GTAPinGAMS and GTAP-EG: Global Datasets for Economic Research and Illustrative Models

Thomas F. Rutherford and Sergey V. Paltsev *

Department of Economics, University of Colorado †

Working Paper No.

September 2000

Abstract

This paper documents a new version of the GTAPinGAMS package which includes two global economic datasets, GTAPinGAMS and GTAP-EG, corresponding illustrative models, and several ancillary programs for dataset management. The GTAPinGAMS is a dataset based on a general GTAPv4 dataset. The GTAP-EG is the energy-economy dataset adjusted for OECD International Energy Agency (IEA) statistics. The illustrative models of the GTAPinGAMS package have been implemented as a nonlinear complementarity problem in the GAMS programming language. Having the dataset in GAMS is helpful for researchers because of its open-architecture approach which permits to modify easily the dataset and the model for their own purposes. The document contains a description and directions for installing and using the GAMSinGAMS package.

---

* Development of these tools has been supported by the Electric Power Research Institute and the United States Department of Energy. The authors are indebted to Randy Wigle and Jared Carbone for their discussions and comments. The authors can be reached at: rutherford@colorado.edu, sergey.paltsev@colorado.edu

† University of Colorado, Department of Economics, Boulder, CO 80309-0256, USA.

1 Introduction

The Global Trade Analysis Project (GTAP) is a research program initiated in 1992 to provide the economic research community with a global economic dataset for use in the quantitative analyses of international economic issues. The Project's objectives include the provision of a documented, publicly available, global, general equilibrium data base, and to conduct seminars on a regular basis to inform the research community about how to use the data in applied economic analysis. GTAP has lead to the establishment of a global network of researchers who share a common interest of multi-region trade analysis and related issues. The GTAP research program is coordinated by Thomas Hertel, Director of the Center for Global Trade Analysis at Purdue University. As Deputy Director of this Center, Robert McDougall oversees the data base work. Software development within the GTAP project has been assisted greatly by the efforts of Ken Pearson and other Australian researchers from Centre of Policy Studies, Monash University. (See Hertel [1997], McDougall [1998]. A list of applications based on the GTAP framework can be found at the GTAP home page: `http://www.agecon.purdue.edu/gtap/`.)

The standard programming language for GTAP data and modeling work has been GEMPACK (Harrison and Pearson [1996]). In the GEMPACK framework the model is solved as a system of nonlinear equations. The present paper describes a version of the GTAP datasets and illustrative models which have been implemented as a nonlinear complementarity problem in the GAMS programming language. This version of the GTAPinGAMS package combines two separate packages GTAPinGAMS (Rutherford [1998]) and GTAP-EG (Rutherford and Paltsev [2000]). Combination of the packages is done due to a development of the new build routine. In addition, users of both packages can save some space and avoid potential confusion with different versions of the programs because both datasets use some similar files. The previous separate versions of these packages can be found at `http://debreu.colorado.edu/gtap/gtapgams.html`.

Along with the datasets and illustrative models, several ancillary programs for dataset management are developed. These programs should be useful to economists who program in GAMS and wish to use GTAP in applied work. These programs include tools for translation of the GTAP files into GAMS readable form, GAMS programs for dataset aggregation, filtering and the imposition of alternative tax rates on trade or domestic transactions.

The GTAP version 4 database[1] represents global production and trade for 45 country/regions, 50 commodities and 5 primary factors. The data characterize intermediate demand and bilateral trade in 1995, including tax rates on imports and exports. Development of the separate energy-economic dataset is based on the fact that the GTAP data alone are unsuitable for assessing issues such as energy use or climate change (Babiker and Rutherford [1997]). In particular, the GTAP data is at variance with the energy statistics of the International Energy Agency (IEA). In addition, the GTAP database is expressed in terms of *values*, i.e. price times quantity. The IEA has data on energy quantities, where the energy balances are expressed in a common unit, tonnes of oil equivalent. Information on energy prices and taxes at the level consistent with the GTAP data has been collected by Babiker and Malcolm [1998]. However, since only two out of three variables (price, quantity, value) can be regarded as independent, it is problematic to incorporate both price and quantity data into the GTAP database. Special procedures have been developed in order to incorporate the energy data into the GTAP 4 database. The resulting dataset (called GTAP-EG) is a balanced set of economic accounts (expressed in value terms) which is calibrated to energy quantity and price data.

This paper consists of three sections following this overview. Section 2 describes the GTAP datasets. This section provides information about the data organization and differences between the GTAPinGAMS and GTAP-EG datasets. Section 3 presents the illustrative static models. It starts with a description of the GTAPinGAMS model using Mathiesen's format for the Arrow-Debreu model. This section provides notation and equations describing technology, preferences and equilibrium conditions. It also describes how the GTAPinGAMS model can be expressed in

[1] A current version of GTAP database is GTAP 4. The fifth version is announced to be released in 2000.

GAMS, either as an MPSGE model or as a system of algebraic equations. This material provides a short but complete overview of how the technology and preferences are calibrated along with GAMS code which performs this task. The GTAP-EG illustrative model, which differs by its energy representation, is also described.

Section 4 has a practical perspective with step-by-step instructions on how to install the GTAPinGAMS package. The intent of this material is to provide as short as possible a learning curve for economists who wish to perform calculations using the GTAP datasets.

This section also describes ancillary GAMS programs which have been developed for use with the GTAP 4 dataset. These include GAMS libinclude programs which read and write GTAP header-array files[2]; FILTER.GMS, a GAMS program which removes small trade flows and intermediate demands from a GTAP dataset to increase sparsity and provide improved computational performance in large scale models; IMPOSE.GMS, a GAMS program which permits arbitrary adjustment of benchmark tax rates with least-squares recalibration; and GTAPAGGR.GMS, a GAMS program which aggregates any GTAP dataset to a smaller number of goods, factors or regions.

Distribution files for the GTAPinGAMS package are located as follows:

- A zip archive (http://debreu.colorado.edu/gtap/gtapgams.zip).
- A PDF version of this document (http://debreu.colorado.edu/papers/gtaptext.pdf).

An HTML document which describes the GTAPinGAMS package and provides an access to the distribution files is located at http://debreu.colorado.edu/gtap/gtapgams.html.

2 The GTAP Datasets

All GTAP datasets are defined in terms of three primary sets: r - the set of countries and regions, i - the set of sectors and produced commodities, and f - the set of primary factors. The GTAPinGAMS dataset, as the original GTAP4 dataset, has 45 regions, 50 goods[3], and 5 primary factors. The GTAP-EG dataset has 45 regions, 23 goods (5 of which are energy goods), and 5 primary factors. Lesser number of goods is determined by the structure of the energy statistics collected in OECD by Complainville [1998]. Identifiers for regions, sectors, and primary factors, as they are defined in GTAPinGAMS and GTAP-EG, are presented in Appendix 1.

An important feature of GTAPinGAMS package is that datasets may be freely aggregated into fewer regions, fewer sectors and even fewer primary factors. This feature permits a modeller to do preliminary model development using a small dataset to ensure rapid response and a short debug cycle. After having implemented a small model, it is then a simple matter to expand the number of sectors and/or regions in order to obtain a more precise empirical estimate.

2.1 The GTAPinGAMS Dataset

The GTAP data describe economic transactions in 1995. All parameters in GTAP are expressed in terms of values (i.e. price times quantity). Units of account in GTAP in its original GEMPACK representation are millions of 1995$. The units in GTAPinGAMS (and GTAP-EG) are different by a factor of 10,000. GTAPinGAMS measures transactions in tens of billions of 1995$. Scaling units in this way assures better numerical precision in equilibrium calculations.

GAMS statements which declare all parameters in a GTAP dataset is shown in Table 1. The parameters beginning with `v` are base year (1995) value data, most of which are from the original GEMPACK implementation of GTAP. Not all value data from the original dataset are included here. The principal difference is that GTAPinGAMS dataset stores tax *rates* rather than gross

[2] These tools have been implemented with the assistance of Ken Pearson using modified versions of his SEEHAR.EXE and MODHAR.EXE programs.

[3] GTAPinGAMS has 51 goods/production sectors: 50 goods + Investment composite (CGD)

and net of tax transaction values. The tax parameters, beginning with `t` are not in the original GEMPACK dataset.

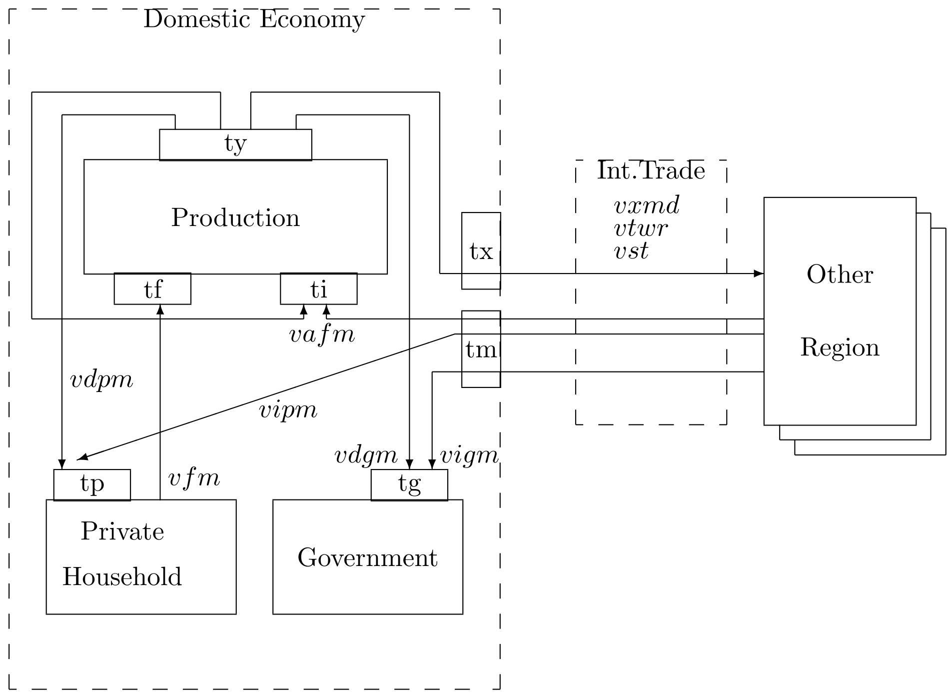

Fig. 1. GTAP flows explicitly represented in the dataset.

Figure 1 presents the general GTAP database flows, which are explicitly represented in the dataset. Whenever the GTAPinGAMS dataset is read, additional intermediate parameter values are assigned. Additional parameters are calculated based on the general flows. Declarations for the computed parameters are presented in Table 2.

Table 3 describes the GAMS parameter assignment statements for the computed items. Briefly, this is done as follows: (i) aggregate exports at market prices (`vxm`) are defined from the matrix of bilateral trade flows; (ii) aggregate imports at market prices (`vim`) are defined by bilateral exports, export taxes, transportation margins and tariff rates; (iii) domestic output (`vdm`) is determined as a residual through the zero profit condition; (vi) domestic supply to the intermediate demand (`vdfm`) is defined as a residual given domestic production and other demands for domestic output; (vii) import supply to intermediate demand (`vifm`) is also defined as a residual given aggregate imports, private and public import demand. This sequence of assignments implies that any imbalance in the dataset shows up as either a discrepancy in the demand and supply for intermediate inputs or as an imbalance between demand and supply of transportation services. The parameter `market` is created to generate a report of consistency of the benchmark data. Primary factor markets always balance because endowments are computed residually given benchmark factor demands across sectors. Likewise, regional current account balances are computed from the income-expenditure identity.

In the GTAP models we use Armington [1969] assumption that goods produced in different regions are qualitatively distinct. The GTAPinGAMS model uses the computed parameters `va`, `vm`, and `vd` which are defined over the market segment (intermediate, public, or private) represented by the set `d`.

Table 4 lists declarations and assignments of reference prices for each of the benchmark transactions which are subject to tax. These parameters are used in the GAMS model as part of the

Table 1: Parameters Explicitly Represented in a GTAPinGAMS Dataset

```
alias (i,j), (r,s);

PARAMETER
        ty(i,r)         Output tax
        ti(j,i,r)       Intermediate input tax
        tf(f,i,r)       Factor tax
        tx(i,s,r)       Export tax rate (defined on a net basis)
        tm(i,s,r)       Import tariff rate
        tg(i,r)         Tax rates on government demand
        tp(i,r)         Tax rate on private demand

        vafm(j,i,r)     Aggregate intermediate inputs
        vfm(f,i,r)      Value of factor inputs (net of tax)
        vxmd(i,r,s)     Value of commodity trade (fob - net export tax)
        vtwr(i,r,s)     Transport services
        vst(i,r)        Value of international transport sales
        vdgm(i,r)       Government demand (domestic)
        vigm(i,r)       Government demand (imported)
        vdpm(i,r)       Aggregate private demands (domestic)
        vipm(i,r)       Aggregate private demands (domestic);
```

Table 2: Computed Benchmark Parameters

```
parameter
        vim(i,r)        Total value of imports (gross tariff)
        vxm(i,r)        Value of export (gross excise tax)
        vdm(i,r)        Value of domestic output (net excise tax)
        vdfm(i,r)       Aggregate intermediate demand (domestic)
        vifm(i,r)       Aggregate intermediate demand (imported)
        vom(i,r)        Aggregate output value (gross of tax)
        vgm(i,r)        Public expenditures
        vpm(i,r)        Private expenditures
        vg(r)           Total value of public expenditure
        vp(r)           Total value of private expenditure
        vi(r)           Total value of investment
        vt              Value of international trade margins
        vb(*)           Net capital inflows
        market(*,*)     Consistency check for calibrated benchmark
        evoa(f,r)       Value of factor income
        va(d,i,r)       Armington supply
        vd(d,i,r)       Domestic supply
        vm(d,i,r)       Imported supply;
```

Table 3: Assignments for Computed Benchmark Parameters

```
vxm(i,r) = sum(s, vxmd(i,r,s)) + vst(i,r);

vim(i,r) = sum(s,(vxmd(i,s,r)*(1+tx(i,s,r))+vtwr(i,s,r))*(1+tm(i,s,r)));

vdm(i,r) = ( sum(j, vafm(j,i,r)*(1+ti(j,i,r)))
           + sum(f,  vfm(f,i,r)*(1+tf(f,i,r)))) / (1-ty(i,r)) - vxm(i,r);

vdfm(i,r) = vdm(i,r)  - vdgm(i,r) - vdpm(i,r) - vdm(i,r)$cgd(i);

vi(r) = sum(cgd, vdm(cgd,r));

vifm(i,r) = vim(i,r) - vipm(i,r) - vigm(i,r);

vom(i,r) = vdm(i,r) + vxm(i,r);

vgm(i,r) = vigm(i,r)+vdgm(i,r);

vpm(i,r) = vipm(i,r)+vdpm(i,r);

vg(r) = sum(i, vgm(i,r) * (1 + tg(i,r)));

vp(r) = sum(i, vpm(i,r) * (1 + tp(i,r)));

vt = sum((i,r), vst(i,r));

evoa(f,r) = sum(i, vfm(f,i,r));

vb(r) = vp(r) + vg(r) + vdm("cgd",r)
        - sum(f, evoa(f,r))
        - sum(i,     ty(i,r)   * vom(i,r))
        - sum((i,j), ti(j,i,r) * vafm(j,i,r))
        - sum((i,f), tf(f,i,r) * vfm(f,i,r))
        - sum((i,s), tx(i,r,s)  * vxmd(i,r,s))
        - sum((i,s), tm(i,s,r)  * (vxmd(i,s,r)*(1+tx(i,s,r)) + vtwr(i,s,r)) )
        - sum(i,     tg(i,r)*vgm(i,r))
        - sum(i,     tp(i,r)*vpm(i,r));

vm("c",i,r) = vipm(i,r);        vd("c",i,r) = vdpm(i,r);
vm("g",i,r) = vigm(i,r);        vd("g",i,r) = vdgm(i,r);
vm("i",i,r) = vifm(i,r);        vd("i",i,r) = vdfm(i,r);
va(d,i,r) = vm(d,i,r) + vd(d,i,r);
market(r,i) = vdfm(i,r) + vifm(i,r) - sum(j, vafm(i,j,r));
market("world","t") = vt - sum((i,r,s), vtwr(i,r,s));
```

Table 4: Benchmark Prices

```
parameter

        pc0(i,r)        Reference price index for private consumption
        pf0(f,i,r)      Reference price index for factor inputs
        pg0(i,r)        Reference price index for public
        pi0(j,i,r)      Reference price index for intermediate inputs
        pt0(i,s,r)      Reference price index for transport
        px0(i,s,r)      Reference price index for imports;

px0(i,s,r) = (1+tx(i,s,r))*(1+tm(i,s,r));
pt0(i,s,r) = 1+tm(i,s,r);
pc0(i,r)   = 1+tp(i,r);
pg0(i,r)   = 1+tg(i,r);
pi0(j,i,r) = 1+ti(j,i,r);
pf0(f,i,r) = 1+tf(f,i,r);
```

calibration of demand functions. For more discussion about the GAMS implementation of the static model, see Section 3.

2.2 The GTAP-EG Dataset

The GTAP-EG dataset is the GAMS version of the energy-economic dataset GTAP-E. It has been observed that the GTAP economic data provide a poor representation of energy flows (Babiker and Rutherford [1997]). The process of economic and energy data integration has proceeded in parallel at Purdue University and the University of Colorado at Boulder. Reconciling the data requires heroic adjustments. Two different approaches for calibration have been used. As a result, two energy datasets have been created.

An approach for calibration taken at Purdue University is to use the RAS procedure (United Nations [1973]) to fit energy quantities with "target" quantities, and then use FIT procedure to adjust the single region input-output coefficients. The process of incorporating energy data into GTAP is described in detail by Malcolm and Truong [1999]. We denote the Purdue dataset as GTAP-E-FIT. As a result of the FIT procedure, the information from all three data sources (GTAP economic data, IEA energy quantities, and price data) has been changed in the process of calibration.

In contrast to Purdue approach, we apply standard optimization techniques for calibrating the GTAP data to energy statistics. The resulting dataset which is described in this paper called GTAP-EG (GTAP-Energy in GAMS). Accordingly, the dataset and an illustrative model are presented in the GAMS programming language (Brook, Kendrick, Meeraus [1992]). The process of GTAP-EG creation by incorporating energy statistics into GTAP format is described in Rutherford and Paltsev [2000]. The GTAP-EG approach is to modify the GTAP value data as little as required while preserving the IEA energy quantity statistics and most of the prices.

The energy statistics collected in OECD by Complainville [1998] have 135 regions, 32 goods, and 7 energy commodities. The resulting GTAP-EG dataset has 45 regions, 23 goods (5 of which are energy goods), and 5 primary factors. An aggregation of 135 IEA-format regions into 45 GTAP regions is shown in Appendix 2. Most of the regional identifiers in the dataset correspond to standard UN three-character country codes[4].

[4]Users can define their own aggregations of the GTAP data and use any labels to describe regions. For technical

Table 5: Differences between GTAP-E-FIT and GTAP-EG sectoral identifiers.

Sector	GTAP-E-FIT	GTAP-EG
Electricity and heat	ELY	ELE
Refined oil products	P_C	OIL
Crude oil	OIL	CRU

Table 6: Differences between GTAP-E-FIT and GTAP-EG primary factor identifiers.

Sector	GTAP-E-FIT	GTAP-EG
Land	Land	LND
Skilled labor	SkLab	SKL
Unskilled labor	UnSkLab	LAB
Capital	Capital	CAP
Natural resources	NatRes	RES

To combine energy and trade data, 32 IEA-format sectors are aggregated into 22 sectors. In order to comply with IEA aggregation, the original 50 industrial sectors of GTAP data are also aggregated into the same 22 sectors. A sector for the investment composite is added to the original GTAP-GEMPACK representation. Table A.4 in Appendix 1 presents the identifiers for the 23 GTAP-EG sectors. The sectoral identifiers for energy are different from the GTAP-E-FIT identifiers[5]. The differences are noted in Table 5.

