulgaria	BGR	88	Iran	IRN	80	Russian Federation	RUS
87	Bahrain	BHR	89	Irag	IRQ	104	South Africa CU	SAF
122	Bosnia Herzegovina	BIH	61	Iceland	ISL	96	Saudi Arabia	SAU
72	Belarus	BLR	90	Israel	ISR	119	Sudan	SDN
36	Bolivia	BOL	56	Italy	ITA	131	Senegal	SEN
40	Brazil	BRA	30	Jamaica	JAM	132	Serbia	SER
123	Brunei	BRN	91	Jordan	JOR	9	Singapore	SGP
20	Canada	CAN	4	Japan	JAP	33	El Salvador	SLV
60	Switzerland	CHE	75	Kazakhstan	KAZ	68	Slovakia	SVK
41	Chile	CHL	117	Kenya	KEN	69	Slovenia	SVN
12	China	CHN	76	Kyrgyzstan	KGZ	47	Sweden	SWE
111	Cote d'Ivoire	CIV	5	South Korea	KOR	97	Syria	SYR
112	Cameroon	CMR	92	Kuwait	KWT	10	Thailand	THA
113	Congo	COG	93	Lebanon	LBN	81	Tajikistan	TJK
1	Colombia	COL	102	Libya	LBY	82	Turkmenistan	TKM
24	Costa Rica	CRI	16	Sri Lanka	LKA	34	Trinidad & Tobago	TTO
25	Cuba	CUB	77	Lithuania	LTU	103	Tunisia	TUN
124	Cyprus	CYP	57	Luxembourg	LUX	85	Turkey	TUR
64	Czech Republic	CZE	78	Latvia	LVA	14	Taiwan	TWN
45	Germany	DEU	99	Morocco	MAR	107	Tanzania	TZA
46	Denmark	DNK	79	Moldova	MDA	83	Ukraine	UKR
26	Dominican Republic	DOM	22	Mexico	MEX	42	Uruguay	URY
100	Algeria	DZA	126	Macedonia	MKD	21	USA	USA
37	Ecuador	ECU	127	Malta	MLT	84	Uzbekistan	UZB
101	Egypt	EGY	128	Myanmar	MMR	35	Venezuela	VEN
51	Spain (Espanol)	ESP	106	Mozambique	MOZ	11	Viet Nam	VNM
73	Estonia	EST	7	Malaysia	MYS	98	Yemen	YEM
114	Ethiopia	ETH	118	Nigeria	NGA	120	Zaire	ZAR
48	Finland	FIN	31	Nicaragua	NIC	108	Zambia	ZMB
52	France	FRA	58	Netherlands	NLD	109	Zimbabwe	ZWE
115	Gabon	GAB	62	Norway	NOR	133	(a)	OTHERAFRIC
44	Great Britain	GBR	18	Nepal	NPL	134	(b)	OTHERASIA
74	Georgia	GEO	3	New Zealand	NZL	135	(c)	OTHERLATIN

(a) (b) (c) : See Table A3 for details.

TABLE A3: List of “OTHER” countries.

OTHERAFRIC		OTHERASIA		OTHERLATIN	
Burundi	BDI	Afghanistan	AFG	Antigua & Barbuda	ATG
Burkina Faso	BFA	Bhutan	BTN	Bahamas	BHS
Botswana	BWA	Fiji	FJI	Barbados	BRB
Central African Rep.	CAF	Cambodia	KHM	Comoros	COM
Cape Verde	CPV	Kiribati	KIR	Dominica	DMA
Djibouti	DJI	Lao (PDR)	LAO	Guinea	GIN
Eritrea	ERI	Mongolia	MNG	Guinea-Bissau	GNB
Gambia	GMB	Nauru	NAU	Grenada	GRD
Equatorial Guinea	GNQ			Guyana	GUY
Liberia	LBR			St Christopher (St Kitts) & Nevis	KNA
Lesotho	LSO			Saint Lucia	LCA
Madagascar	MDG			Niger	NER
Maldives	MDV			Solomon Is.	SLB
Mali	MLI			Sierra Leone	SLE
Mauritania	MRT			Sao Tome & Principe	STP
Mauritius	MUS			Suriname	SUR
Malawi	MWI			Seychelles	SYC
Namibia	NAM			Chad	TCD
Rwanda	RWA			Tonga	TON
Somalia	SOM			Saint Vincent & the Grenadines	VCT
Swaziland	SWZ			Samoa	WSM
Togo	TGO				
Uganda	UGA				
Vanuatu	VUT				