A concordance between IEA, GTAP 4, and GTAP-EG production sectors is presented in Appendix 3. The process of incorporating IEA statistics into GTAP-EG format is described in detail in Rutherford and Paltsev [2000]. Sectors may be aggregated to produce more compact datasets. The aggregation routine is described in Section 4.

Table 6 presents the three-character identifiers used for primary factors. Note that these differ from the primary factor names employed in the GEMPACK model.

The GTAP-EG dataset has a similar structure to the GTAPinGAMS dataset with the addition of energy quantities. The general database flows are shown in Figure 1. The parameters explicitly represented in the GTAP-EG are listed in Table 1 and Table 5. The energy parameters, beginning with ``e'' are in neither the original GTAP nor in the GTAPinGAMS dataset. Energy prices can be recovered by division of the respective values by the energy quantities. IEA statistics are expressed in a common unit, tonnes of oil equivalent. In the GTAP-EG units for electricity are converted into trillion kilowatt hour (TKWH) and units for other energy flows are converted into exajoules (EJ)[6].

A summary of economic activity by production sectors and regions in the GTAP-EG dataset is presented in Appendix 4. These numbers differ slightly from GTAP-E-FIT dataset[7]. The two energy datasets are different even though they are based on the same initial data, such as the GTAP version 4 (Hertel [1997]) expressed in terms of values (i.e. price times quantity), OECD International Energy Agency statistics (Complainville [1998]) expressed in terms of quantity, and energy price and tax data (Babiker and Malcolm [1998]). The reason for this discrepancy is the

reasons, if a GTAP dataset is to be used with MPSGE, then regional identifiers can have at most four characters.

[5]GTAP-E-FIT has the same identifiers as the GTAP4 dataset.

[6]Energy is defined as the capacity to do work. One joule (J) is a unit of energy equal to the work done when a force of 1 newton acts through a distance of 1 meter. One joule is approximately equivalent to the potential energy of one apple one meter above the floor. 1 exajoule (EJ) = $10^{18} J$. For conversion: 1 EJ = 23.88 million tonnes of oil equivalent (MTOE). For electricity: 1kwh = $3.61 \cdot 10^6 J$, or 1EJ = 0.2778 trillion kwh.

[7]A summary of economic activities from GTAP-E-FIT dataset can be found at http://debreu.colorado.edu/download/gtap-eg.html

Table 7: Energy Parameters Explicitly Represented in a GTAP-EG Dataset

```
PARAMETER
        eind(i,i,r)      Industrial energy demand,
        efd(i,r)         Final energy demand,
        eexp(i,r)        Energy exports,
        eimp(i,r)        Energy imports;
```

different calibration procedures that have been used. Since only two out of three variables (price, quantity, value) can be regarded as independent, it is problematic to incorporate both price and quantity data into the GTAP database.

To illustrate the difference between GTAP-EG and GTAP-E-FIT, we calculate carbon dioxide emissions and then compare the results with the IEA [1997] publication where the carbon dioxide emissions from fuel combustion are reported. It should be noted that the results from the IEA publication [1997] and the IEA statistics collected by Complainville are different. One source of the difference is International Marine Bunkers which are present in IEA book but not in the datasets. The International Marine Bunkers contains emissions from fuels burned by sea-going ships of all flags that are engaged in international transport. These emissions are excluded from national totals in IEA publication. As a result, the data for countries with big sea fleet differs substantially.

The CO_2 emissions for the full list of GTAP countries are presented in Appendix 4. Table 8 shows the results for the countries where differences in calculated CO_2 emissions are substantial. We report carbon dioxide emissions from the IEA publication. Then we compare them with the calculated emissions based on IEA statistics, GTAP-E-FIT, and GTAP-EG energy datasets. We have also provided the numbers for the GTAP-EG dataset without a fix for agriculture in USA (an ad hoc adjustment) described in Rutherford and Paltsev [2000]. It should be noted that there is a discrepancy between all four sources of the energy data. The calibration procedures employed in both the GTAP-E-FIT and the GTAP-EG do not reconcile precisely the IEA statistics. The carbon dioxide emissions are underestimated in the GTAP-E-FIT while they are overestimated slightly in the GTAP-EG.

3 Illustrative Models

In this section the core static models are presented. The purpose of the models is largely to illustrate how the benchmark data in the GTAPinGAMS and GTAP-EG datasets are organized. Any application of the GTAP data to a specific policy question should involve the development of a model tailored to the issues. A modeller can change the model to suit his specific purpose because he has a full access to the code of the build routine and the model.

3.1 The GTAPinGAMS Model

The core static model described in this paper does not have precisely the same structure as the GTAP model implemented in GEMPACK. There are several immediate differences between the standard GTAP model and the GTAPinGAMS model. First, the units of account are different by a factor of 10000. GTAP measures all transactions in millions of 1995$. GTAPinGAMS measure transactions in tens of billions of 1995$.

Second, there is a potentially important difference concerns the structure of final demand. In the GEMPACK model final demand is represented by a constant-difference-elasticity (CDE) demand system whereas in the GAMS model final demand is Cobb-Douglas. Given differences in functional forms, even if benchmark value shares and reference prices are identical the two models

Table 8: Carbon dioxide emissions (selected countries) - billion of tonnes

	IEA book	IEA stat	E-FIT	EG before fix	EG
JPN	1.151	1.208	1.145	1.257	1.257
KOR	0.353	0.449	0.396	0.449	0.449
SGP	0.059	0.085	0.085	0.085	0.085
CHN	3.007	3.098	2.902	3.112	3.112
IND	0.803	0.771	0.765	0.773	0.773
CAN	0.471	0.505	0.472	0.506	0.506
USA	5.228	5.339	5.175	5.340	5.460
MEX	0.328	0.328	0.309	0.328	0.328
BRA	0.287	0.269	0.256	0.289	0.289
GBR	0.565	0.605	0.540	0.607	0.607
DEU	0.884	0.973	0.865	0.973	0.973
REU	1.560	1.734	1.628	1.735	1.735
FSU	2.483	2.542	2.341	2.549	2.549
RME	0.817	0.788	0.755	0.827	0.827
ROW	0.518	0.208	0.183	0.208	0.208
total	22.150	22.482	21.272	22.644	22.764

may produce somewhat different estimates of policy changes due to differences in income and substitution elasticities.

A third set of differences concerns the representation of investment demand and global capital markets. The standard GTAP model assumes that a "global bank" allocates international capital flows in response to changes in regional rates of return. The GTAPinGAMS model makes the simplest possible assumptions regarding investment demand, international capital flows and the time path of adjustment: all of these variables are exogenously fixed at base year levels.[8]

In the next section, the GTAPinGAMS model in Mathiesen's format is presented. After describing the mathematics, the MPSGE formulation of the model is provided. This presentation includes a fair amount of explanatory text so that we hope it will be comprehensible to non-MPSGE programmers. Thereafter follows a description of the model as it may be expressed in GAMS algebra as an MCP model.[9]

3.1.1 GTAP in Mathiesen's Equilibrium Format

An Arrow-Debreu model concerns the interaction of consumers and producers in markets. Lars Mathiesen [1985] proposed a representation of this class of models in which two types of equations define an equilibrium: zero profit and market clearance. The corresponding variables defining an equilibrium are activity levels (for constant-returns-to-scale firms) and commodity prices.[10]

Commodity markets merge primary endowments of households with producer outputs. In equilibrium the aggregate supply of each good must be at least as great as total intermediate and final demand. Initial endowments are exogenous. Producer supplies and demands are defined by

[8] In extensions of the core static model, the GTAPinGAMS framework can be readily employed to study adjustment paths, but a description of these techniques lies beyond the scope of the present paper. See Rutherford, Lau and Pahlke [1998] for a pedagogic introduction to dynamic general equilibrium analysis within the GAMS framework.

[9] The distribution files provide representations of the core model as a constrained nonlinear system (CNS) and a square system of nonlinear constraints within a conventional nonlinear program (NLP).

[10] Under a maintained assumption of perfect competition, Mathiesen may characterize technology as CRTS without loss of generality. Decreasing returns are accommodated through introduction of a specific factor, while increasing returns are inconsistent with the assumption of perfect competition. In this environment zero excess profit is consistent with free entry for atomistic firms producing an identical product.

producer activity levels and relative prices. Final demands are determined by market prices.

Economists who have worked with conventional textbook equilibrium models can find Mathiesen's framework to be somewhat opaque because many quantity variables are not explicitly specified in the model. Variables such as final demand by consumers, factor demands by producers and commodity supplies by producers, are defined implicitly in Mathiesen's model. For example, given equilibrium prices for primary factors, consumer incomes can be computed, and given income and goods prices, consumers' demands can then be determined. The consumer demand functions are written down in order to define an equilibrium, but quantities demanded need not appear in the model as separate variables. The same is true of inputs or outputs from the production process: relative prices determine conditional demand, and conditional demand times the activity level represents market demand. Omitting decisions variables and suppressing definitional equations corresponding to intermediate and final demand provides significant computational advantages at the cost of a somewhat more complex model statement.

The core GTAPinGAMS model described here is a static, multi-regional model which tracks the production and distribution of goods in the global economy. In GTAP the world is divided into regions, and each region's final demand structure is portrayed by a representative agent who allocates expenditure across goods so as to maximize welfare, with fixed levels of investment and public output. Production incorporates intermediate inputs, and primary factors include skilled and unskilled labor, land, resources and physical capital. The dataset includes a full set of bilateral trade flows with associated transport costs, export taxes and tariffs.

In the following section, before writing down the equilibrium conditions per se, production technology and producer choices are described. Then the structure of private and public final demand are outlined. Finally, the zero profit and market clearance equations are written down.

It is a matter of personal taste in mathematics and computing, but we generally use one or two character identifiers in an algebraic exposition while employing GAMS parameters with as many as 10 characters. In order to avoid potential confusion due to differences in notation, Table 9 gives a cross-reference of symbols used in the algebraic formulation in this paper to the GAMS parameters which define the benchmark value of these variables in the GTAPinGAMS dataset.

3.1.2 Production

In the GTAP model there are two types of produced commodities, goods produced for domestic markets and goods produced for export. In the base GTAPinGAMS model these goods are assumed to be imperfect substitutes produced as joint products with a constant elasticity of transformation.[11] Specifically, if D_{ir} is domestic output and X_{ir} is exports, then

$$Y_{ir} = \left[\alpha_{ir}^{Y} D_{ir}^{1+1/\eta} + \beta_{ir}^{Y} X_{ir}^{1+1/\eta}\right]^{1/(1+1/\eta)}$$

where Y_{ir} is the activity level for good i in region r. Producers are competitive, implying that given a value of Y_{ir}, supplies to the domestic and export markets are given by:[12]

$$D_{ir} = Y_{ir}\ a_{ir}^{D}(p_{ir}^{D}, p_{ir}^{X})$$

and

$$X_{ir} = Y_{ir}\ a_{ir}^{X}(p_{ir}^{D}, p_{ir}^{X}).$$

[11] Model files in the GTAPinGAMS distribution accommodate an infinite elasticity of transformation between domestic and export markets as they are treated in the GTAP implementation in GEMPACK. For simplicity, my algebraic exposition in this paper focuses on the case in which the elasticity of transformation is finite.

[12] For the sake of brevity, I present functional forms explicitly but represent unit demand and supply functions in reduced form, e.g. $a_{ir}^{D}(p_{ir}^{D}, p_{ir}^{X})$. The next section of the paper presents detailed specific functions in the GAMS/MCP implementation.

Table 9: Algebraic Symbols and Related Benchmark Parameters

Symbol	GAMS Parameter
t_{ir}^{Y}	`ty(i,r)`
t_{ir}^{ID}	`ti(j,i,r)`
t_{ir}^{F}	`tf(f,i,r)`
t_{isr}^{X}	`tx(i,s,r)`
t_{isr}^{M}	`tm(i,s,r)`
t_{ir}^{G}	`tg(i,r)`
t_{ir}^{C}	`tp(i,r)`
$Y_{ir}a_{jir}$	`vafm(j,i,r)`
FD_{fir}	`vfm(f,i,r)`
M_{irs}	`vxmd(i,r,s)`
T_{irs}	`vtwr(i,r,s)`
TD_{ir}	`vst(i,r)`
DG_{ir}	`vdgm(i,r),`
MG_{ir}	`vigm(i,r)`
DC_{ir}	`vdpm(i,r)`
MC_{ir}	`vipm(i,r)`
X_{ir}	`vxm(i,r)`
D_{ir}	`vdm(i,r)`
DI_{ir}	`vdfm(i,r)`
MI_{ir}	`vifm(i,r)`
CD_{ir}	`vpm(i,r)`
GD_{ir}	`vgm(i,r)`
MI_{ir}	`vm("i",i,r)`
MG_{ir}	`vm("g",i,r)`
MC_{ir}	`vm("c",i,r)`
DI_{ir}	`vm("i",i,r)`
DG_{ir}	`vm("g",i,r)`
DC_{ir}	`vm("c",i,r)`
B_{r}	`B(r)`
τ_{irs}	`vtwr(i,r,s)/vxmd(i,r,s)`

Inputs to production include primary factors and intermediate inputs. Intermediate demands are proportional to the level of activity, so the total intermediate demand for good i in region r is:

$$ID_{ir} = \sum_j Y_{jr} a_{ijr}.$$

In the core model we assume that all intermediate input coefficients (a_{ijr}) are fixed, unresponsive to price.[13]

Following Armington [1969] intermediate demand is represented as a composite of imported and domestic goods as imperfect substitutes. Thus, we have:

$$ID_{ir} = \left[\alpha^I_{ir} DI^\rho_{ir} + \beta^I_{ir} MI^\rho_{ir}\right]^{1/\rho}$$

in which DI_{ir} is domestic intermediate and MI_{ir} is imported intermediate demand.

A Cobb-Douglas production function relates activity level and factor inputs. Producers minimize unit cost given factor prices and applicable taxes. The factor demands solve:

$$\min \sum_f p^F_{fr}(1 + t^F_{fir}) FD_{fir} \quad \text{s.t.} \quad \psi_{ir} \prod_f FD^{\theta_{fir}}_{fir} = Y_{ir}$$

taking Y_{ir} as given. Linear homogeneity of the production function implies that factor demands may be expressed as the product of an activity level and compensated demand function depending on factor prices and factor taxes:

$$FD_{fir} = Y_{ir}\ a^F_{fir}(p^F_r, t^F_{ir})$$

3.1.3 Public and Private Demand

Public sector output is assumed to represent a Cobb-Douglas aggregation of market commodities:

$$G_r = \Gamma_r \prod_i GD^{\theta^G_{ir}}_{ir}$$

As is the case for intermediate demand, an Armington aggregation of domestic and imported inputs defines public sector demand:

$$GD_{ir} = \left[\alpha^G_{ir} DG^\rho_{ir} + \beta^G_{ir} MG^\rho_{ir}\right]^{1/\rho}$$

Public sector output is exogenous, however the composition of public sector inputs responds to relative prices, gross of applicable tax, hence:

$$GD_{ir} = \overline{G}_r\ a^G_{ir}(p^D_{ir}, p^M_{ir}, t^G_{ir})$$

A representative agent determines final demand in each region. These consumers are endowed with primary factors, tax revenue, and an exogenously-specified net transfer from other regions. This income is allocated to investment, public demand and private demand. Investment and public output are exogenous while private demand is determined by utility maximizing behavior. The utility function is Cobb-Douglas:

$$U_r = \sum_i \theta^C_{ir} \log(CD_{ir})$$

[13] There is no reason that this functional form should be employed in every study. For example, when we use the GTAP dataset to study energy and environmental issues, it is important to account for the nature of substitution possibilities among energy carriers as well as between energy and non-energy inputs to production; so in those applications a nested CES function is employed in which energy trades off against value-added with a non-zero elasticity of substitution.

As in the case of intermediate and public demand, an Armington aggregation of domestic and imported inputs defines each commodity, so

$$CD_{ir} = \left[\alpha_{ir}^{C} DC_{ir}^{\rho} + \beta_{ir}^{C} MC_{ir}^{\rho}\right]^{1/\rho}$$

Aggregate final demand is then defined by regional expenditure and the unit price of aggregate of domestic and imported goods, gross of applicable tax:

$$CD_{ir} = \frac{\theta_{ir}^{C}\ M_r}{p_{ir}^{C}(1+t_{ir}^{C})}$$

Regional expenditure (M_r) includes factor income, net capital flows and tax revenue, net of the cost of investment and public expenditure.

3.1.4 Bilateral Trade

There are three types of imports in the model: imports to intermediate demand (MI_{ir}), imports to public sector demand (MG_{ir}) and imports to final consumer demand (MC_{ir}). The maintained assumption is that while the aggregate import share may differ between these three functions, each of these shares have the same regional composition within the import aggregate. A CES aggregation across imports from different regions s forms the total import composite:

$$MI_{ir} + MG_{ir} + MC_{ir} = \left[\sum_{s} \alpha_{isr}^{M} M_{isr}^{\rho}\right]^{1/\rho}$$

Two tax margins and a transportation cost apply on bilateral trade in the model. Real transport costs are proportional to trade:

$$T_{irs} = \tau_{irs} M_{irs}$$

and these inputs are defined by a Cobb-Douglas aggregate of international transport inputs supplied by different countries:

$$\sum_{irs} T_{irs} = \psi_T \prod_{i,r} TD_{ir}^{\theta_{ir}^{T}}$$

It is helpful to think of international transportation margins as transportation services which are provided by perfectly competitive producers from different regions with an Armginton aggregation across services from different countries and an elasticity of substitution equal to unity. The technology providing transportation services exhibits constant returns to scale, so we can specify a price p^T representing the unit cost of transportation on all commodity trade flows.[14]

Bilateral trade flows are determined by cost-minimizing choice, given the fob export price from region r, p_{ir}^{X}, the export tax rate, t_{ir}^{X}, and the import tariff rate, t_{ir}^{M}.[15] We then may write the demand for bilateral imports as:

$$M_{irs} = M_{is}\ a_{irs}^{M}(p_{ir'}^{X}, t_{ir's}^{X}, p^{T}, t_{ir's}^{M})$$

[14]There are some simplifications here. For example, the regional composition of transportation services is identical across all bilateral trade flows. Furthermore, while the dataset incorporates explicit trade and transport margins on international trade flows, wholesale and retail margins on domestic sales are ignored in the dataset, so there is some asymmetry in the database's price level.

[15]The model formulation assumes that the export tax applies on the fob price (net of transport margins), while the import tariff applies on the cif price, gross of export tax and transport margin.

3.1.5 Income and Expenditure

Consumer expenditure for a representative agent are the sum of factor earnings and tax revenue, net the cost of investment, public sector output and net capital outflows:

$$
\begin{array}{lll}
M_r = & \sum_f p^F_{fr}\ F_{fr} & \text{! factor income} \\
& +\sum_i t^Y_{ir}\ (p^D_{ir} D_{ir} + p^X_{ir} X_{ir}) & \text{! indirect taxes} \\
& +\sum_{ij} t^{ID}_{ijr}\ p^{ID}_{ir}\ Y_{jr}\ a_{ijr} & \text{! taxes on intermediate goods} \\
& +\sum_{fi} t^F_{fir}\ p^F_{fr}\ FD_{fir} & \text{! factor tax revenue} \\
& +\sum_i t^G_{ir}\ p^{GD}_{ir}\ GD_{ir} & \text{! public tax revenue} \\
& +\sum_i t^C_{ir}\ p^{CD}_{ir}\ CD_{ir} & \text{! consumption tax revenue} \\
& +\sum_{is} t^X_{irs}\ p^X_{ir}\ M_{irs} & \text{! export tax revenue} \\
& +\sum_{is} t^M_{isr}\ (p^X_{is}\ M_{isr}\ (1+t^X_{isr}) + p^T T_{isr}) & \text{! tariff revenue} \\
& -\sum_i p^D_{ir} I_{ir} & \text{! investment demand} \\
& -\sum_i p^G_{ir}\ (1+t^G_{ir})\ GD_{ir} & \text{! public sector demand} \\
& -p^C_n B_r & \text{! current account balance}
\end{array}
$$

Capital flows in the base year are represented by B_r in this expression, and in a counterfactual equilibrium these are held fixed and denominated in terms of the numeraire price index, the consumer price level in region n (USA).

3.1.6 Market Clearance

Having defined technology, preferences and compensated demand functions, we now may turn to the market clearance conditions.

- Domestic Output

  Domestic output equals demand for intermediate inputs to production, public sector use, final consumer demand plus domestic investment:[16]

$$
\begin{array}{lllll}
D_{ir} & = DI_{ir} & +DG_{ir} & +DC_{ir} & +I_{ir} \\
& = ID_{ir}\ a^{D,I}_{ir} & +GD_{ir}\ a^{D,G}_{ir} & +CD_{ir}\ a^{D,C}_{ir} & +I_{ir}
\end{array}
$$

  in which $a^{D,I}_{ir}$, $a^{D,G}_{ir}$, and $a^{D,C}_{ir}$ represent the compensated demands for domestic inputs by submarket, each of which are functions of p^D_{ir} and p^M_{ir}.

- Imports

  Aggregate supply of imports, defined by the Armington aggregation across imports from different regions must equal aggregate import demand for intermediate, public and private consumption:

$$
\begin{array}{llll}
M_{ir} & = MI_{ir} & +MG_{ir} & +MC_{ir} \\
& = ID_{ir}\ a^{M,I}_{ir} & +GD_{ir}\ a^{M,G}_{ir} & +CD_{ir}\ a^{M,C}_{ir}
\end{array}
$$

  in which $a^{M,I}_{ir}$, $a^{M,G}_{ir}$, and $a^{M,C}_{ir}$ represent compensated demands for imported inputs by submarket, each functions of p^D_{ir} and p^M_{ir}.

[16] Within the dataset investment inputs flow to the `cgd` sector, and demand for `cgd` sectoral output appears as the sole non-zero in the I_{ir} vector for each region r.

- Exports

  Export supplies equals import demand across all trading partners plus demands for international transport:[17]

$$\begin{aligned} X_{ir} &= \sum_s M_{irs} & +TD_{ir} \\ &= \sum_s M_{is} a^M_{irs} & +T\ a^T_{ir} \end{aligned}$$

  In the second equation a^M_{irs} represents the unit demand for region r output per unit of region s aggregate imports.

- Armington Aggregate Supply

  The model includes supply-demand conditions for the Armington composite goods entering intermediate demand, public and private demand, as has already been specified above in the equations defining ID_{ir}, GD_{ir} and CD_{ir}.

- Primary Factors

  Primary factor (labor, capital, land, resource) endowment equals primary factor demand:

$$F_{fr} = \sum_i Y_{ir}\ a^F_{fir}$$

3.1.7 Zero Profit

- Production

  Competitive producers operating constant-returns technology earn zero profit in equilibrium. For the GTAP producer, the value of output to the firm equals the value of sales in the domestic and export markets net of applicable indirect taxes. Costs of production include factor inputs (taxed at rate t^F) and intermediate inputs (taxed at rate t^{ID}):

$$(p^D_{ir}\ a^D_{ir} + p^X_{ir}\ a^X_{ir})(1 - t^Y_{ir}) = \sum_f a^F_{fir} p^F_{fr}(1 + t^F_{fir}) + \sum_j a_{jir} p^{ID}_{jr}(1 + t^{ID}_{jir})$$

- Imports

  Zero profit conditions apply to trade activities as well as production. In equilibrium, the value of imports at the domestic cif price therefore equals the fob price gross of export tax, the transportation margin and the applicable tariff:

$$p^M_{ir} = \sum_s a^M_{irs} \left[p^X_{is}(1 + t^X_{isr}) + \tau_{irs} p^T \right] (1 + t^M_{isr})$$

- Intermediate, Public and Private Demand

  Armington aggregation functions transform domestic and imported goods into composite goods for intermediate demand, public sector demand and private demand. Zero profit for these activities provide the following equilibrium identities:

[17]When the elasticity of transformation between goods produced for the domestic and export markets is infinite, the market clearance conditions for D_{ir} and X_{ir} are merged, i.e.

$$Y_{ir} = DI_{ir} + DG_{ir} + DC_{ir} + I_{ir} + \sum_s M_{irs} + TD_{ir}.$$

and prices p^D_{ir} and p^X_{ir} are replaced throughout the model by a single price index, p^Y_{ir}.

$$p^I_{ir} = c(p^D_{ir}, p^M_{ir}, \alpha^I_{ir}, \beta^I_{ir})$$

$$p^G_{ir} = c(p^D_{ir}, p^M_{ir}, \alpha^G_{ir}, \beta^G_{ir})$$

$$p^C_{ir} = c(p^D_{ir}, p^M_{ir}, \alpha^C_{ir}, \beta^C_{ir})$$

in which

$$\begin{aligned} c(p^D, p^M, \alpha, \beta) &\equiv \min_{D,M} \quad p^D D + p^M M \quad \text{s.t.} \quad \left(\alpha D^\rho + \beta M^\rho\right)^{1/\rho} = 1 \\ &= \left(\alpha^\sigma p_D^{1-\sigma} + \beta^\sigma p_M^{1-\sigma}\right)^{1/1-\sigma} \end{aligned}$$

is the unit cost function defined by the constant-elasticity-of-substitution aggregate of domestic and imported inputs.

3.1.8 The MPSGE Formulation

Table 10 contains variable declarations for the GTAPinGAMS model as implemented in MPSGE.[18] The model includes sectors related to production by commodity and region (`Y(i,r)`); Armington aggregation across imports from different trading partners (`M(i,r)`); Armington aggregation between domestic and imported varieties by market segment (`A(d,i,r)` in which `d` refers to intermediate, public and private demand); public demand by region (`G(r)`); private demand by region (`C(r)`); and the provision of international transport margins (`YT`).

The production activities for public and private demand are associated with outputs which represent the marginal cost of public and private consumption, `PG(r)` and `PC(r)`. For each commodity and region there are six different price indices. `PD(i,r)` represents the cost index for a unit of domestic output; `PX(i,r)` represents the cost index for exports; `pm(i,r)` represents the cost of a unit of imports (aggregated across all trading partners), and `PA(d,i,r)` represents the cost index of a unit of composite Armington supply by submarket. Primary factor prices are represented by `PF(f,r)`, and the market price of a unit of international transport services is represented by `PT`.

The final class of variables in the MPSGE model are consumers, and in this model there is one representative consumer for each region. In a solution `RA.L(r)` returns the equilibrium expenditure on household consumption by region `r`.

An MPSGE model is specified by a sequence of function "blocks", one for each production sector and consumer in the model. In this model there are six *classes* of production sectors. (See Table 10.) The first of these refers to production by commodity and region, `Y(i,r)`. This production activity has a nested-CES cost function with a Leontief aggregation across intermediate inputs at the top level (see `s:0`) and unity within the value-added aggregate (`va:1`), and an elasticity of transformation across outputs equal to 2 (see `T:2`). There are inputs and outputs associated with the `Y(i,r)` production function. Outputs correspond to production for the domestic market, `O:PD(i,r)`, and production for the export market, `O:PX(i,r)`. The reference quantity entries for these coefficients are the benchmark values of domestic and export sales. A tax at an ad-valorem rate `ty(i,r)` is applied to both domestic and export sales.[19]

[18] I have omitted exception operators from the variable and function declarations to make the code easier to read. In most aggregations of the dataset, the model shown here is operational. In highly disaggregate models, however, not all goods are produced in all regions, and it is necessary to specify, for example, `Y(i,r)$(vdm(i,r)+vxm(i,r))`.

[19] The output tax is defined on a *gross basis*. For example, the value of sales in the domestic market gross of tax equals `vdm(i,r)` of which `(1-ty(i,r))*vdm(i,r)` is returned to producers and `ty(i,r)*vdm(i,r)` is paid to the government.

Table 10: Variable Declarations for GTAP Implemented in MPSGE

```
$MODEL:GTAP

$SECTORS:
        Y(i,r)          ! Output
        A(d,i,r)        ! Armington aggregation of domestic and imports
        M(i,r)          ! Import aggregation
        C(r)            ! Private consumption
        G(r)            ! Public provision
        YT              ! Transport

$COMMODITIES:
        PG(r)           ! Public provision
        PC(r)           ! Private demand
        PD(i,r)         ! Output price
        PX(i,r)         ! Export price
        PM(i,r)         ! Import price
        PA(d,i,r)       ! Armington composite price
        PF(f,r)         ! Factor price
        PT              ! Transport services

$CONSUMERS:
        RA(r)           ! Representative agent
```

There are two types of inputs to the `Y(i,r)` production function, corresponding to goods and factors. Intermediate inputs are taken from the market for Armington aggregates into production. Substitution between factor inputs is created by assigning those inputs to the `va:` input aggregate.[20]

Taxes are levied on intermediate demand inputs at net rate `ti` and taxes apply to primary factor inputs at net rate `tf`. For example, the market value of primary factor services purchased by firms is `vfm(f,i,r)`, but the total cost to firms equals `vfm(f,i,r)*(1+tf(f,i,r))`, of which `vfm(f,i,r)` is paid as wages or dividends to factor owners while `vfm(f,i,r)*tf(f,i,r)` is paid as a tax to `RA(r)`.[21]

The Armington aggregation activity `A(d,i,r)` generates three functions for each commodity in each region. For simplicity, we have specified a domestic-import elasticity of substitution equal to 4 for all goods, commodities and Armington submarkets.

The import aggregation activity, `M(i,r)`, is the most complex component of the model. First, it defines the aggregation of imports by trading partner. Second, it applies export taxes and import tariffs on all bilateral trades.[22] Third, it applies transportation margins which are proportional to

[20] "`va:`" is a nesting identifier. These names are arbitrary and may have from one to four characters. Two reserved names are "`s:`" which represents the elasticity of substitution at the root of the inputs tree and "`T:`" which represents the elasticity of transformation at the root of the output tree.

[21] Readers unfamiliar with the MPSGE model representation may wish to refer back to the algebraic equilibrium conditions. The specification of the `$PROD:Y(I,R)` block automatically generates a zero profit condition for Y_{ir}. It also generates terms in the market clearance equations for all associated inputs and outputs. In this function the affected markets include the domestic output market, the market for export of good i from region r, markets for Armington composites entering intermediate demand, and primary factors markets. For this reason the tabular format is very compact – in essence, the user only needs to specify the dual (zero-profit) conditions and the modeling language automatically generates the primal (market clearance) equations.

[22] Note that export taxes on sales from region `s` in region `r` are accrue to the representative agent in region `s`

Table 11: Function Declarations for GTAP Implemented in MPSGE

```
$PROD:Y(i,r)  S:0  T:eta  va:1
        O:PD(i,r)       Q:vdm(i,r)    A:RA(r) T:ty(i,r)
        O:PX(i,r)       Q:vxm(i,r)    A:RA(r) T:ty(i,r)
        I:PA("i",j,r)   Q:vafm(J,i,r) P:pi0(j,i,r) A:RA(r) T:ti(j,i,r)
        I:PF(f,r)       Q:vfm(f,i,r)  P:pf0(f,i,r) A:RA(r) T:tf(f,i,r) va:

$REPORT:
        V:FD(f,i,r)     I:PF(f,r)       PROD:Y(i,r)
        V:YD(i,r)       O:PD(i,r)       PROD:Y(i,r)
        V:YX(i,r)       O:PX(i,r)       PROD:Y(i,r)

$PROD:A(d,i,r)          S:esubdm
        O:PA(d,i,r)     Q:va(d,i,r)
        I:PD(i,r)       Q:vd(d,i,r)
        I:PM(i,r)       Q:vm(d,i,r)

$PROD:M(i,r)            S:esubmm   s.TL:0
        O:PM(i,r)       Q:vim(i,r)
        I:PX(i,s)       Q:vxmd(i,s,r)  P:px0(i,s,r)     s.TL:
+                       A:RA(S) T:TX(i,s,r) A:RA(r) T:(tm(i,s,r)*(1+tx(i,s,r)))
        I:PT#(s)        Q:vtwr(i,s,r)  P:pt0(i,s,r)     s.TL:
+                       A:RA(r) T:tm(i,s,r)

$PROD:G(r)  S:1
        O:PG(r)         Q:vg(r)
        I:PA("g",i,r)   Q:vgm(i,r)  P:pg0(i,r)  A:RA(r) T:tg(i,r)

$PROD:C(r)  S:1
        O:PC(r)         Q:vp(r)
        I:PA("c",i,r)   Q:vpm(i,r)  P:pc0(i,r)  A:RA(r) T:tp(i,r)

$PROD:YT  S:1
        O:PT            Q:vt
        I:PX(i,r)       Q:vst(i,r)

$DEMAND:RA(r)
        E:PF(f,r)       Q:evoa(f,r)
        E:PC(num)       Q:vb(r)
        E:PD(cgd,r)     Q:(-vi(r))
        E:PG(r)         Q:(-vg(r))
        D:PC(r)         Q:vp(r)
```

quantities traded. The output market PM(i,r) serves as an input to the Armington aggregation sectors. There are two types of inputs to the M(i,r) activity. The I:PX(i,s) input represents *fob* payments to producers in region s.

The I:PT#(s) input represents multiple inputs of transportation services, one for each element of set s. There are multiple inputs of transportation services into each imported good simply because every bilateral trade flow demands its own transportation services. Using a Leontief aggregate on each bilateral trade flow assures that transport costs and imports remain strictly proportional to the base year level, $\tau(i,r,s) = \text{vtwr(i,s,r)}/\text{vxmd(i,s,r)}$.

The function declaration indicates a top-level substitution elasticity equal to esubmm (S:esubmm), and it also indicates a *vector* of second level input nests, each with an elasticity of substitution equal to zero (s.tl:0).[23]

Final consumption by consumers and producers in region r is characterized by production activities c(r) and g(r), respectively. The elasticity of substitution across goods in final demand is specified to be unity (S:1).

The yt production activity provides international transportation services as a Cobb-Douglas composite of goods provided in the domestic markets of each region.

The model statement concludes with a specification of endowments and preferences for each region's representative agent ($DEMAND:RA(r)). Each agent is endowed with primary factors and capital inflows. They are also "endowed" with a fixed negative quantities of the domestic cgd commodity and public sector outputs representing exogenously-specified demands for investment and public sector output. All remaining income is allocated to private consumption.

The closure adopted here in which investment and public demand are both exogenous is adopted for simplicity, and also because the welfare estimates from this closure seem to most closely match the infinite-horizon model (see Rutherford and Tarr [1998]). In the MPSGE model it is quite simple to adopt alternative assumptions regarding investment. For example, investment could be modeled by a constant marginal propensity to save as:

```
$DEMAND:RA(r)  s:1
        E:PF(f,r)        Q:evoa(f,r)
        E:PC(num)        Q:vb(r)
        E:PG(r)          Q:(-vg(r))
        D:PD(cgd,r)      Q:vi(r)
        D:PC(r)          Q:vp(r)
```

As stated above, the MPSGE formulation of an equilibrium model follows Mathiesen's modeling format in which intermediate demand and supply functions can be captured as functions of prices and activity levels. The computational advantage of this approach is that fewer variables are needed, and it is considerably less costly to solve the resulting smaller system of equations.[24]

Intermediate demands and supplies from MPSGE models can be computed by the modeller using equilibrium prices and activity levels. For example, in this model it is quite simple to compute private, public and transportation demands from the solution of an MPSGE model, because all of these activities are Cobb-Douglas. (See Table 12.) It is, however, possible to extract equilibrium demands directly from the MPSGE function evaluation through use of the $REPORT: statement, listed in Table 11 immediately after the Y_{ir} production block. In this model three demand and supply quantity variables are declared, representing primary factor demand by sector, supply to

(A:RA(s)) while import tariffs are paid to the representative agent in region r (A:RA(r)).

[23] A .tl suffix alerts MPSGE that a set of nests are being declared. When an input is to be associated with one of these nests, the set label flag must be specified on the input line.

[24] In terms of computational complexity, the cost of solving a system of equations increases somewhere between the square and the cube of the number of dimensions, although in large-scale implementations such as the GAMS/MCP solver PATH or MILES, computational complexity depends on both the number of equations and their density.

Table 12: Computing Demand Quantities from an MPSGE Equilibrium

```
parameter       cd(i,r) Private demand
                gd(i,r) Public demand
                td(i,r) Transportation demand;

cd(i,r) = vpm(i,r) * C.L(r) * PC.L(r) * pc0(i,r)
                        / ( PA.L("c",i,r) * (1 + tp(i,r)) );
gd(i,r) = vgm(i,r) * G.L(r) * PG.L(r) * pg0(i,r)
                        / ( PA.L("g",i,r) * (1 + tg(i,r)) );
td(i,r) = vst(i,r) * YT.L * PT.L / PX.L(i,r);
```

the domestic market and supply to the export market. These values are returned in `FD.L(f,i,r)`, `YD.L(i,r)`, and `YX.L(i,r)` respectively.

3.1.9 The Algebraic Formulation

Compactness of representation leaves fewer opportunities to make mistakes. For this reason alone we prefer to implement equilibrium models with MPSGE. We recognize, however, that the tabular MPSGE syntax can be impenetrable for many competent modellers, therefore, we conclude my discussion of the core GTAPinGAMS model by going through the implementation in GAMS algebra. We pose the model here as a mixed complementarity problem, but in this formulation all of the market clearance and zero profit conditions will hold with equality, so the model can be solved in GAMS as a nonlinear program (NLP) with a vacuous objective or a constrained nonlinear system (CNS).

In order to write out GTAP in algebraic form, it is helpful to introduce some additional benchmark data structures which simplify demand function algebra. The extra parameters include benchmark value shares for all of the nonlinear demand and supply functions in the model. (See Table 13.)

The variables used in the GAMS/MCP model are listed in Table 14, separated by blank lines into four subsets. The first set of variables are unit demand and supply functions, corresponding to the symbols from the algebraic formulation above. (See Table 9.) Demand and supply functions are represented implicitly in the MPSGE model, but for simplicity these are introduced as separate symbols in the algebraic model. The remaining three subsets of variables in the model correspond precisely to the sectors, commodities and consumers in the MPSGE model.

Table 15 presents equations defining the unit demand and supply functions. Using benchmark quantities, prices and value shares, we use the calibrated share form to express demands as a function of input prices. In order to keep track of what is what, we are following Michael Saunders' suggestions for GAMS program style, listing variables and GAMS variables in upper case, parameters and sets in lower case. The three symbols in Table 15 have not yet been defined. `eta=2` represents the elasticity of transformation between production for the domestic and export markets, `esubdm=4` is the domestic-import Armington elasticity of substitution, and `esubmm=8` is the import-import Armington elasticity.

Having defined compensated demands, it is then straightforward to write down zero profit equations. For sector Y_{ir} this means that the cost of inputs to production (intermediate demand plus primary factors, gross of tax) must equal the value of outputs (domestic sales plus exports, net of tax).