TABLE A4: Concordance between energy volume countries and GTAP Version 4 regions

Volume country	V4 region	Volume country	V4 region	Volume country	V4 region	Volume country	V4 region
ADO	ROW	DZA	RNF	LBN	RME	RUS	FSU
AFG	ROW	ECU	RAP	LBR	RSS	RWA	RSS
AGO	RSA	EGY	RNF	LBY	RNF	SAU	RME
ALB	ROW	ERI	RSS	LCA	CAM	SDN	RSS
ARE	RME	ESP	REU	LEI	EFT	SEN	RSS
ARG	ARG	EST	FSU	LKA	LKA	SGP	SGP
ARM	FSU	ETH	RSS	LSO	SAF	SLB	ROW
ATG	CAM	FIN	FIN	LTU	FSU	SLE	RSS
AUS	AUS	FJI	ROW	LUX	REU	SLV	CAM
AUT	REU	FRA	REU	LVA	FSU	SMR	REU
AZE	FSU	GAB	RSS	MAR	MAR	SOM	RSS
BDI	RSS	GBR	GBR	MDA	FSU	STP	RSS
BEL	REU	GEO	FSU	MDG	RSS	SUR	RSM
BEN	RSS	GHA	RSS	MDV	RAS	SVK	CEA
BFA	RSS	GIN	RSS	MEX	MEX	SVN	CEA
BGD	RAS	GMB	RSS	MKD	ROW	SWE	SWE
BGR	CEA	GNB	RSS	MLI	RSS	SWZ	SAF
BHR	RME	GNQ	RSS	MLT	ROW	SYC	RSS
BHS	CAM	GRC	REU	MMR	ROW	SYR	RME
BIH	ROW	GRD	CAM	MNG	ROW	TCD	RSS
BLR	FSU	GTM	CAM	MON	REU	TGO	RSS
BOL	RAP	GUY	RSM	MOZ	RSA	THA	THA
BRA	BRA	HKG	HKG	MRT	RSS	TJK	FSU
BRB	CAM	HND	CAM	MUS	RSS	TKM	FSU
BRN	ROW	HRV	ROW	MWI	RSA	TON	ROW
BTN	RAS	HTI	CAM	MYS	MYS	TTO	CAM
BWA	SAF	HUN	CEA	NAM	SAF	TUN	RNF
CAF	RSS	IDN	IDN	NAU	ROW	TUR	TUR
CAN	CAN	IND	IND	NER	RSS	TUV	ROW
CHE	EFT	IRL	REU	NGA	RSS	TWN	TWN
CHL	CHL	IRN	RME	NIC	CAM	TZA	RSS
CHN	CHN	IRQ	RME	NLD	REU	UGA	RSS
CIV	RSS	ISL	EFT	NOR	EFT	UKR	FSU
CMR	RSS	ISR	RME	NPL	RAS	URY	URY
COG	RSS	ITA	REU	NZL	NZL	USA	USA
COL	COL	JAM	CAM	OMN	RME	UZB	FSU
COM	RSS	JOR	RME	PAK	RAS	VCT	CAM
CPV	RSS	JPN	JPN	PAN	CAM	VEN	VEN
CRI	CAM	KAZ	FSU	PER	RAP	VNM	VNM
CUB	CAM	KEN	RSS	PHL	PHL	VUT	ROW
CYP	ROW	KGZ	FSU	PNG	ROW	WSM	ROW
CZE	CEA	KHM	ROW	POL	CEA	YEM	RME
DEU	DEU	KIR	ROW	PRK	ROW	YUG	ROW
DJI	RSS	KNA	CAM	PRT	REU	ZAF	SAF
DMA	CAM	KOR	KOR	PRY	RSM	ZAR	RSS
DNK	DNK	KWT	RME	QAT	RME	ZMB	RSA
DOM	CAM	LAO	ROW	ROM	CEA	ZWE	RSA