The zero profit conditions for `A(d,i,r)`, and `YT` are based on CES and Cobb-Douglas cost func-

Table 13: Benchmark Share Parameters used in the Algebraic Model

```
PARAMETER
        vad(i,r)        Sectoral value-added,
        tau(i,r,s)      Unit transport cost coefficient

        thetaf(f,i,r)   Value added,
        thetad(i,r)     Domestic output,
        thetag(i,r)     Government demand,
        thetap(i,r)     Private demand,
        thetat(i,r)     Transport,
        thetam(d,i,r)   Import value share,
        beta(i,s,r)     Value share of bilateral imports,
        gamma(i,s,r)    Goods share of unit import cost;

vad(i,r) = sum(f, vfm(f,i,r)*pf0(f,i,r));
tau(i,r,s)$vxmd(i,r,s) = vtwr(i,r,s) / vxmd(i,r,s);

thetam(d,i,r)$va(d,i,r) = vm(d,i,r) / va(d,i,r);
thetaf(f,i,r)$vad(i,r) = pf0(f,i,r) * vfm(f,i,r) / vad(i,r);
thetad(i,r)$vom(i,r) = vdm(i,r) / vom(i,r);
thetag(i,r) = pg0(i,r) * vgm(i,r) / vg(r);
thetap(i,r) = pc0(i,r) * vpm(i,r) / vp(r);
thetat(i,r)$sum((j,s), vst(j,s)) = vst(i,r) / vt;
beta(i,s,r)$vxmd(i,s,r) =
           (vxmd(i,s,r)*pmx0(i,s,r)+vtwr(i,s,r)*pmt0(i,s,r) )  / vim(i,r);
gamma(i,s,r)$vxmd(i,s,r) = vxmd(i,s,r)*pmx0(i,s,r) /
           (vxmd(i,s,r)*pmx0(i,s,r) + vtwr(i,s,r)*pmt0(i,s,r) );
```

Table 14: Variables in the MCP Model

```
VARIABLES

        A_G(i,r)        Public sector unit demand
        A_C(i,r)        Private unit demand
        A_F(f,i,r)      Factor unit demand
        A_X(i,r)        Export unit supply
        A_D(i,r)        Domestic unit supply
        A_M(i,r,s)      Import unit demand

        C(r)            Private consumption
        G(r)            Public provision
        Y(i,r)          Aggregate Output
        M(i,r)          Import aggregation
        A(d,i,r)        Armington aggregations
        YT              Transport

        PC(r)           Private demand
        PG(r)           Public provision
        PD(i,r)         Domestic Output price
        PX(i,r)         Export price
        PM(i,r)         Import price
        PA(d,i,r)       Armington composite price
        PF(f,r)         Factor price
        PT              Transport services

        RA(r)           Representative agent income;
```

Table 15: Compensated Unit Demand and Supply Functions

```
DEF_G(i,r)..

  A_G(i,r) =E= vgm(i,r) *
        PROD(j, (PA("g",j,r)*(1+tg(j,r))/pg0(j,r))**thetag(j,r)) /
        ( PA("g",i,r)*(1+tg(i,r))/pg0(i,r) );

DEF_C(i,r)..

  A_C(i,r) =E= vpm(i,r) *
        PROD(j, (PA("c",j,r)*(1+tp(j,r))/pc0(j,r))**thetap(j,r)) /
        ( PA("c",i,r)*(1+tp(i,r))/pc0(i,r) );

DEF_F(f,i,r)..

  A_F(f,i,r) =E= vfm(f,i,r) *
        PROD(ff, (PF(ff,r)*(1+tf(ff,i,r))/pf0(ff,i,r) )**thetaf(ff,i,r)) /
        ( PF(f,r)*(1+tf(f,i,r)) / pf0(f,i,r) );

DEF_X(i,r)..

  A_X(i,r) =E=  vxm(i,r) * (PX(i,r) /
        (thetad(i,r)  * PD(i,r)**(1+eta) +
        (1-thetad(i,r)) * PX(i,r)**(1+eta))**(1/(1+eta)) )**eta;

DEF_D(i,r)..

  A_D(i,r) =E=  vdm(i,r) * (PD(i,r) /
        (thetad(i,r)  * PD(i,r)**(1+eta) +
        (1-thetad(i,r)) * PX(i,r)**(1+eta))**(1/(1+eta)) )**eta;

DEF_M(i,r,s)..

  A_M(i,r,s) =E= vxmd(i,r,s) * ( PM(i,s) /
        ( gamma(i,r,s)  * PX(i,r)*(1+tx(i,r,s))*(1+tm(i,r,s))/pmx0(i,r,s)
        + (1-gamma(i,r,s)) * PT*(1+tm(i,r,s)) / pmt0(i,r,s) ) )**esubmm;
```

Table 16: Exhaustion of Production Equations in the MCP Formulation

```
*       Production:

PRF_Y(i,r)..

  SUM(j, vafm(j,i,r) * PA("i",j,r) * (1+ti(j,i,r)) ) +
        SUM(f, A_F(f,i,r) * PF(f,r) * (1 + tf(f,i,r)) )
        =E= (1 - ty(i,r)) * (PD(i,r) * A_D(i,r) + PX(i,r) * A_X(i,r));


*       Armington aggregation across imports from different countries:

PRF_M(i,r)..

  SUM(s, (1 + tm(i,s,r)) * A_M(i,s,r) *
        (PX(i,s) * (1 + tx(i,s,r)) + PT * tau(i,s,r))) =E= PM(i,r) * vim(i,r) ;

*       Public output:

PRF_G(r)..      SUM(i, PA("g",i,r) * (1+tg(i,r)) * A_G(i,r)) =E= PG(r) * vg(r);

*       Private consumption:

PRF_C(r)..      SUM(i, PA("c",i,r) * (1+tp(i,r)) * A_C(i,r)) =E= PC(r) * vp(r);


*       Import-domestic aggregation by submarket:

PRF_A(d,i,r)..

  ((1-thetam(d,i,r)) * PD(i,r)**(1-esubdm) +
        thetam(d,i,r)  * PM(i,r)**(1-esubdm) )**(1/(1-esubdm)) =E= PA(d,i,r);

*       Inter-national transport services (Cobb-Douglas):

PRF_YT..        PROD((i,r), PD(i,r)**thetat(i,r)) =E= PT;
```

tions because the associated unit demand functions are not defined explicitly in the model.[25] Market clearance equations are displayed in Table 17. The only tricky equation here is `MKT_PA(d,i,r)` in which we use three different subsets to equate submarket supply with intermediate, public and private demand. The subsets are declared:

```
set i_d(d)/i/, c_d(d) /c/, g_d(d) /g/;
```

This notation permits us to define a different right-hand-side expression for each element of set d in the model definition.

The final equation for this model is an expression defining regional income as a function of factor prices, transfers, and tax revenue. The complexity in this equation concerns accounting for revenue from each of seven different tax instruments.[26] Not to belabor the point, but the income expression in Table 18 illustrates the usefulness of MPSGE for tax policy analysis.[27]

[25]Of course it is mathematically equivalent to use the cost function or an expression for cost based on the unit demand functions, i.e. if:

$$c(p) \equiv min_x \sum p_i x_i \quad \text{s.t.} f(x) = 1$$

then $c(p) = \sum_i p_i x_i^*(p)$ where $x_i^*(p)$ is the unit demand function.

[26]There is a subtle but important point with regard to the complex system of taxes in GTAP. Users should not assume that because the dataset has a tax *instrument* the associated tax *rates* have a strong empirical basis. The research work in putting together GTAP has tended to focus on trade taxes (import tariffs and export taxes), and all other tax rates come directly from the national input-output tables. If you undertake an analysis in which the structure of the domestic tax system plays an important role, it is highly recommended to collect and update the benchmark tax rates. For an example of how a domestic tax system may be introduced in a GTAP model, see Harrison, Rutherford and Tarr [1997].

[27]In the MPSGE model a single entry in the import activity introduces both the import and export taxes, and given a description of taxes applying to the producer, the modelling language automatically generates the appropriate income entries, greatly reducing the likelihood of an accounting error.

Table 17: Market Clearing Equations in the MCP Formulation

```
*       Exports:

MKT_PX(i,r)..

  YX(i,r)*Y(I,R) =E= SUM(s, A_M(i,r,s)*M(i,s)) + VST(i,r)*YT*(PT/PX(i,r));

*       Domestic supply:

MKT_PD(i,r)..

  YD(i,r) * Y(I,R) =E=
        SUM(d, A(d,i,r) * vd(d,i,r) * ( PA(d,i,r)/PD(i,r) )**esubdm )
        + vi(r)$cgd(i);

*       Imports:

MKT_PM(i,r)..

  vim(i,r) * M(i,r) =E=
        SUM(d, A(d,i,r) * vm(d,i,r) * ( PA(d,i,r)/PM(i,r) )**esubdm );

*       International transport:

MKT_PT..        YT * vt =E=  sum((i,r,s), A_M(i,r,s) * M(i,s) * tau(i,r,s));

*       Armington supply:

MKT_PA(d,i,r)..

  va(d,i,r) * A(d,i,r) =E= sum(j, vafm(i,j,r) * Y(j,r))$i_d(d)
        + (A_C(i,r) * C(r))$c_d(d) + ( A_G(i,r) *  G(r))$g_d(d);

*       Government provision:

MKT_PG(r)..     G(r) =E= 1;

*       Factor market:

MKT_PF(f,r)..   evoa(f,r) =E= sum(i, A_F(f,i,r) * Y(i,r));


*       Private demand:

MKT_PC(r)..     C(r) * vp(r) =E= RA(r) / PC(r) ;
```

Table 18: Income Balance Equations in the MCP Formulation

```
INC_RA(r)..

  RA(r) =E= sum(f,   PF(f,r) * evoa(f,r))
        + sum(num, PC(num) * vb(r))
        - sum(cgd, PD(cgd,r) * vdm(cgd,r))
        - PG(r) * vg(r)

* Output tax:

        +    sum(i, ty(i,r)*(PX(i,r)*A_X(i,r)+PD(i,r)*A_D(i,r))*Y(I,R) )

* Tax on intermediate demand:

        +    sum((i,j), ti(j,i,r) * PA("i",j,r) * vafm(j,i,r) * Y(i,r) )

* Taxes on factor use:

        +    sum( (i,f),  tf(f,i,r) * PF(f,r) * A_F(f,i,r) * Y(i,r))

* Export tax:

        +     sum((i,s), tx(i,r,s) * PX(i,r) * A_M(i,r,s) * M(i,s) )

* Import tariff applies to mechandise gross of export tax
* and transport cost:

        +    sum((i,s), tm(i,s,r) * A_M(i,s,r) * M(i,r) *
                (  PX(i,s) * (1+tx(i,s,r)) + PT * tau(i,s,r) ) )

* Taxes on government consumption:

        +    sum(i, tg(i,r) * PA("g",i,r) * A_G(i,r) * G(r))

* Taxes on private consumption:

        +    sum(i, tp(i,r) * PA("c",i,r) * A_C(i,r) * C(r));
```

3.2 The GTAP-EG Model

In this section, an illustrative static model based on the GTAP-EG dataset is presented. We start with a description of the flows of goods and factors in the model. Then a general structure of the Arrow-Debreu model in Mathiesen format is discussed. We present basic blocks of the core model and their nesting structure as implemented in GAMS-MPSGE. MPSGE (Mathematical Programming System for General Equilibrium) is a compact and powerful programming language for economic modelling developed by Rutherford [1999][28].

The simplified structure of the regional flows of goods and factors is presented in Figure 2. The world is divided into regions. Each region incorporates markets for electricity, E, and non-electric energy, N. Non-electric energy includes: oil, gas and coal. Crude oil may be produced domestically or imported, and it is then refined prior to delivery as an input to production and final demand. Electricity is not traded, and produced using coal, oil, gas or non-fossil inputs. Final energy products are supplied both as inputs to production and final demand.

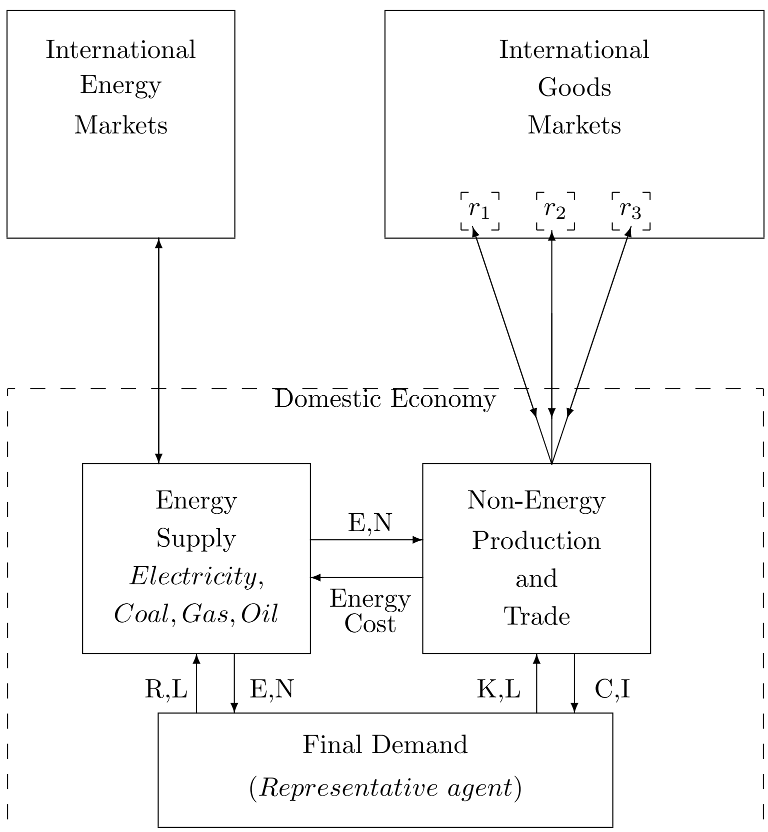

Fig. 2. Regional Flows of Goods and Factors

Consumption in each region is associated with utility maximization by a representative agent subject to a budget constraint. The agent supplies primary factors (capital, K, labor, L, and energy resources, R) to non-energy and energy sectors. Factor income of each representative agent is then allocated to the purchase of energy (E and N), non-energy goods (C), and investment (I). Regions are connected with the global economy through trade in energy and non-energy

[28]MPSGE syntax can be found at http://debreu.colorado.edu/mainpage/mpsge.htm

goods. Energy trade involves primarily crude oil and coal which can be exported or imported in international markets.

The core model described here is a static, multi-regional model which tracks the production and distribution of goods in the global economy. The model is an Arrow-Debreu general economic equilibrium model concerning the interaction of consumers and producers in markets. Lars Mathiesen [1985] proposed a representation of this class of models in which two types of equations define an equilibrium: zero profit and market clearance. The corresponding variables defining an equilibrium are activity levels (for constant-returns-to-scale firms) and commodity prices.[29]

Commodity markets merge primary endowments of households with producer outputs. In equilibrium the aggregate supply of each good must be at least as great as total intermediate and final demand. Initial endowments are exogenous. Producer supplies and demands are defined by producer activity levels and relative prices. Final demands are determined by market prices.

Economists who have worked with conventional textbook equilibrium models can find Mathiesen's framework to be somewhat opaque because many quantity variables are not explicitly specified in the model. Variables such as final demand by consumers, factor demands by producers and commodity supplies by producers, are defined implicitly in Mathiesen's model. For example, given equilibrium prices for primary factors, consumer incomes can be computed, and given income and goods prices, consumers' demands can then be determined. The consumer demand functions are written down in order to define an equilibrium, but quantities demanded need not appear in the model as separate variables. The same is true of inputs or outputs from the production process: relative prices determine conditional demand, and conditional demand times the activity level represents market demand. Omitting decisions variables and suppressing definitional equations corresponding to intermediate and final demand provides significant computational advantages at the cost of a somewhat more complex model statement.

The flows represented in Figure 2 are implemented in the GTAP-EG model in the following way. In the model there are two types of produced commodities, fossil-fuel and non-fossil fuel commodities. The model assumes that goods produced in different regions are qualitatively distinct (Armington [1968]). This implies that trade in goods is represented as flows between pairs of countries rather than from individual countries and an integrated global market. Every bilateral trade flow requires its own transportation services. Primary factors in each region include labor, capital and fossil-fuel resources. Labor is mobile within domestic borders but cannot move between regions. Capital can be global or region-specific. Natural resources are sector-specific.

Now we turn to a formulation of the GTAP-EG model in MPSGE format. The MPSGE framework is based on nested constant elasticity of substitution utility functions and production functions. MPSGE uses a concept of representing these functions as separate "blocks". We describe the basic blocks only. Some exception operators are omitted here to make the code easier to read[30]. Appendix 5 contains listing of the GAMS-MPSGE code.

In the GTAP-EG model an economy in region r consists of three production blocks. The block $y(i,r)$ is related to production, where fossil-fuel production has a different structure from other production sectors. We implicitly introduced a production block for Armington supply which represents an aggregation between domestic and import varieties and across imports from different trading partners. Armington aggregation is described by the block $a(i,r)$. Armington supply is used then for private consumption and as an intermediate input to production. Private consumption is presented by the block $c(r)$. Finally, a production block yt describes the provision of international transport services.

In order to represent consumption, another class of the MPSGE variables is introduced. In

[29]Under a maintained assumption of perfect competition, Mathiesen may characterize technology as CRTS without loss of generality. Decreasing returns are accommodated through introduction of a specific factor, while increasing returns are inconsistent with the assumption of perfect competition. In this environment zero excess profit is consistent with free entry for atomistic firms producing an identical product.

[30]GAMS has a special operator used for exception handling. It is denoted as a dollar sign. The exception operator is very useful, for example, in the cases when we want to represent some sectors of an economy which may not be active in a benchmark. For more information, see GAMS User's Guide.

each region the representative agent (described by a consumption block $ra(r)$) depicts a collective decision process for allocating income to households and to a government. Both $c(r)$ and $ra(r)$ MPSGE blocks are needed because final consumption is taxed and taxes cannot be imposed on a demand block.

Regions may apply domestic carbon taxes. Carbon tax revenue is collected by the representative agent in each region. Within this model, the carbon tax policy is equivalent to an emission permit system where the permit price coincides with the carbon tax. There are also taxes on output ty, intermediate inputs ti, consumption tc, export tx, and import tm. Figure 3 depicts the structure of the GTAP-EG model.

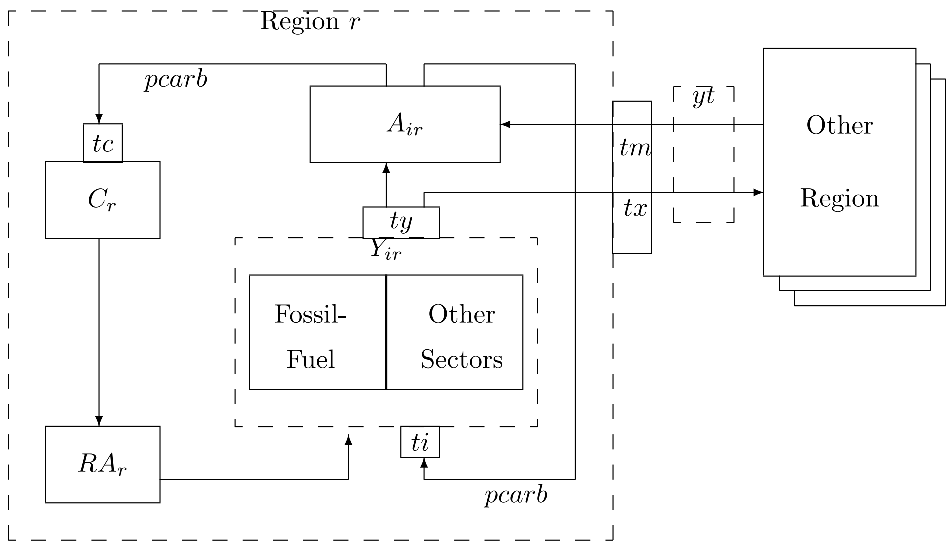

Fig. 3. Structure of the GTAP-EG model.

An MPSGE model is specified by endogenous variable declarations and a sequence of function "blocks", one for each production sector and consumer in the model. A declaration of the GTAP-EG variables is presented below. Key words in the declaration as follows. $\$SECTORS$: describes production activities that convert commodity inputs into commodity outputs. The variable associated with a sector is the activity level. $\$COMMODITIES$ are related to a good or factor. The variable associated with a commodity is its price, not its quantity. $\$CONSUMERS$ denote individuals who demand commodities, supply factors and receive tax and other revenues. The variable associated with a consumer is income from all sources.

```
$sectors:
        c(r)                    ! Private consumption
        y(i,r)                  ! Output
        a(i,r)                  ! Armington aggregation
        yt                      ! Transport

$commodities:
        pc(r)                   ! Final demand
        py(i,r)                 ! Output price
        pa(i,r)                 ! Armington composite price
        pl(r)                   ! Wage rate
        pr(i,r)                 ! Energy resource
        rkr(r)$rsk              ! Return to regional capital
        rkg$gk                  ! Return to global capital
```

```
        pt                      ! Transport services
        pcarb(r)                ! Carbon permits -- non-tradable

$consumers:
        ra(r)                   ! Representative agent
```

According to Figure 3 and the declarations above, the GTAP-EG model includes sectors related to production by commodity and region $y(i,r)$; Armington aggregation between domestic and import varieties and across imports from different trading partners $a(i,r)$; the provision of international transport services yt; and private demand by region $c(r)$.

The production activity for private demand is associated with an output which represent the marginal cost of private consumption $pc(r)$. For each commodity and region there are three different price indices: $py(i,r)$ represents the cost index for a unit of output; $pa(i,r)$ is the cost index of a unit of composite Armington supply; $pr(i,r)$ represents the cost index for energy resource. Labor is mobile within a region and the wage rate is $pl(r)$. Capital may be region-specific $rkr(r)$ or global rkg. The market price of a unit of international transport services is represented by pt. Emission permit price is $pcarb(r)$.

The final class of variables in the MPSGE model are the consumers, and in this model there is one representative consumer for each region. In equilibrium, $ra(r)$ is a variable representing income of the consumer in the region r.

3.2.1 Production

Fossil fuel production activities includes crude, gas, and coal. Production has the structure shown in Figure 4, where a value to the right of the arc represents an elasticity. Fossil fuel output ($y(xe)$, where xe is one type of exhaustible energy: crude, gas, coal) is produced as an aggregate of a resource input ($pr(xe)$) and a non-resource input composite. The non-resource input for the production is a fixed - coefficient (Leontief) composite of labor (pl) and the Armington aggregation ($pa(i)$) of domestic and imported intermediate input from a production sector i. The elasticity of substitution between pa and pl equals to zero ($id : 0$), which characterizes a Leontief composite. The elasticity of substitution ($s : esub\_es$) between the resource input and a non-resource input composite depends on the value share of resource inputs in fossil fuel supply.

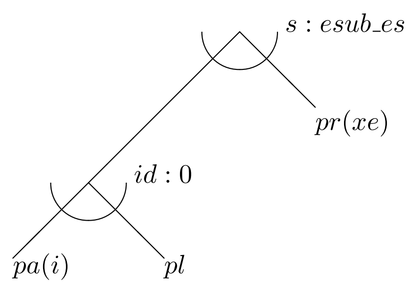

Fig. 4. Fossil fuel production

Production block for the fossil fuel production $y(xe,r)$ (where xe is a set of exhaustible energy) has the following implementation in MPSGE. There are inputs (i : fields) and outputs (o : field) associated with a production block. Each of them has an associated reference quantity (q : field) and reference price (p : field). If a reference price is equal to zero, then the price field can be omitted. The nesting structure consists of two nests with top level elasticity equal to $s : esub\_es$, and the elasticity between intermediate and labor inputs equals 0 ($id : 0$).

Output taxes ty and intermediate input taxes ti are collected by a representative agent in region r. The field a : shows who collects a tax, and the field t : determines a tax rate. For example, taxes are levied on intermediate demand inputs at net rate ti. The market value of intermediate inputs purchased by firms is $vafm(j, xe, r)$, but the total cost to firms equals $vafm(j, xe, r) * (1 + ti(j, xe, r))$, of which $vafm(j, xe, r)$ is paid to sellers of intermediate inputs $vafm(j, xe, r) * ti(j, xe, r)$ is paid as a tax to $ra(r)$.

```
*       Fossil fuel production activity (crude, gas and coal):

$prod:y(xe,r)$vom(xe,r)  s:(esub_es(xe,r))  id:0

o:py(xe,r)      q:vom(xe,r)     a:ra(r) t:ty(xe,r)
i:pa(j,r)       q:vafm(j,xe,r)  p:pai0(j,xe,r)  a:ra(r) t:ti(j,xe,r) id:
i:pl(r)         q:ld0(xe,r)  id:
i:pr(xe,r)      q:rd0(xe,r)
```

Non-fossil fuel production (including electricity and refining) has a different structure. Figure 5 illustrates the nesting and typical elasticities employed in production sectors other than fossil fuels. Output is produced with fixed-coefficient (Leontief) inputs of intermediate non-energy goods and an energy-primary factor composite. The energy-primary factor composite is a constant-elasticity of substitution (CES) function with elasticity = 0.5. Primary factor inputs of labor and capital are aggregated through a Cobb-Douglas production function ($va : 1$). The energy composite is a CES function of electricity versus other energy inputs, coal versus liquid fuels, and oil versus gas.

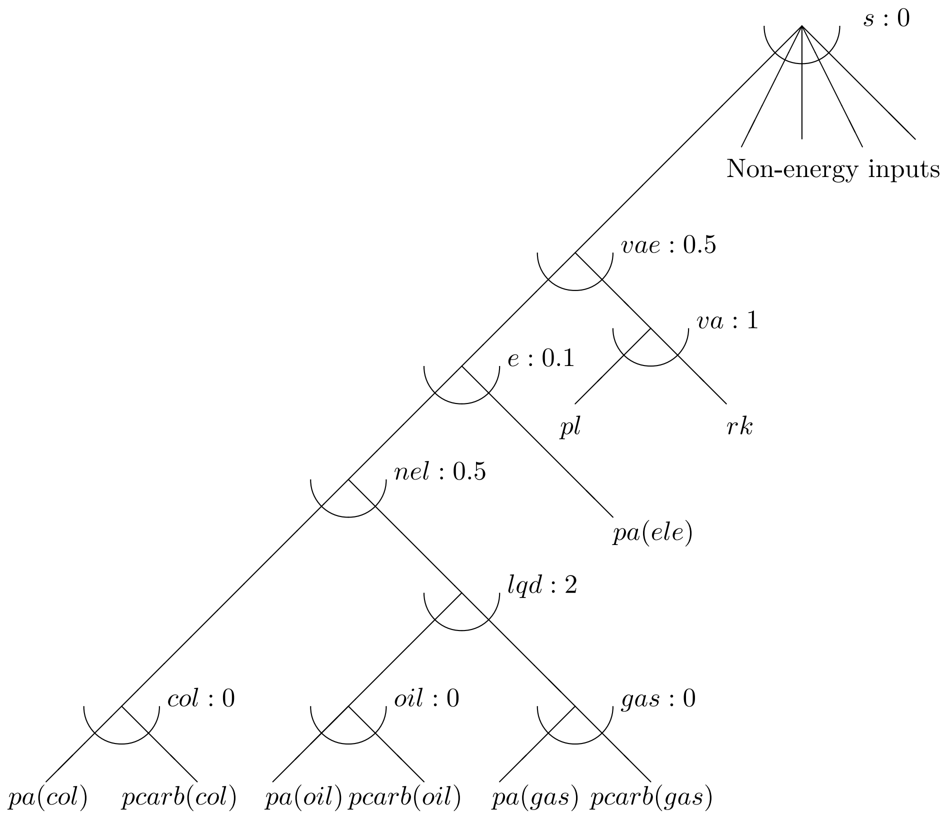

Fig. 5. Non-fossil fuel production

The following MPSGE block describes production of non-energy products ($y(i,r)$). This block is generated only for non-fossil fuel sectors, as noted by the inclusion of the exception operator $\$nr(i,r)$, where $nr(i,r)$ is a set created as `nr(i,r) = yes$(vom(i,r)$(not xe(i)));`. This GAMS line means that a sector i in a region r is included in the set $nr(i,r)$ if an aggregate output value $vom(i,r)$ of sector i is not equal to zero and the sector i is not an exhaustible energy sector xe.

```
*       Non-fossil fuel production (includes electricity and refining):

$prod:y(i,r)$nr(i,r)   s:0  vae(s):0.5  va(vae):1
+                       e(vae):0.1  nel(e):0.5 lqd(nel):2
+                       oil(lqd):0 col(nel):0 gas(lqd):0

o:py(i,r)       q:vom(i,r)  a:ra(r) t:ty(i,r)
i:pa(j,r)$(not fe(j)) q:vafm(j,i,r) p:pai0(j,i,r) e:$ele(j) a:ra(r) t:ti(j,i,r)
i:pl(r)         q:ld0(i,r)                      va:
i:rkr(r)$rsk    q:kd0(i,r)                      va:
i:rkg$gk        q:kd0(i,r)                      va:
i:pcarb(r)#(fe) q:carbcoef(fe,i,r) p:1e-6   fe.tl:
i:pa(fe,r)      q:vafm(fe,i,r)  p:pai0(fe,i,r) fe.tl: a:ra(r) t:ti(fe,i,r)
```

The nesting of the production block is clearly more complicated than for fossil-fuel production. First, note that $i : pa$ appears in two places of the production block. This is because Armington composite enters into production differently for different sectors. The line $i : pa(j,r)\$(not fe(j))$ defines it for non final energy sectors, where fe denotes a set of a final energy (oil, coal, gas). The top level elasticity ($s :$) equals to zero. It has a subnest $vae(s)$, which in turn has two subnests $va(vae)$ and $e(vae)$. An elasticity $e :$ is only applied for electricity, which is shown by an exception operator $e : ele(j)$. It means that the elasticity for all non final energy sectors except electricity is equal to the top level elasticity ($s : 0$). Capital and labor are in $va :$ nest.

The final energy has a special treatment in the line $i : pa(fe,r)$. The line of elasticities $nel(e) : lqd(nel) : oil(lqd) : col(nel) : gas(lqd) :$ shows that the final energy is a subnest of $e :$. As such, intermediate inputs in the form of each final energy and its associated carbon tax enter as fixed-coefficient composites defined by an elasticity of substitution equal to zero ($fe.tl : 0$). The suffix ($.tl$) represents a GAMS text label for a set element. It is used here to represent the nest generated for a set of final energy fe as shown in Figure 5.

Again, a representative agent in region r collects output taxes ty and intermediate input taxes ti. Carbon tax $pcarb(r)$ is levied if production uses oil, gas, or coal as an intermediate input. The $i : pcarb(r)\#(fe)$ input represents the fact that the tax is applied for each element of the set fe.

3.2.2 Armington Supply

Armington aggregation activity generates intermediate demand for production and final demand for consumption as a mix of domestic and imported goods as imperfect substitutes. We assume that the domestic-imports elasticity of substitution (d) equals to 4, while the elasticity of substitution among import sources (m) equals to 8. Imports from every region require transportation services (pt) which are implemented as shown in Figure 6 for region S.

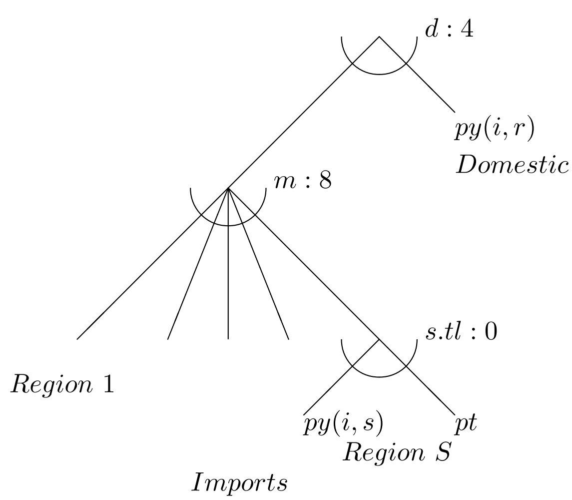

Fig. 6. Armington aggregation

The MPSGE function declaration indicates a top-level substitution elasticity between domestic ($py(i,r)$) and imported ($py(i,s)$) goods equal to four ($s : 4$). Then it defines the aggregation of imports from a trading partner with the second-level substitution elasticity between imported commodities equal to eight ($m : 8$). It applies export taxes (tx) and import tariffs (tm) on all bilateral trades. Note that the $i : py(i,s)$ input also represents fob payments to producers in region s, and as such export taxes on sales from region s to region r are accrued to the representative agent in region s ($a : ra(s)$) while import tariffs are paid to the representative agent in region r ($a : ra(r)$).

The Armington supply block also applies transportation margins which are proportional to quantities traded. The `i:pt#(s)` input represents multiple inputs of transportation services, one for each element of set s. There are multiple inputs of transportation services into each imported good because every bilateral trade flow demands its own transportation services. Using a Leontief aggregate on each bilateral trade flow assures that transport costs and imports remain strictly proportional to the base year level.

```
*       Armington aggregation over domestic versus imports:

$prod:a(i,r)$a0(i,r)  s:4  m:8 s.tl(m):0
        o:pa(i,r)       q:a0(i,r)
        i:py(i,r)       q:d0(i,r)
        i:py(i,s)       q:vxmd(i,s,r)  p:pmx0(i,s,r) s.tl:
+               a:ra(s) t:tx(i,s,r)     a:ra(r) t:(tm(i,s,r)*(1+tx(i,s,r)))
        i:pt#(s)       q:vtwr(i,s,r)  p:pmt0(i,s,r) s.tl: a:ra(r) t:tm(i,s,r)
```

3.2.3 International Transport

The international transport services are assumed to be a Cobb-Douglas composite of goods provided in the domestic markets in each region, as shown in Figure 7.

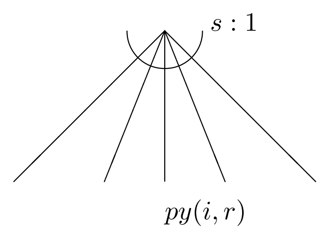

Fig. 7. International transport services

The MPSGE representation shows yt as a Cobb-Douglas ($s:1$) composite of goods provided in the domestic markets of each region.

```
*       International transport services (Cobb-Douglas):

$prod:yt  s:1
        o:pt           q:(sum((i,r), vst(i,r)))
        i:py(i,r)       q:vst(i,r)
```

3.2.4 Final Demand

Final demand has the structure shown in Figure 8. Utility in each country is a constant elasticity aggregate of non-energy consumption and energy. The non-energy composite is in turn a Cobb-Douglas aggregate of different goods while final energy is a Cobb-Douglas aggregate of electricity, oil, gas, and coal.

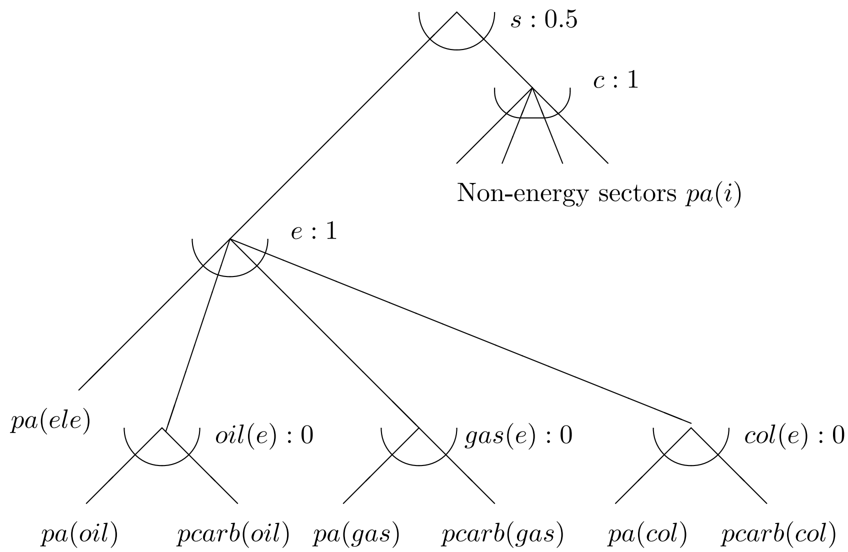

Fig. 8. Final demand

Final consumption in region r is characterized by activity $c(r)$, which is a constant elasticity aggregate ($s : 0.5$) of non-energy consumption and energy. The non-energy composite is in turn a Cobb-Douglas ($c : 1$) aggregate of different goods while final energy is a Cobb-Douglas ($e : 1$) aggregate of electric energy, oil, gas, and coal. Carbon tax $pcarb(r)$ is applied to a final demand.

```
*       Final demand

$prod:c(r)  s:0.5 c:1 e:1  oil(e):0 col(e):0 gas(e):0

o:pc(r)   q:ct0(r)
i:pa(i,r) q:c0(i,r) p:pc0(i,r) i.tl:$fe(i) c:$(not e(i)) e:$ele(i) a:ra(r) t:tc(i,r)
i:pcarb(r)#(fe)     q:carbcoef(fe,"final",r) p:1e-6  fe.tl:
```

The model statement concludes with a specification of endowment and demand for each region's representative agent. Each agent is endowed with primary factors, capital inflows, non-tradable carbon permits and collects tax revenue. The income is allocated to investment and private demand. Representative agents are "endowed" with a fixed negative quantity of the domestic "CGD" commodity representing the exogenously-specified demand for investment. Private demand is determined by utility maximizing behavior.

```
$demand:ra(r)
        d:pc(r)         q:ct0(r)
        e:py("cgd",r)   q:-vom("cgd",r)
        e:rkr(r)$rsk    q:(sum(i, kd0(i,r)))
        e:rkg$gk        q:(sum(i, kd0(i,r)))
        e:pl(r)         q:evoa("lab",r)
        e:pr(xe,r)      q:rd0(xe,r)
        e:pc("usa")     q:vb(r)
        e:pcarb(r)      q:carblim(r)
```

3.2.5 An Illustrative Calculation: Leakage Rate

In this section we show a calculation of a leakage rate based on the GTAP-EG model. The leakage rate is defined as the ratio of total carbon emissions by non-Annex B countries to total emissions abatement by the Annex B. This means that if the leakage rate is 50%, then a decrease in carbon emissions by the Annex B countries of 100 million tons will lead to the increase in carbon emissions by the non-Annex B countries of 50 million tons.

The formula for the leakage rate is

$$LeakageRate = \sum_n 100 \cdot \frac{CARB_{1,n} - CARB_{0,n}}{\sum_m CARB_{0,m} - CARB_{1,m}} \tag{1}$$

where n denotes the non-Annex B countries, m represents the Annex B, $CARB_0$ is the level carbon emissions in a benchmark, $CARB_1$ denotes carbon emissions in a counterfactual scenario. The calculation of the leakage rate can be implemented in GAMS in the following way.

```
leakage(r)$(not annexb(r)) = 100 * (scncarbon(r) - baucarbon(r)) /
   sum(annexb(rr), baucarbon(rr) - scncarbon(rr));

leakage("total")=sum(r, leakage(r));
```

The results of calculations based on the GTAP-EG model using the two different aggregated versions of the energy-economy datasets are as follows: GTAP-E-FIT - 8.1%, GTAP-EG - 11.5%. Once again, it shows a small difference between the datasets in their representation of the energy flows.