TABLE A5: Concordance between GTAP Version 4 regions and energy volume countries

Volume country	V4 region	Volume country	V4 region	Volume country	V4 region	Volume country	V4 region
ARG	ARG	FSU	LTU	RME	LBN	RSS	CMR
AUS	AUS	FSU	LVA	RME	OMN	RSS	COG
BRA	BRA	FSU	MDA	RME	QAT	RSS	COM
CAM	ATG	FSU	RUS	RME	SAU	RSS	CPV
CAM	BHS	FSU	TJK	RME	SYR	RSS	DJI
CAM	BRB	FSU	TKM	RME	YEM	RSS	ERI
CAM	CRI	FSU	UKR	RNF	DZA	RSS	ETH
CAM	CUB	FSU	UZB	RNF	EGY	RSS	GAB
CAM	DMA	GBR	GBR	RNF	LBY	RSS	GHA
CAM	DOM	HKG	HKG	RNF	TUN	RSS	GIN
CAM	GRD	IDN	IDN	ROW	ADO	RSS	GMB
CAM	GTM	IND	IND	ROW	AFG	RSS	GNB
CAM	HND	JPN	JPN	ROW	ALB	RSS	GNQ
CAM	HTI	KOR	KOR	ROW	BIH	RSS	KEN
CAM	JAM	LKA	LKA	ROW	BRN	RSS	LBR
CAM	KNA	MAR	MAR	ROW	CYP	RSS	MDG
CAM	LCA	MEX	MEX	ROW	FJI	RSS	MLI
CAM	NIC	MYS	MYS	ROW	HRV	RSS	MRT
CAM	PAN	NZL	NZL	ROW	KHM	RSS	MUS
CAM	SLV	PHL	PHL	ROW	KIR	RSS	NER
CAM	TTO	RAP	BOL	ROW	LAO	RSS	NGA
CAM	VCT	RAP	ECU	ROW	MKD	RSS	RWA
CAN	CAN	RAP	PER	ROW	MLT	RSS	SDN
CEA	BGR	RAS	BGD	ROW	MMR	RSS	SEN
CEA	CZE	RAS	BTN	ROW	MNG	RSS	SLE
CEA	HUN	RAS	MDV	ROW	NAU	RSS	SOM
CEA	POL	RAS	NPL	ROW	PNG	RSS	STP
CEA	ROM	RAS	PAK	ROW	PRK	RSS	SYC
CEA	SVK	REU	AUT	ROW	SLB	RSS	TCD
CEA	SVN	REU	BEL	ROW	TON	RSS	TGO
CHL	CHL	REU	ESP	ROW	TUV	RSS	TZA
CHN	CHN	REU	FRA	ROW	VUT	RSS	UGA
COL	COL	REU	GRC	ROW	WSM	RSS	ZAR
DEU	DEU	REU	IRL	ROW	YUG	SAF	BWA
DNK	DNK	REU	ITA	RSA	AGO	SAF	LSO
EFT	CHE	REU	LUX	RSA	MOZ	SAF	NAM
EFT	ISL	REU	MON	RSA	MWI	SAF	SWZ
EFT	LEI	REU	NLD	RSA	ZMB	SAF	ZAF
EFT	NOR	REU	PRT	RSA	ZWE	SGP	SGP
FIN	FIN	REU	SMR	RSM	GUY	SWE	SWE
FSU	ARM	RME	ARE	RSM	PRY	THA	THA
FSU	AZE	RME	BHR	RSM	SUR	TUR	TUR
FSU	BLR	RME	IRN	RSS	BDI	TWN	TWN
FSU	EST	RME	IRQ	RSS	BEN	URY	URY
FSU	GEO	RME	ISR	RSS	BFA	USA	USA
FSU	KAZ	RME	JOR	RSS	CAF	VEN	VEN
FSU	KGZ	RME	KWT	RSS	CIV	VNM	VNM

TABLE A6: A list of energy products in the energy volume data base and their concordance with the GTAP energy commodities

No.	Energy product Code	GTAP energy commodity code	Description
1	COL	COL	Coal
2	OIL	OIL	Crude oil
3	GAS	GAS	Natural and manufactured gas
4	P_C	P_C	Petroleum, coal products
5	ELY	ELY	Electricity
6	HEAT		Heat
7	RENEW		Renewable

TABLE A7: A list of energy products in the energy price data base and their concordance with the GTAP energy commodities