4 Using the Dataset

This section describes a usage of the GTAPinGAMS package. There are two versions of the GTAP datasets in the package: the full versions (file `gtap001.zip` for the GTAPinGAMS and `aspen.zip` for the GTAP-EG) and the aggregated versions (file `gtap_small.zip` for the GTAPinGAMS and `aspen_small.zip` for the GTAP-EG). The full versions of the datasets have a restricted distribution. In order to use the full versions, a modeller needs to obtain the standard GTAP4 data file `gsddat.har`[31], and then download the GTAPinGAMS distribution archive with a build routine which prepares the GTAP data for calibration, and combines the GTAP with the energy data. The instructions for using the GTAPinGAMS package are given below[32]. For testing purposes, aggregated versions of the GTAPinGAMS and GTAP-EG datasets are provided with the archive. A modeller can test the aggregated datasets by running illustrative models. The modeller can change the models to suit his specific purpose because he has a full access to the code of the build routines and the models.

4.1 System Requirements

You will need to have the following GAMS system components:

- GAMS compiler version 2.50
- PATH complementarity solver
- MINOS5 nonlinear optimizer and nonlinear system solver (for creating the GTAP-EG dataset and running NLP models)

[31] The instructions for obtaining the GTAP data can be found at `http://www.agecon.purdue.edu/gtap/`

[32] Short directions are also given in the file `README.TXT` of the `GTAPinGAMS` archive.

- LIBINCLUDE Tools for Writing GAMS-Readable Data Files (optional)[33]
- A Pentium computer running Windows 95 or NT with more than 100 MB of free disk space.

4.2 Download

The GAMSinGAMS package is distributed in a zip archive (`gtapgams.zip`) file. You can download it from `http://debreu.colorado.edu/gtap/gtapgams.zip`. The archive has the directory structure presented in Tables 19 and 20.

After downloading the file `gtapgams.zip` into your computer, unzip the file making sure that the archive's directory structure is preserved[34]. Both the GTAPinGAMS and GTAP-EG have two versions: full and aggregated. In order to get the full versions, a user needs to run the build routines `BUILD.BAT` and `ASPEN.BAT` described below. The aggregated datasets are provided for testing purposes. They are located in the `DATA` subdirectory and ready to use. A description of the aggregated datasets is given below.

4.3 The build routine BUILD.BAT

The GTAPinGAMS dataset is built on the standard GTAP4 database, which is not distributed freely. In order to construct the full GTAP datasets, a user needs to contact GTAP at:
`http://www.agecon.purdue.edu/gtap/` to obtain the GTAP4 dataset (the file `gsddat.har`).

The file `gsddat.har` should be placed into the `DATA` subdirectory of the GTAPinGAMS package. To create the full version of the GTAPinGAMS, a user needs to run `build.bat` file. For this, in MS-DOS prompt[35] type `build` and press `Enter`. Run time on Pentium 500 is about 8 minutes[36].

The file `build.bat` is intended to:

- Read `gsddat.har` file.
- Convert `gsddat.har` into `gtap.gms`.
- Relabel and scale the data to create `gtap.zip`.
- Filter and recalibrate the data to `gtap001.zip`.

Once you successfully run the BUILD routine, you can use the full GTAPinGAMS dataset which is placed in the DATA subdirectory. Your computer may not have enough memory to run your models on the full dataset. Therefore, the aggregation routine GTAPAGGR.BAT (described below) can aggregate the original GTAP data into the smaller datasets. You can choose what aggregation is appropriate for studying your particular application.

4.4 The build routine ASPEN.BAT

The GTAP-EG dataset is built on the standard GTAP4 database and calibrated to energy quantity and price data. In order to construct the full GTAP-EG dataset, a user needs to run `aspen.bat` file. For this, In MS-DOS prompt, type `aspen` and press `Enter`. Run time on Pentium 500 is

[33]The GTAP-EG build routine and the model use the LIBINCLUDE tools located in the INCLIB directory of the GTAP-EG distribution package. In order to be able to use the tools in your own applications, you need to install them into GAMS directory. The latest version of the LIBINCLUDE tools is distributed as a file `inclib.pck`. To install it on your computer download the file from http://nash.colorado.edu/tomruth/inclib/inclib.pck into your GAMS system directory, and run GAMSINST. A description of `inclib.pck` can be found at http://nash.colorado.edu/tomruth/inclib/gams2txt.htm

[34]The files from the ZIP archive can be extracted by using `WinZip.exe` or `unzip.exe`. WinZip can be downloaded from `http://www.winzip.com`

[35]Make sure that you are connected to a proper directory.

[36]Make sure that GAMS is included in the PATH variable of your computer's MS-DOS. To check it, in MS-DOS prompt type `path` and press `Enter`.

Table 19: Structure of the archive GTAPGAMS.ZIP

Directory	Purpose	File	Purpose
BUILD	build dataset	aspen.bat	Batch file to build the GTAP-EG dataset
		build.bat	Batch file to build the GTAPinGAMS dataset
		chkdata.bat	Check benchmark data
		chkdata.gms	Check benchmark data
		chkeq.gms	Check equilibrium
		filter.gms	Filter data
		gtapaggr.bat	Aggregation
		gtapaggr.gms	Aggregation GAMS file
		gtapsets.gms	Make sets for aggregation
		iea_vail.gms	Calibration
		ieo.dat	IEO Projections Data
		ieodata.gms	IEO Projections
		impose.bat	New dataset creation
		impose.gms	New dataset creation
		recalib.gms	Recalibration
		regbal.gms	Regional Balance
		relabel.gms	Relabel to GAMS format
		seehar.exe	Read HAR-file utility
DATA	initial and resulting data	vail.dat	Aggregated Energy Data
		aspen_small.zip	13x8 GTAP-EG dataset
		gtap_small.zip	10x10 GTAPinGAMS dataset
		(gsddat.har)	GTAP Data (not included)
		(aspen.zip)	Created after running ASPEN.BAT
		(gtap001.zip)	Created after running BUILD.BAT
DEFINES	stores .set, .map, and .def files	aspen.map	Full GTAP-EG mapping
		aspen.set	Full GTAP-EG sets
		aspen_small.map	13x8 GTAP-EG mapping
		aspen_small.set	13x8 GTAP-EG sets
		gtap_small.map	10x10 GTAPinGAMS sets
		gtap_small.set	10x10 GTAPinGAMS sets
		gtap.set	GTAP4 sets
		iea.map	IEA-GTAP mapping
		iea.set	IEA-GTAP sets
		notax.def	New tax rates definitions

Table 20: Structure of the archive GTAPGAMS.ZIP (cont.)

INCLIB	standard GAMS utilities	aggr.gms	Aggregation
		checkset.gms	Check set
		chktarget.gms	Check target
		gams2har.gms	Move to HAR format
		gams2prm.gms	Move to GAMS parameter
		gams2tbl.gms	Move to table
		gams2txt.gms	Move to text file
		gdpreport.gms	Report GDP
		har2gams.gms	Move from HAR format
		mrtdata.gms	GTAP parameters calculation
		unzip.gms	Call unzip
		zip.gms	Call zip
MODELS	Illustrative models	gtap-eg.gms	The GTAP-EG Model
		runall.bat	Run all GTAPinGAMS Models
		mrttest.gms	Test all GTAPinGAMS Models
		cnstest.gms	Test CNS GTAPinGAMS Model
		mcptest.gms	Test MCP GTAPinGAMS Model
		mgetest.gms	Test MPSGE GTAPinGAMS Model
		nlptest.gms	Test NLP GTAPinGAMS Model
		mrtcns.gms	CNS GTAPinGAMS Model
		mrtmcp.gms	MCP GTAPinGAMS Model
		mrtmge.gms	MPSGE GTAPinGAMS Model
		mrtnlp.gms	NLP GTAPinGAMS Model
readme.txt	Installation Directions		

about 4 minutes. If the full version of the GTAPinGAMS dataset has not been created yet, then the BUILD routine (described above) will be automatically called.

The file `aspen.bat` is intended to:

- Aggregate `gtap001.zip` to a dataset compatible with the IEA data: `iea.zip`.
- Calibrate the GTAP and IEA energy data to create `gtap1000.zip`.
- Relabel the energy commodities, translating `gtap1000.zip` to `aspen.zip`.
- Delete work files.
- Give to a user an option of creating the aggregated GTAP-EG dataset and running an illustrative model.

Users can edit `aspen.bat` to suit their specific applications. In particular, a pause option can be uncommented for every step to see all the stages of the dataset creation[37]. The original build routine has a pause in one place only - before the aggregation and running the illustrative model. At this point you'll see a message

```
 Aggregate to 13x8, include energy
projections and create aspen_small.zip Will aggregate to
aspen_small.zip: Press any key to continue . . .
```

If you press any key, then in addition to the full dataset an aggregated dataset will be created (a new file `aspen_small.zip` will replace an old one) and an illustrative model will run on an aggregated data. If you press "Ctrl-C", then the following message appears.

```
 Terminate batch job (y/n)?
```

If you terminate the batch job at this time (by pressing "y" and then "Enter"), `aspen.bat` will stop and only the full version of the GTAP-EG dataset will be created and placed into `DATA` subdirectory under the name `aspen.zip`.

In the process of building the dataset, several echo files are placed in the `ASPEN` subdirectory:

- `iea.ech` - Report on economic activity by sector and region from GTAP data;
- `energy1000.ech` - Energy statistics;
- `aspen.ech` - Report on economic activity by sector and region from the full GTAP-EG dataset;
- `aspen_small.ech` - Report on economic activity by sector and region from the aggregated GTAP-EG dataset.

4.5 The Aggregated GTAPinGAMS Dataset: GTAP_SMALL.ZIP

The GTAPinGAMS package contains an aggregated version of the GTAPinGAMS dataset. It is provided for testing purposes. A user can check the installation by running the models on the aggregated dataset (file `gtap_small.zip`), which is located in the DATA subdirectory and ready to use. The archive contains the data file `GTAP_SMALL.GMS`, and associated `SET` and `MAP` files. The dataset has 10 sectors, 10 regions, and 5 primary factors. Table 21 lists the set definitions. The dataset is provided by the DOE and is intended to focus on competitiveness impacts of measures intended to reduce global carbon dioxide emissions. The regional aggregation for the DOE dataset is motivated by the nature of Kyoto agreement with separate representation of the US, Japan, the EU and China. Aggregate regions in the model include Other OECD, Former Soviet Union, Central European Associates, Other Asia, Mexico plus OPEC and Rest of World. Commodities in the DOE dataset include the investment aggregate, five energy goods, metals-related industry, other energy-intensive, other manufactures and services.

[37] To uncomment a `pause` command, delete a :(column) sign, i.e. change a line from `:pause` to `pause`.

Table 21: Set Definitions for The Aggregated GTAPinGAMS Dataset

```
SET   I   Sectors/
        MTL     Metals-related industry (IRONSTL & NONFERR)
        EIS     Other energy intensive (CHEMICAL & PAPERPRO)
        MFR     Other manufactures
        SER     Other Services
        COL     Coal
        OIL     Petroleum and coal products (refined)
        CRU     Crude oil
        GAS     Natural gas
        ELE     Electricity
        CGD     Savings good /;

SET R Aggregated Regions /
        USA     United States
        JPN     Japan
        EUR     Europe
        OOE     Other OECD
        CHN     China
        FSU     Former Soviet Union
        CEA     Central European Associates
        ASI     Other Asia
        MPC     Mexico plus OPEC
        ROW     Other countries /;

SET   F   Factors of production
/       LND     Land,
        SKL     Skilled labor,
        LAB     Unskilled labor,
        CAP     Capital,
        RES     Natural resources /;
```

4.6 The Aggregated GTAP-EG Dataset: ASPEN_SMALL.ZIP

The `GTAPGAMS.ZIP` archive contains an aggregated version of the GTAP-EG dataset. It is located in `DATA` subdirectory and named `ASPEN_SMALL.ZIP`. The archive contains the data file `ASPEN_SMALL.GMS`, and associated `SET` and `MAP` files. The aggregated dataset has 13 regions, 8 goods, and two primary factors. The identifiers for the aggregated GTAP-EG dataset are contained in the `SET` file, which is provided in Table 22.

```
Table 22. Set Definitions for The Aggregated GTAP-EG Dataset

SET   I   Sectors/
Y Other manufactures and services
EIS Energy-intensive sectors
COL Coal
OIL     Petroleum and coal products (refined)
CRU Crude oil
GAS Natural gas
ELE Electricity
CGD Savings good/;

SET R Aggregated Regions /
USA United States
CAN Canada
EUR Europe
JPN Japan
OOE Other OECD
FSU Former Soviet Union
CEA Central European Associates
CHN China (including Hong Kong + Taiwan)
IND India
BRA Brazil
ASI Other Asia
MPC Mexico + OPEC
ROW Rest of world /

Set   F  Aggregated factors /
       LAB      Labor,
       CAP      Capital /;
```

4.7 The Aggregation Routine GTAPAGGR.BAT

Once you have built the initial GTAPinGAMS or GTAP-EG dataset, you can begin to think about a particular application and which aggregations of the original GTAP-EG data would be appropriate for studying those issues. Typically it is useful to create two aggregations for any new model, one with a minimal number of regions and commodities and another with a larger number of dimensions. The small aggregation can then be used for model development.

The `gtapaggr.bat` program is used to aggregate the GTAPinGAMS and the GTAP-EG datasets. A command line argument defines the name of the target aggregation. You only need to provide the batch file with the target because the target's mapping file defines the source. Before running `gtapaggr.bat`, you must create two files, one defining the sets of commodities, regions and primary factors in the target dataset, and another defining the name of the source dataset and a correspondence between elements of the source and target. The aggregation routine produces a brief report of GDP and trade shares in the new dataset. The SET and MAP files for a new dataset are GAMS-readable files located in the `DEFINES` subdirectory. An example of aggregating the full GTAP-EG dataset to `ASPEN_SMALL` is given below.

Step 1. Creating SET and MAP files. Appendix 6 shows a sample set file `aspen_small.set` defining the identifiers of the resulting dataset `ASPEN_SMALL`. The file defines the sets of goods, regions, and primary factors which are in the model. Appendix 7 presents the associated mapping file, `aspen_small.map`. The file provides a definition of the source dataset together with mapping definitions for commodities and factors. When no mapping is defined for the set of regions, the aggregation routine retains the same set as in the source data. In order to run the GTAP-EG model on the aggregated dataset two requirements[38] should be fulfilled: (a) commodity `CGD`, the investment-savings composite, must be included in every aggregation; (b) primary factors should be aggregated into capital and labor. For aggregating the GTAPinGAMS dataset only the requirement (a) should be met.

Step 2. Placing the files into the proper subdirectories. The files should placed into the proper subdirectories: the files `aspen_small.set` and `aspen_small.map` into the `DEFINES` subdirectory, and the source datafile `aspen.zip` into the `DATA` subdirectory. Make sure that you have the files `gtapaggr.bat`, `gtapaggr.gms`, and `gtapsets.gms` in the `BUILD` subdirectory.

Step 3. Running the aggregation routine. To run the aggregation routine, go to `BUILD` subdirectory, type `gtapaggr aspen_small` at MS-DOS prompt, and press "Enter". The target dataset `aspen_small.zip` will be placed in the `DATA` subdirectory [39]. Now you can define your own SET and MAP and create your own aggregated datasets using `gtapaggr.bat`. It should be noted that the aggregation routine also includes the energy projections into the aggregated dataset if the source file is `aspen.zip`.

4.8 The recalibration routine IMPOSE.BAT

The program `impose.bat` is used to create a new balanced dataset by imposing a new set of benchmark tax rates on an existing GTAP dataset. The syntax of the command is: `impose target source`, where `target` is the name of the new dataset which is going to be created and `source` is the name of the dataset from which the data is taken. An example of creation the dataset in which all domestic taxes are eliminated is given below. The newly created dataset is called `notax`. We use `gtap_small.zip` as a source file.

Step 1. Creating DEF file. A user needs to create a definitions file and place it into the DEFINES subdirectory. The new dataset will have the same name as the definitions file. An example of the file where all domestic taxes are eliminated is provided in the DEFINES subdirectory of the GTAPinGAMS package and presented in Table 23.

```
Table 23. Benchmark Tax Definitions File: notax.def

* Set up a benchmark equilibrium in which we eliminate all domestic taxes:

ty(i,r) = 0;
tp(i,r) = 0;
tf(f,i,r) = 0;
tg(i,r) = 0;
ti(i,j,r) = 0;
```

When you write the definitions file for adjusting tax rates, bear in mind that a gross basis tax (`ty`) is defined as a percentage of the gross-of-tax price, hence these tax rates have a maximum

[38] To run the GTAP-EG model "as is", a region "USA" should be present in every aggregation. Otherwise, a user needs to change a numeraire region in the line `e:pc(''usa'')` in the demand block of the GTAP-EG model.

[39] SET and MAP files are provided with the GTAPinGAMS package. An aggregation to `aspen_small.zip` is done automatically if you run `aspen.bat`

value of 100% and no minimum. A net basis tax, such as `tf`, `tp`, `tg`, `tx` or `tm` is defined as a percentage of the net-of-tax price, hence these tax rates have no maximum value and a minimum value of -100%.

Step 2. Running the recalibration routine. Before running the `impose.bat`, check that you have the following files: `gtap_small.map`, `gtap_small.set`, and `notax.def` in the DEFINES subdirectory; `gtap_small.zip` in the DATA subdirectory. Then, in MS-DOS prompt type: `impose notax gtap_small` and press "Enter".

The `impose` command generates a dataset named `notax.zip` from `gtap_small.zip` using information from a benchmark revision file `notax.def` in the DEFINES subdirectory. This command also copies the set definition file `gtap_small.set` to `notax.set` and the mapping file `gtap_small.map` to `notax.map`, all in the DEFINES subdirectory.[40] The program also generates a summary echo-print of trade and GDP shares for the new dataset and places this file in the BUILD subdirectory.

4.9 The benchmark data testing routine CHKDATA.BAT

The program `chkdata.bat` reads the dataset and checks benchmark consistency, producing an echoprint of base year GDP and trade shares. The syntax is: `chkdata dataset`, where `dataset` is the name of the dataset which is going to be checked for consistency.

4.10 The model testing routine RUNALL.BAT

The program `runall.bat` tests the GTAPinGAMS model in four different formulations and compares the results. These four models files are located in the MODELS subdirectory. They are:

- `mrtmge.gms` The static multiregional model specified as an mixed complementarity model using an MPSGE representation of demand and supply functions.
- `mrtmcp.gms` The static multiregional model specified as an mixed complementarity model with GAMS algebra.
- `mrtcns.gms` The static multiregional model specified as an constrained nonlinear system with the GAMS CNS model type.
- `mrtnlp.gms` The static multiregional model specified as an square nonlinear system within a GAMS nonlinear program.

To test the GTAPinGAMS installation, type in MS-DOS prompt: `runall gtap_small`. It calls the `mrttest.gms` program. This GAMS program processes four other programs, `mgetest.gms`, `mcptest.gms`, `cnstest.gms`, and `nlptest.gms` each of which executes a benchmark replication check, a benchmark clean-up and a single counter-factual scenario. Output from these four programs is collected in `mrt.sol` and the results are compared by `mrttest.gms`. If your installation is correct, you should see the message:

```
No difference detected in the solutions from alternative solvers.
```

displayed on the screen.[41]

These tests are conducted with the `gtap_small` dataset. You can choose to use a different existing dataset (for example, called `another`) for the test if you change the argument of the batch file call to: `runall another`.

---

[40] The mapping file is copied if one can be found. This is done to assure that it is always possible to trace the aggregation definitions for any dataset.

[41] The first calculation which is performed is a benchmark replication check in which a solver may report "INFEASIBLE". This simply means that there is some imprecision in the data, as is subsequently reported in the listing as "Benchmark tolerance". Any number on the order of 1.e-4 or smaller indicates a reasonably precise dataset.

The GTAPinGAMS models also can be called directly. For example, if you need to run MPSGE formulation of the model, type in MS-DOS prompt: `gams mrtmge.gms`. By default, the models are set to the aggregated GTAPinGAMS dataset. The dataset can be changed by correcting the following line in the model GAMS code: `$IF NOT SETGLOBAL dataset $SETGLOBAL dataset gtap_small`

4.11 Running the GTAP-EG model

An illustrative static GTAP-EG model (file `gtap-eg.gms`) is included in the GTAPinGAMS package. It recreates the benchmark equilibrium and also calculates a leakage rate, which is the ratio of total increased carbon emissions by non-Annex B countries to total emissions abatement by the Annex B.

In the GTAP-EG illustrative model, we restrict carbon emissions by 25% and calculate the leakage rate. You can run the model by typing `gams gtap-eg` in MS-DOS prompt. The results are placed in the listing file `gtap-eg.lst`. There is an option of using global (GK) or region-specific (RSK) capital in the model. By default, region-specific capital option is used, which is defined by the following GAMS scalar.

```
SCALAR  RSK     Flag for region-specific capital /1/,
        GK      Flag for global capital          /0/;
```

A switch to the global capital specification may be done by changing the scalars. The GTAP-EG model is set for the calculations based on the aggregated `ASPEN_SMALL` dataset. A user may run the model on a dataset created from his own aggregation. It can be done by changing the name of the dataset in the line `$setglobal dataset aspen_small` in the file `gtap-eg.gms`.

References

[1] Armington, P. (1969). *A Theory of Demand for Products Distinguished by Place of Production.* IMF Staff Papers, 16, 159-178.

[2] Babiker M.H. and T.F. Rutherford, "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. 1997.

[3] Brooke A., D. Kendrick, and A. Meeraus, *GAMS: A User's Guide, Release 2.25*, Scientific Press, 1992.

[4] Bruce J.P., H. Lee, and E.F.Haites (eds.), *Climate Change 1995: Economic and Social Dimensions of Climate Change*, Cambridge University Press, 1996.

[5] Complainville C. and D. van der Mensbrugghe, "Construction of an Energy Database for GTAP V4: Concordance with IEA Energy Statistics", OECD Development Centre, 1998.

[6] Ferris, M.C. and T.F. Rutherford, "Matrix Balancing: A Practical Approach", *University of Colorado Working Paper*, 1998.

[7] Harrison, W.J. and K.R. Pearson, "Computing solutions for large general equilibrium models using GEMPACK", *Computational Economics*, 9:83-127, 1996.

[8] Harrison, G.W., T.F. Rutherford, and D. Tarr, "Opciones de poltica comercial para Chile: una evaluacin cuantitiva" *Cuadernos de Economa*, August 1997. A related working paper in English is available on the web, (see `http://theweb.badm.sc.edu/glenn/addreg.pdf`).

[9] Hertel, T.W. (ed.) *Global Trade Analysis: Modeling and Applications*, Cambridge University Press, Cambridge and New York, 1997.

[10] International Energy Agency. *CO_2 Emissions from Fuel Combustion. A New Basis for Comparing Emissions of a Major Greenhouse Gas.* OECD/IEA, Paris, 1997.

[11] Malcolm G. and T.P.Truong, "The Process of Incorporating Energy Data into GTAP", GTAP Technical Paper, 1999.

[12] Mathiesen, L. "Computation of Economic Equilibrium by a Sequence of Linear Complementarity Problems", *Mathematical Programming Study 23*, North-Holland, 1985, pp. 144-162.

[13] McDougall, R. "The GTAP Database", *Draft documentation.* See the GTAP 4 release (http://www.agecon.purdue.edu/gtap/database).

[14] Rutherford, T.F. "Applied General Equilibrium Modeling with MPSGE as a GAMS Subsystem: An overview of the Modeling Framework and Syntax", Computational Economics, V.14, Nos. 1-2, 1999.

[15] Rutherford, T.F. and S.V.Paltsev, "GTAP-EG: Incorporating energy statistics into GTAP format", University of Colorado Department of Economics, 2000.

[16] Rutherford, T.F., M. Lau and A. Pahlke, "A Primer in Dynamic General Equilibrium Analysis", Working Paper, University of Colorado Department of Economics, 1998.

[17] Rutherford, T.F., and D. Tarr, "Regional Trading Arrangements for Chile: Do the Results Differ with a Dynamic Model?", University of Colorado Department of Economics, 1998.

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

Appendix 1. GTAP Identifiers

Appendix 1 presents sectoral (Tables A.1 and A.2), regional (Table A.3), and primary factors (Table A.4) identifiers in the GTAPinGAMS and GTAP-EG datasets. They both have 45 regions and 5 primary factors. The GTAPinGAMS dataset has 50 sectors, while the GTAP-EG dataset has 23 sectors (5 of which are energy sectors).

Table A.1. Sectoral identifiers in the Full GTAPinGAMS Dataset

PDR	Paddy rice,	B_T	Beverages and tobacco,
WHT	Wheat,	TEX	Textiles,
GRO	Grains (except rice-wheat),	WAP	Wearing apparel,
V_F	Vegetable fruit nuts,	LEA	Leather goods,
OSD	Oil seeds,	LUM	Lumber and wood,
C_B	Sugar cane and beet,	PPP	Pulp and paper,
PFB	Plant-based fibers,	P_C	Petroleum and coal products,
OCR	Crops n.e.c.,	CRP	Chemicals rubber and plastics,
CTL	Bovine cattle,	NMM	Non-metallic mineral products,
OAP	Animal products n.e.c.,	I_S	Primary ferrous metals,
RMK	Raw milk,	NFM	Non-ferrous metals,
WOL	Wool,	FMP	Fabricated metal products,
FRS	Forestry,	MVH	Motor vehicles,
FSH	Fishing,	OTN	Other transport equipment,
COL	Coal,	ELE	Electronic equipment,
OIL	Oil,	OME	Machinery and equipment,
GAS	Natural Gas,	OMF	Other manufacturing products,
OMN	Other Minerals,	ELY	Electricity,
CMT	Bovine cattle meat products,	GDT	Gas manuf. and distribution,
OMT	Meat products n.e.c.,	WTR	Water,
VOL	Vegetable oils,	CNS	Construction,
MIL	Dairy products,	T_T	Trade and transport,
PCR	Processed rice,	OSP	Other services (private),
SGR	Sugar,	OSG	Other services (public),
OFD	Other food products,	DWE	Dwellings,
		CGD	Investment composite

Table A.2. Sectoral identifiers in the Full GTAP-EG Dataset

GAS	Natural gas works	FPR	Food products
ELE	Electricity and heat	PPP	Paper-pulp-print
OIL	Refined oil products	LUM	Wood and wood-products
COL	Coal	CNS	Construction
CRU	Crude oil	TWL	Textiles-wearing apparel-leather
I_S	Iron and steel industry	OMF	Other manufacturing
CRP	Chemical industry	AGR	Agricultural products
NFM	Non-ferrous metals	T_T	Trade and transport
NMM	Non-metallic minerals	SER	Commercial and public services
TRN	Transport equipment	DWE	Dwellings,
OME	Other machinery	CGD	Investment composite
OMN	Mining		

Table A.3. Regional identifiers in the Full GTAPinGAMS and GTAP-EG Datasets

AUS	Australia (*),	ARG	Argentina,
NZL	New Zealand (*),	BRA	Brazil,
JPN	Japan (*),	CHL	Chile,
KOR	Republic of Korea,	URY	Uruguay,
IDN	Indonesia,	RSM	Rest of South America,
MYS	Malaysia,	GBR	United Kingdom (*),
PHL	Philippines,	DEU	Germany (*),
SGP	Singapore,	DNK	Denmark (*),
THA	Thailand,	SWE	Sweden (*),
VNM	Vietnam,	FIN	Finland (*),
CHN	China,	REU	Rest of EU (*),
HKG	Hong Kong,	EFT	European Free Trade Area(*),
TWN	Taiwan,	CEA	Central European Associates (*),
IND	India,	FSU	Former Soviet Union (*),
LKA	Sri Lanka,	TUR	Turkey,
RAS	Rest of South Asia,	RME	Rest of Middle East,
CAN	Canada (*),	MAR	Morocco,
USA	United States of America (*),	RNF	Rest of North Africa,
MEX	Mexico,	SAF	South Africa,
CAM	Central America and Caribbean,	RSA	Rest of South Africa,
VEN	Venezuela,	RSS	Rest of Sub-Saharan Africa,
COL	Columbia,	ROW	Rest of World
RAP	Rest of Andean Pact,		

The Annex B regions are denoted by (*). CEA includes Bulgaria, Czech Republic, Hungary, Poland, Romania, Slovakia, and Slovenia. REU includes Austria, Belgium, Spain, France, Giblartar, Greece, Ireland, Italy, Luxembourg, Netherlands, and Portugal. EFT includes Switzerland, Iceland, and Norway. FSU includes Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Lithuania, Latvia, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan.

Table A.4. Primary Factor Identifiers in the Full GTAPinGAMS Dataset

LND	Land,
SKL	Skilled labor,
LAB	Unskilled labor,
CAP	Capital,
RES	Natural resources

Appendix 2. Aggregation of IEA regions into GTAP format

Country	IEA code	Region	GTAP-EG code
Australia	AUS	Australia	AUS
New Zealand	NZL	New Zealand	NZL
Japan	JPN	Japan	JPN
Korea	KOR	Korea	KOR
Indonesia	IDN	Indonesia	IDN
Malaysia	MYS	Malaysia	MYS
Philippines	PHL	Phillipines	PHL
Singapore	SGP	Singapore	SGP
Thailand	THA	Thailand	THA
Vietnam	VNM	Vietnam	VNM
China	CHN	China	CHN
Hong Kong	HKG	Hong Kong	HKG
Taiwan	TWN	Taiwan	TWN
India	IND	India	IND
Sri Lanka	LKA	Sri Lanka	LKA
Bangladesh	RAS_BGD	Rest of South Asia	RAS
Nepal	RAS_NPL	Rest of South Asia	RAS
Pakistan	RAS_PAK	Rest of South Asia	RAS
Canada	CAN	Canada	CAN
USA	USA	USA	USA
Mexico	MEX	Mexico	MEX
Antilles	CAM_ANT	Central America and Carribean	CAM
Costa Rica	CAM_CRI	Central America and Carribean	CAM
Cuba	CAM_CUB	Central America and Carribean	CAM
Dominican Republic	CAM_DOM	Central America and Carribean	CAM
Guatemala	CAM_GTM	Central America and Carribean	CAM
Honduras	CAM_HND	Central America and Carribean	CAM
Haiti	CAM_HTI	Central America and Carribean	CAM
Jamaica	CAM_JAM	Central America and Carribean	CAM
Nicaragua	CAM_NIC	Central America and Carribean	CAM
Panama	CAM_PAN	Central America and Carribean	CAM
El Salavador	CAM_SLV	Central America and Carribean	CAM
Trinidad & Tobago	CAM_TTO	Central America and Carribean	CAM
Venezuela	VEN	Venezuela	VEN
Columbia	COL	Columbia	COL
Bolivia	RAP_BOL	Rest of Andean Pact	RAP
Ecuador	RAP_ECU	Rest of Andean Pact	RAP
Peru	RAP_PER	Rest of Andean Pact	RAP
Argentina	ARG	Argentina	ARG
Brazil	BRA	Brazil	BRA
Chile	CHL	Chile	CHL
Uruguay	URY	Uruguay	URY
Paraguay	RSM_PRY	Rest of South America	RSM
Great Britain	GBR	Great Britain	GBR
Germany	DEU	Germany	DEU
Denmark	DNK	Denmark	DNK
Sweden	SWE	Sweden	SWE
Finland	FIN	Finland	FIN

Austria	REU_AUT	Rest of European Union	REU
Belgium	REU_BEL	Rest of European Union	REU
Spain	REU_ESP	Rest of European Union	REU
France	REU_FRA	Rest of European Union	REU
Giblartar	REU_GIB	Rest of European Union	REU
Greece	REU_GRC	Rest of European Union	REU
Ireland	REU_IRL	Rest of European Union	REU
Italy	REU_ITA	Rest of European Union	REU
Luxembourg	REU_LUX	Rest of European Union	REU
Netherlands	REU_NLD	Rest of European Union	REU
Portugal	REU_PRT	Rest of European Union	REU
Switzerland	EFT_CHE	European Free Trade Area	EFT
Iceland	EFT_ISL	European Free Trade Area	EFT
Norway	EFT_NOR	European Free Trade Area	EFT
Bulgaria	CEA_BGR	Central European Associates	CEA
Czech Republic	CEA_CZE	Central European Associates	CEA
Hungary	CEA_HUN	Central European Associates	CEA
Poland	CEA_POL	Central European Associates	CEA
Romania	CEA_ROM	Central European Associates	CEA
Slovakia	CEA_SVK	Central European Associates	CEA
Slovenia	CEA_SVN	Central European Associates	CEA
Armenia	FSU_ARM	Former Soviet Union	FSU
Azerbaijan	FSU_AZE	Former Soviet Union	FSU
Belarus	FSU_BLR	Former Soviet Union	FSU
Estonia	FSU_EST	Former Soviet Union	FSU
Georgia	FSU_GEO	Former Soviet Union	FSU
Kazakhstan	FSU_KAZ	Former Soviet Union	FSU
Kyrgyzstan	FSU_KGZ	Former Soviet Union	FSU
Lithuania	FSU_LTU	Former Soviet Union	FSU
Latvia	FSU_LVA	Former Soviet Union	FSU
Moldova	FSU_MDA	Former Soviet Union	FSU
Russia	FSU_RUS	Former Soviet Union	FSU
Tajikistan	FSU_TJK	Former Soviet Union	FSU
Turkmenistan	FSU_TKM	Former Soviet Union	FSU
Ukraine	FSU_UKR	Former Soviet Union	FSU
Uzbekistan	FSU_UZB	Former Soviet Union	FSU
Turkey	TUR	Turkey	TUR
United Arab Emirates	RME_ARE	Rest of Middle East	RME
Bahrain	RME_BHR	Rest of Middle East	RME
Iran	RME_IRN	Rest of Middle East	RME
Iraq	RME_IRQ	Rest of Middle East	RME
Israel	RME_ISR	Rest of Middle East	RME
Jordan	RME_JOR	Rest of Middle East	RME
Kuwait	RME_KWT	Rest of Middle East	RME
Lebanon	RME_LBN	Rest of Middle East	RME
Oman	RME_OMN	Rest of Middle East	RME
Qatar	RME_QAT	Rest of Middle East	RME
Saudi Arabia	RME_SAU	Rest of Middle East	RME
Syria	RME_SYR	Rest of Middle East	RME
Yemen	RME_YEM	Rest of Middle East	RME

Morocco	MAR	Morocco	MAR
Algeria	RNF_DZA	Rest of North Africa	RNF
Egypt	RNF_EGY	Rest of North Africa	RNF
Libya	RNF_LBY	Rest of North Africa	RNF
Tunisia	RNF_TUN	Rest of North Africa	RNF
South Africa CU	SAF	South Africa	SAF
Angola	RSA_AGO	Rest of South Africa	RSA
Mozambique	RSA_MOZ	Rest of South Africa	RSA
Tanzania	RSA_TZA	Rest of South Africa	RSA
Zambia	RSA_ZMB	Rest of South Africa	RSA
Zimbabwe	RSA_ZWE	Rest of South Africa	RSA
Benin	RSS_BEN	Rest of South-Saharan Africa	RSS
Cote d'Ivoire	RSS_CIV	Rest of South-Saharan Africa	RSS
Cameroon	RSS_CMR	Rest of South-Saharan Africa	RSS
Congo	RSS_COG	Rest of South-Saharan Africa	RSS
Ethiopia	RSS_ETH	Rest of South-Saharan Africa	RSS
Gabon	RSS_GAB	Rest of South-Saharan Africa	RSS
Ghana	RSS_GHA	Rest of South-Saharan Africa	RSS
Kenya	RSS_KEN	Rest of South-Saharan Africa	RSS
Nigeria	RSS_NGA	Rest of South-Saharan Africa	RSS
Sudan	RSS_SDN	Rest of South-Saharan Africa	RSS
Senegal	RSS_SEN	Rest of South-Saharan Africa	RSS
Zaire	RSS_ZAR	Rest of South-Saharan Africa	RSS
Albania	ROW_ALB	Rest of World	ROW
Bosnia	ROW_BIH	Rest of World	ROW
Brunei	ROW_BRN	Rest of World	ROW
Cyprus	ROW_CYP	Rest of World	ROW
Croatia	ROW_HRV	Rest of World	ROW
Macedonia	ROW_MKD	Rest of World	ROW
Malta	ROW_MLT	Rest of World	ROW
Myanmar	ROW_MMR	Rest of World	ROW
Papua New Guinea	ROW_PNG	Rest of World	ROW
North Korea	ROW_PRK	Rest of World	ROW
Serbia	ROW_SER	Rest of World	ROW
Other Africa	OTHERAFRIC	Rest of World	ROW
Other Asia	OTHERASIA	Rest of World	ROW
Other Latin America	OTHERLATIN	Rest of World	ROW

Appendix 3. An aggregation of production sectors into GTAP-EG format

Appendix 3 describes the mapping of IEA and GTAP 4 production sectors into GTAP-EG format. For more details, see Rutherford and Paltsev [2000] where the process of incorporating of IEA statistics into GTAP-EG is described. The original IEA statistics has 35 sectors. The following table presents a concordance between IEA and GTAP-EG production sectors.

IEA code	Sector	GTAP-EG sector
COL	Coal	COL
AGR	agriculture	AGR
CNS	Construction	CNS
CRP	Chemical and Petrochemical	CRP
DWE	Dwellings	DWE and final consumption
ELY	Electricity	ELE
EXPORTS	Exports	goes to export data
FPR	Food and Tobacco	FRP
GAS	Gas	GAS
HEAT	Heat	Not used
I_S	Iron and steel	I_S
IMPORTS	Imports	goes to import data
INDPROD	Indigenous production	Not used
LUM	Wood products	LUM
NEINTREN	Non energy use in industry	CRP
NEOTHER	Non-energy use in other sectors	AGR
NETRANS	Non-energy use in transport	T_T
NFM	Non ferrous metals	NFM
NMM	Non metallic minerals	NMM
NONROAD	Other (non road) transport	T_T
OIL	Oil	CRU
OME	Machinery	OME
OMF	Other manufacturing	OMF
OMN	Mining	OMN
OWNUSE	Ownuse	Not used
P_C	Petroleum	OIL
PPP	Paper, Pulp, and Print	PPP
RENEW	Renewable	Not used
ROAD	Road	Part to T_T and part to final consumption
SER	Services	SER
TRN	Transport equipment	TRN
TWL	Textile and leather	TWL

An aggregation of GTAP 4 into GTAP-EG is done with the aggregation routine `gtapaggr`, described in Section 4. The following table shows the mapping.