No.	Energy product Code	GTAP energy commodity code	Description
1	S	COL	Steaming coal
2	C	COL	Coking coal
3	O	OIL	Crude oil
4	N	GAS	Natural gas
5	H	P_C	Heavy fuel oil
6	L	P_C	Light fuel oil
7	G	P_C	Gasoline
8	D	P_C	Diesel
9	E	ELY	Electricity

TABLE A8: A list of energy flow codes and their descriptions(1)

No.	Flow Code	Description
1	AGR	Agriculture
2	CNS	Construction
3	COL	Coal
4	CRP	Chemical and petrochemical
5	DWE	Dwellings
6	ELY	Electricity
7	EXPORTS	Exports
8	FPR	Food and tobacco
9	GAS	Gas
10	HEAT	Heat
11	I_S	Iron and steel
12	IMPORTS	Imports
13	INDPROD	Indigenous production
14	LUM	Wood and wood products
15	NEINTREN	Non-energy use: in industry/transf./energy
16	NEOTHER	Non-energy use in other sectors (agriculture, commercial & public services, residential)
17	NETRANS	Non-energy use in transport
18	NFM	Non ferrous metals
19	NMM	Non metallic minerals
20	NONROAD	Other (non road) transport
21	OIL	Oil
22	OME	Machinery
23	OMF	Other manufacturing
24	OMN	Mining an quarrying
25	OWNUSE	Own use
26	P_C	Petroleum and coal products
27	PPP	Paper, pulp and print
28	RENEW	Renewable
29	ROAD	Road transport
30	SER	Commercial and public services
31	TRN	Transport equipment
32	TWL	Textile and leather

(1) Only 21 flows from this Table correspond to GTAP industries: {AGR, CNS, COL, CRP, DWE, ELY, FPR, GAS, I_S, LUM, NFM, NMM, OIL, OME, OMF, OMN, P_C, PPP, SER, TRN, TWL}. A new 'flow' category T_T is created which corresponds to GTAP industry T_T. This makes up 22 total flows.

TABLE A9: Conversion factors for energy

To:	TJ	Gcal	Mtoe	Mbtu	GWh
From:	multiply by:				
TJ	1	238.8	2.388×10^{-5}	947.8	0.2778
Gcal	4.1868×10^{-3}	1	10^{-7}	3.968	1.163×10^{-3}
Mtoe	4.1868×10^{-4}	10^{7}	1	3.968×10^{7}	11630
Mbtu	1.0551×10^{-3}	0.252	2.52×10^{-8}	1	2.931×10^{-4}
GWh	3.6	860	8.6×10^{-5}	3412	1

Source: IEA (1997), *Energy Balances of OECD Countries 1994-1995*, OECD, Paris, pp xv-xvi

TABLE A10: Conversion factors for Mass

To:	Kg	T	lt	St	lb
From:	multiply by:				
Kilogram(kg)	1	0.001	9.84×10^{-4}	1.102×10^{-3}	2.2046
Tonne (t)	1000	1	0.984	1.1023	2204.6
Long ton (lt)	1016	1.016	1	1.120	2240.0
Short ton (st)	907.2	0.9072	0.893	1	2000.0
Pound (lb)	0.454	4.54×10^{-4}	4.46×10^{-4}	5.0×10^{-4}	1

Source: IEA (1997), *Energy Balances of OECD Countries 1994-1995*, OECD, Paris, pp xv-xvi

TABLE A11: Conversion factors for Volume

To:	gal U.S.	Gal U.K.	Bbl	ft^3	l	m^3
From:	Multiply by:					
U.S. gallon (gal)	1	0.8327	0.02381	0.13379	3.785	0.0038
U.K. gallon (gal)	1.201	1	0.02859	0.1605	4.546	0.0045
Barrel (bbl)	42.0	34.97	1	5.615	159.0	0.159
Cubic foot (ft^3)	7.48	6.229	0.1781	1	28.3	0.0283
Litre (l)	0.2642	0.220	0.0063	0.0353	1	0.001
Cubic metre (m^3)	264.2	220.0	6.289	35.3147	1000.0	1

Source: IEA (1997), *Energy Balances of OECD Countries 1994-1995*, OECD, Paris, pp xv-xvi