GTAP 4	GTAP-EG	Sector
GDT, GAS	GAS	Natural gas works
ELY	ELE	Electricity and heat
P_C	OIL	Refined oil products
COL	COL	Coal transformation
OIL	CRU	Crude oil
I_S	I_S	Iron and steel industry
CRP	CRP	Chemical industry
NFM	NFM	Non-ferrous metals
NMM	NMM	Non-metallic minerals
MVH, OTN	TRN	Transport equipment
ELE, OME, FMP	OME	Other machinery
OMN	OMN	Mining
OMT, VOL, MIL, PCR, SGR, OFD, B_T, CMT	FPR	Food products
PPP	PPP	Paper-pulp-print
LUM	LUM	Wood and wood-products
CNS	CNS	Construction
TEX, WAP, LEA	TWL	Textiles-wearing apparel-leather
OMF, WTR	OMF	Other manufacturing
PDR, WHT, GRO, V_F, OSD, C_B, PFB, OCR, CTL, OAP, RMK, WOL, FRS, FSH	AGR	Agricultural products
T_T	T_T	Trade and transport
OSP, OSG	SER	Commercial and public services
DWE	DWE	Dwellings
CGD	CGD	Investment composite

Appendix 4. GTAP-EG: Basic statistics

Table A.4.1. Economic activity by sector

	gdp	gdp%	trade	trade%
DWE	104.0	4.1		
ELE	93.8	3.7		
CNS	159.9	6.3	2.2	0.4
COL	12.0	0.5	2.3	0.4
GAS	14.6	0.6	3.2	0.5
NMM	21.0	0.8	7.3	1.2
OIL	18.4	0.7	8.5	1.4
OMN	5.8	0.2	9.1	1.5
LUM	19.1	0.7	11.0	1.8
NFM	5.5	0.2	11.3	1.8
OMF	25.5	1.0	15.3	2.5
PPP	41.6	1.6	16.1	2.6
I_S	20.6	0.8	18.5	3.0
CRU	37.1	1.5	21.3	3.4
AGR	120.3	4.7	25.9	4.2
FPR	76.0	3.0	35.1	5.6
TWL	44.2	1.7	46.4	7.5
SER	892.3	35.0	46.4	7.5
T_T	505.5	19.8	53.3	8.6
TRN	55.0	2.2	58.0	9.3
CRP	84.4	3.3	64.1	10.3
OME	190.9	7.5	165.8	26.7

Table A.4.2. Economic activity by region

	gdp	gdp%	trade	trade%
RSM	0.4	0.0	0.4	0.1
URY	1.4	0.1	0.4	0.1
LKA	1.2	0.0	0.5	0.1
VNM	1.2	0.0	0.7	0.1
MAR	2.6	0.1	1.0	0.2
COL	6.9	0.3	1.5	0.2
RSA	1.6	0.1	1.5	0.2
RAP	7.4	0.3	1.6	0.3
RAS	6.9	0.3	1.7	0.3
CHL	5.5	0.2	2.0	0.3
VEN	6.8	0.3	2.0	0.3
NZL	5.1	0.2	2.2	0.3
PHL	5.9	0.2	2.8	0.4
ARG	24.9	1.0	2.9	0.5
ROW	22.0	0.9	3.3	0.5
SAF	12.7	0.5	3.5	0.6
TUR	15.6	0.6	3.8	0.6
RNF	10.7	0.4	3.9	0.6

RSS	13.6	0.5	4.3	0.7
CAM	7.2	0.3	4.4	0.7
IND	27.7	1.1	4.4	0.7
FIN	11.6	0.5	4.9	0.8
IDN	19.6	0.8	5.7	0.9
BRA	62.9	2.5	6.2	1.0
DNK	15.5	0.6	6.4	1.0
AUS	31.8	1.2	7.2	1.2
THA	14.9	0.6	7.5	1.2
HKG	9.9	0.4	8.2	1.3
MEX	25.2	1.0	8.9	1.4
SWE	19.3	0.8	9.2	1.5
MYS	7.1	0.3	9.3	1.5
FSU	44.8	1.8	11.4	1.8
CEA	27.8	1.1	11.7	1.9
SGP	6.0	0.2	13.3	2.1
TWN	24.6	1.0	15.1	2.4
RME	39.8	1.6	15.8	2.5
KOR	39.7	1.6	16.0	2.6
EFT	40.8	1.6	16.6	2.7
CAN	49.7	2.0	21.1	3.4
CHN	55.5	2.2	23.7	3.8
GBR	101.3	4.0	29.6	4.8
JPN	463.1	18.2	54.3	8.7
DEU	222.1	8.7	58.6	9.4
USA	655.8	25.7	79.5	12.8
REU	372.0	14.6	132.2	21.3

Table A.4.3. Carbon inventories -- mton

	total	ind_nele	fd_nele	electric	ind_total	fd_total	kg/$
AUS	78.0	33.2	9.8	35.0	60.8	17.1	0.2
NZL	8.8	6.8	1.2	0.8	7.4	1.4	0.2
JPN	342.8	198.3	54.8	89.7	269.7	73.0	0.1
KOR	122.4	83.5	18.0	20.9	101.4	21.0	0.3
IDN	64.0	40.3	12.3	11.5	48.8	15.2	0.3
MYS	23.1	12.8	3.7	6.6	18.4	4.6	0.3
PHL	12.2	7.2	1.9	3.1	9.7	2.5	0.2
SGP	23.2	16.8	0.8	5.6	21.6	1.6	0.4
THA	38.4	18.2	8.2	12.0	28.1	10.3	0.3
VNM	5.4	4.0	0.6	0.8	4.6	0.8	0.5
CHN	848.8	534.0	78.5	236.4	745.1	103.7	1.6
HKG	13.8	7.5	0.4	5.8	12.2	1.6	0.1
TWN	49.8	28.9	4.8	16.1	42.1	7.7	0.2
IND	210.9	88.1	26.4	96.4	172.4	38.5	0.8
LKA	2.1	1.7	0.3	0	1.7	0.3	0.2
RAS	27.4	14.8	5.5	7.1	20.3	7.1	0.4
CAN	138.1	83.9	28.6	25.6	104.1	34.0	0.3
USA	1489.2	613.2	337.1	539.0	1014.5	474.8	0.2
MEX	89.6	54.5	16.3	18.8	70.1	19.5	0.4

CAM	27.2	17.5	2.7	7.0	23.5	3.8	0.4
VEN	33.1	22.2	5.8	5.1	26.4	6.7	0.5
COL	17.8	10.8	4.1	2.9	12.9	4.8	0.3
RAP	13.8	9.8	2.5	1.5	11.0	2.7	0.2
ARG	33.4	15.6	12.2	5.6	20.0	13.4	0.1
BRA	78.9	61.5	14.1	3.3	64.2	14.7	0.1
CHL	11.3	6.9	2.6	1.9	8.5	2.8	0.2
URY	1.6	1.2	0.3	0	1.3	0.3	0.1
RSM	0.9	0.4	0.5	0	0.4	0.5	0.2
GBR	165.6	84.9	37.4	43.3	117.9	47.7	0.2
DEU	265.4	118.4	64.4	82.6	184.2	81.2	0.1
DNK	18.6	7.7	2.7	8.2	13.9	4.7	0.1
SWE	17.5	11.1	4.4	2.1	12.6	4.9	0.1
FIN	16.2	8.4	2.4	5.4	12.7	3.5	0.1
REU	473.1	267.7	106.9	98.5	346.6	126.4	0.1
EFT	25.3	17.5	7.4	0.3	17.8	7.5	0.1
CEA	208.1	91.3	25.0	91.8	167.2	40.9	0.8
FSU	695.1	324.6	72.3	298.2	576.6	118.5	1.7
TUR	45.9	27.5	7.1	11.3	37.0	8.9	0.3
RME	225.6	133.4	39.4	52.8	175.2	50.4	0.6
MAR	7.3	3.7	1.0	2.7	5.7	1.6	0.3
RNF	56.5	32.3	9.2	15.1	44.5	12.1	0.5
SAF	96.0	44.1	10.9	41.0	79.8	16.2	0.8
RSA	7.2	4.5	0.6	2.1	6.3	0.9	0.5
RSS	22.7	16.0	4.4	2.3	17.9	4.8	0.2
ROW	56.8	32.0	5.6	19.2	47.2	9.6	0.3
total	6208.5	3218.4	1054.9	1935.1	4784.3	1424.1	

Table A.4.4. Carbon emissions as a percentage of global carbon emissions

ANNEX B

	as % of annex	as % of total
AUS	1.978	1.256
NZL	0.222	0.141
JPN	8.696	5.521
CAN	3.503	2.224
USA	37.782	23.987
GBR	4.202	2.668
DEU	6.732	4.274
DNK	0.471	0.299
SWE	0.445	0.282
FIN	0.411	0.261
REU	12.002	7.620
EFT	0.642	0.407
CEA	5.279	3.352
FSU	17.636	11.197
annex b	100.000	63.488

NON-ANNEX B

	as % of non-annex	as % of total
KOR	5.398	1.971
IDN	2.824	1.031
MYS	1.018	0.372
PHL	0.539	0.197
SGP	1.023	0.374
THA	1.694	0.618
VNM	0.237	0.086
CHN	37.446	13.672
HKG	0.607	0.222
TWN	2.195	0.801
IND	9.303	3.397
LKA	0.091	0.033
RAS	1.207	0.441
MEX	3.951	1.442
CAM	1.202	0.439
VEN	1.460	0.533
COL	0.784	0.286
RAP	0.608	0.222
ARG	1.471	0.537
BRA	3.479	1.270
CHL	0.501	0.183
URY	0.070	0.025
RSM	0.039	0.014
TUR	2.024	0.739
RME	9.954	3.634
MAR	0.322	0.118
RNF	2.495	0.911
SAF	4.235	1.546
RSA	0.316	0.115
RSS	1.003	0.366
ROW	2.504	0.914
non-annex b	100.000	36.512

Table A.4.5. Carbon dioxide emissions - billion of tonnes

	IEA book	IEA stat	GTAP-E-FIT	EG with no fix	GTAP-EG
AUS	0.286	0.286	0.283	0.286	0.286
NZL	0.029	0.032	0.033	0.032	0.032
JPN	1.151	1.208	1.145	1.257	1.257
KOR	0.353	0.449	0.396	0.449	0.449
IDN	0.227	0.235	0.212	0.235	0.235
MYS	0.092	0.085	0.084	0.085	0.085
PHL	0.050	0.045	0.044	0.045	0.045
SGP	0.059	0.085	0.085	0.085	0.085

THA	0.156	0.140	0.140	0.141	0.141
VNM	0.022	0.020	0.021	0.020	0.020
CHN	3.007	3.098	2.902	3.112	3.112
HKG	0.044	0.052	0.052	0.050	0.050
TWN	0.167	0.182	0.179	0.182	0.182
IND	0.803	0.771	0.765	0.773	0.773
LKA	0.006	0.008	0.007	0.008	0.008
RAS	0.211	0.100	0.097	0.100	0.100
CAN	0.471	0.505	0.472	0.506	0.506
USA	5.228	5.339	5.175	5.340	5.460
MEX	0.328	0.328	0.309	0.328	0.328
CAM	0.111	0.097	0.100	0.100	0.100
VEN	0.113	0.114	0.112	0.121	0.121
COL	0.065	0.063	0.062	0.065	0.065
RAP	0.052	0.050	0.047	0.051	0.051
ARG	0.128	0.121	0.115	0.122	0.122
BRA	0.287	0.269	0.256	0.289	0.289
CHL	0.042	0.042	0.039	0.042	0.042
URY	0.005	0.006	0.006	0.006	0.006
RSM	0.003	0.003	0.004	0.003	0.003
GBR	0.565	0.605	0.540	0.607	0.607
DEU	0.884	0.973	0.865	0.973	0.973
DNK	0.060	0.067	0.063	0.068	0.068
SWE	0.056	0.064	0.061	0.064	0.064
FIN	0.054	0.059	0.057	0.059	0.059
REU	1.560	1.734	1.628	1.735	1.735
EFT	0.078	0.093	0.082	0.093	0.093
CEA	0.749	0.762	0.707	0.763	0.763
FSU	2.483	2.542	2.341	2.549	2.549
TUR	0.160	0.168	0.156	0.168	0.168
RME	0.817	0.788	0.755	0.827	0.827
MAR	0.026	0.027	0.026	0.027	0.027
RNF	0.213	0.204	0.201	0.207	0.207
SAF	0.321	0.347	0.337	0.352	0.352
RSA	0.025	0.026	0.026	0.026	0.026
RSS	0.081	0.083	0.103	0.083	0.083
ROW	0.518	0.208	0.183	0.208	0.208
total	22.150	22.482	21.272	22.644	22.764

Appendix 5. MPSGE formulation

Appendix 5 presents the function declarations for GTAP-EG model implemented in MPSGE.

```
*       Final demand
$prod:c(r)  s:0.5 c:1 e:1  oil(e):0 col(e):0 gas(e):0
  o:pc(r)   q:ct0(r)
  i:pa(i,r) q:c0(i,r) p:pc0(i,r) i.tl:$fe(i) c:$(not e(i)) e:$ele(i) a:ra(r) t:tc(i,r)
  i:pcarb(r)#(fe)     q:carbcoef(fe,"final",r) p:1e-6  fe.tl:

*       Non-fossil fuel production (includes electricity and refining):
$prod:y(i,r)$nr(i,r)   s:0  vae(s):0.5  va(vae):1
+                       e(vae):0.1  nel(e):0.5 lqd(nel):2
+                       oil(lqd):0 col(nel):0 gas(lqd):0

        o:py(i,r)       q:vom(i,r)  a:ra(r) t:ty(i,r)
        i:pa(j,r)$(not fe(j)) q:vafm(j,i,r) p:pai0(j,i,r) e:$ele(j) a:ra(r) t:ti(j,i,r)
        i:pl(r)         q:ld0(i,r)                      va:
        i:rkr(r)$rsk    q:kd0(i,r)                      va:
        i:rkg$gk        q:kd0(i,r)                      va:
        i:pcarb(r)#(fe) q:carbcoef(fe,i,r) p:1e-6   fe.tl:
        i:pa(fe,r)      q:vafm(fe,i,r)  p:pai0(fe,i,r) fe.tl: a:ra(r) t:ti(fe,i,r)

*       Fossil fuel production activity (crude, gas and coal):
$prod:y(xe,r)$vom(xe,r)  s:(esub_es(xe,r))  id:0
        o:py(xe,r)      q:vom(xe,r)     a:ra(r) t:ty(xe,r)
        i:pa(j,r)       q:vafm(j,xe,r)  p:pai0(j,xe,r)  a:ra(r) t:ti(j,xe,r) id:
        i:pl(r)         q:ld0(xe,r)  id:
        i:pr(xe,r)      q:rd0(xe,r)

*       Armington aggregation over domestic versus imports:
$prod:a(i,r)$a0(i,r)  s:4  m:8 s.tl(m):0
        o:pa(i,r)       q:a0(i,r)
        i:py(i,r)       q:d0(i,r)
        i:py(i,s)       q:vxmd(i,s,r)  p:pmx0(i,s,r) s.tl:
+               a:ra(s) t:tx(i,s,r)     a:ra(r) t:(tm(i,s,r)*(1+tx(i,s,r)))
        i:pt#(s)       q:vtwr(i,s,r)  p:pmt0(i,s,r) s.tl: a:ra(r) t:tm(i,s,r)

*       International transport services (Cobb-Douglas):
$prod:yt  s:1
        o:pt           q:(sum((i,r), vst(i,r)))
        i:py(i,r)       q:vst(i,r)

*       Final demand:
$demand:ra(r)
        d:pc(r)         q:ct0(r)
        e:py("cgd",r)   q:-vom("cgd",r)
        e:rkr(r)$rsk    q:(sum(i, kd0(i,r)))
        e:rkg$gk        q:(sum(i, kd0(i,r)))
        e:pl(r)         q:evoa("lab",r)
        e:pr(xe,r)      q:rd0(xe,r)
        e:pc("usa")     q:vb(r)
        e:pcarb(r)      q:carblim(r)
```

Appendix 6. ASPEN_SMALL.SET

```
$TITLE   Set Definitions for 13 regions and 8 goods

SET   I   Sectors/
Y Other manufactures and services
EIS Energy-intensive sectors
COL Coal
OIL     Petroleum and coal products (refined)
CRU Crude oil
GAS Natural gas
ELE Electricity
CGD Savings good/;

SET R Aggregated Regions /
USA United States
CAN Canada
EUR Europe
JPN Japan
OOE Other OECD
FSU Former Soviet Union
CEA Central European Associates
CHN China (including Hong Kong + Taiwan)
IND India
BRA Brazil
ASI Other Asia
MPC Mexico + OPEC
ROW Rest of world /

Set   F  Aggregated factors /
       LAB      Labor,
       CAP      Capital /;
```

Appendix 7. ASPEN_SMALL.MAP

```
$title Map file

* Aggregating ASPEN dataset (45x23) into ASPEN_SMALL dataset (13x8)

*       -------------------------------------------------------------
*       The target dataset has fewer sectors, so we need to specify how
*       each sector in the source dataset is mapped to a sector in the
*       target dataset:

$SETGLOBAL source aspen

Set mapi Sectors and goods /
```

```
GAS.GAS Natural gas works
ELE.ELE Electricity and heat
OIL.OIL Refined oil products
COL.COL Coal transformation
CRU.CRU Crude oil

I_S.EIS Iron and steel industry (IRONSTL)
CRP.EIS Chemical industry (CHEMICAL)
NFM.EIS Non-ferrous metals (NONFERR)
NMM.EIS Non-metallic minerals (NONMET)
TRN.EIS Transport equipment (TRANSEQ)
PPP.EIS Paper-pulp-print (PAPERPRO)

T_T.Y   Trade margins
AGR.Y   Agricultural products
OME.Y   Other machinery (MACHINE)
OMN.Y   Mining (MINING)
FPR.Y   Food products (FOODPRO)
LUM.Y   Wood and wood-products (WOODPRO)
CNS.Y   Construction (CONSTRUC)
TWL.Y   Textiles-wearing apparel-leather (TEXTILES)
OMF.Y   Other manufacturing (INONSPEC)
SER.Y   Commercial and public services
DWE.Y   Dwellings,

CGD.CGD Investment composite /;

SET MAPR  mapping GTAP regions /
 AUS.OOE    Australia
 NZL.OOE    New Zealand
 JPN.JPN    Japan
 KOR.ASI    Republic of Korea
 IDN.MPC    Indonesia
 MYS.ASI    Malaysia
 PHL.ASI    Philippines
 SGP.ASI    Singapore
 THA.ASI    Thailand
 VNM.ASI    Vietnam
 CHN.CHN    China
 HKG.CHN    Hong Kong
 TWN.CHN    Taiwan
 IND.IND    India
 LKA.ASI    Sri Lanka
 RAS.ASI    Rest of South Asia
 CAN.CAN    Canada
 USA.USA    United States of America
 MEX.MPC    Mexico
 CAM.ROW    Central America and Caribbean
 VEN.ROW    Venezuela
 COL.ROW    Columbia
 RAP.ROW    Rest of Andean Pact
 ARG.ROW    Argentina
```

```
 BRA.BRA    Brazil
 CHL.ROW    Chile
 URY.ROW    Uruguay
 RSM.ROW    Rest of South America
 GBR.EUR    United Kingdom
 DEU.EUR    Germany
 DNK.EUR    Denmark
 SWE.EUR    Sweden
 FIN.EUR    Finland
 REU.EUR    Rest of EU,
 EFT.EUR    European Free Trade Area
 CEA.CEA    Central European Associates
 FSU.FSU    Former Soviet Union
 TUR.ROW    Turkey
 RME.MPC    Rest of Middle East
 MAR.ROW    Morocco
 RNF.MPC    Rest of North Africa
 SAF.ROW    South Africa
 RSA.ROW    Rest of South Africa
 RSS.ROW    Rest of South-Saharan Africa
 ROW.ROW    Rest of World  /;

*       The following statements illustrate how to aggregate
*       factors of production in the model.  Unlike the aggregation
*       of sectors or regions, you need to declare the set of
*       primary in the source as set FF, then you can specify the
*       mapping from the source to the target sets.


set ff /LND,SKL,LAB,CAP,RES/;
SET MAPF mapping of primary factors /LND.CAP,SKL.LAB,LAB.LAB,CAP.CAP,RES.CAP/;
```