cgemod

Deriving a Global Social Accounting Matrix and Satellite Data from GTAP Databases

CGEMOD Technical Paper 12

Scott McDonald[1] and Karen Thierfelder

Abstract

This paper describes how the GTAP databases can be converted into global Social Accounting Matrices (SAM) and linked satellite accounts. The focus of attention is on the generation of a SAM representation of the GTAP database that is consistent with the GTAP model. The resultant SAM can be used to calibrate global CGE models that have been coded in GAMS or GEMPACK. Since GTAP version 7, the GTAPAgg programme has produced the transaction values in the GTAP database as a SAM using the method described in this paper. As of mid-2019, the GAMS codes used to generate the SAM representation of the GTAP database are available for GTAP versions 5 to 10 inclusive.

Keywords: Social Accounting Matrix, Trade Analysis, Computable General Equilibrium.

Correspondence Address:

Dr Scott McDonald, International Agricultural Trade and Development Group, Humboldt-University of Berlin, Berlin, Germany.

e: jrs.mcdonald@gmail.co.uk

---

1 - November 8, 2019

Table of Contents

1.   Introduction

The GTAP database is the major source of data used by global Computable General Equilibrium (CGE) models. The database is supplied in the Header Array (HAR) file format, which is used by GEMPACK, together with two programmes, GTAPAgg and FlexAgg, that can be used to generate aggregations of the database. However, while the HAR file format is appropriate for GEMPACK users, it is not immediately usable by different modeling software, e.g., GAMS for which the GDX file format is preferred, and although some modelers may wish to use the GTAP data, they may not wish to use the GTAP model and/or GEMPACK.

Since the databases for ALL whole economy models can always be represented in a Social Accounting Matrix (SAM) format, and many modelers choose to present their model databases in SAM format, a SAM format has attractions. The main reasons for developing a Social Accounting Matrix (SAM) version of the GTAP database are

1. the greater accessibility of the information in the database to modelers and policy makers;

2. the explicit identification of the price system and processes that generate the price definitions;

3. the increased use by economists of the SAM format, especially now that it is formally part of the System of National Accounts (SNA) (UN, 1993 and 2008); and

4. the greater ease with which the data for a single region can be assessed and related to national account aggregates.

The transformation of the GTAP database to a SAM format does involve some complications. These arise because the GTAP database does not appear, at first sight, to observe the 'law of one price', i.e., 'buyer' prices are not common across all agents, and the fact that GTAP does not adopt the price definitions set out in the SNA.[2] Thus when developing a SAM representation, it is necessary to identify the price system used in the GTAP database.

SAMs that are consistent with the SNA use three key price definitions - purchaser, producer and basic prices – while the GTAP system uses three different price definitions -

---

[2]   The issue of prices is explored further below.

agent, market and world prices. In the GTAP literature for an imported commodity, the values of transactions at 'market' prices are defined as the values at 'world' prices plus trade and transport margins plus import duties, and the values at 'agents' prices are then defined as the values at 'market' prices plus any sales/purchase taxes. For domestic commodities, the values of transactions at 'agents' prices for domestic consumption are the values at 'market' prices plus any sales/purchase taxes, while for exports the values at 'world' prices are the values at 'market' prices less any export taxes. There is no simple relationship between the SNA and GTAP price systems.

The guiding principle followed in converting the GTAP database to a SAM is to ensure that the 'law of one price' holds in the resulting SAM, i.e., that all transactions in each row are valued at a common price. Thus, for example, purchasers of commodities for consumption/demand by domestic agents, make separate payments to the supplying agents (domestic or foreign) and the government with respect to indirect taxes on commodities. This approach ensures that the national accounting identities for the SAM are maintained, while transforming the GTAP database into a global SAM format that is FULLY consistent with the price and tax systems in the GTAP model.

A SAM representation has the further advantage of making it easier to augment the transactions data and relate the information in satellite accounts to specific transactions. Examples of data augmentation include separately identifying various inter regional current account transactions, e.g., migrant remittances, aid transfers, etc., (see McDonald and Sonmez, 2004; Walmsley *et al*., 2007; McDonald *et al*., 2016) and decomposing the transfers between domestic institutions, e.g., savings and income taxes paid by households and government borrowing/savings (see McDonald *et al*., 2004; Walmsley & Minor, 2013).

The programmes used to produce the SAM representation were developed in GAMS and are available, in the form of a GAMS User Library, for free.[3] The programmes include checks to ensure data consistency, but users are encouraged to carry out checks for

---

[3] The GAMS programmes can all be run using demo/student GAMS that is freely available from the GAMS website (www.gams.com). Early versions of the code used HAR2GAMS and GAMS2HAR, to convert between HAR and GDX formats; these programmes were written by Mark Horridge and Tom Rutherford and versions were subsequently made standard parts of the GAMS language. Earlier code for use with version 5.0 of the GTAP database used SEEHAR, written by Ken Pearson for converting the HAR file into a GAMS readable input file.

themselves.[4] The SAM representation detailed in this paper has been implemented in GTAPAgg since version 7: the codes used to produce the SAM by GTAPAgg were developed from the GAMS code reported here by Terrie Walmsley and incorporated into the GTAPAgg programme by Mark Horridge. The GAMS programmes produce the SAMs in GDX and HAR format while GTAPAgg produces the SAM in HAR format.[5] The GAMS programmes also convert all the other data in the standard GTAP database into satellite accounts that are saved in the GDX format.

The rest of this paper is organised as follows. The next section provides a description of the Social Accounting Matrix. The pertinent aspects of the data provided by the GTAP database are discussed in section 3, while section 4 provides a description of how the matrices of data from the GTAP database are used to construct the SAM. There is also an appendix that contains an extract from the GAMS programme that relates to the construction of the SAM. The SAM produced by the programme does not balance exactly; since this is a necessary condition for a 'true' SAM section 5 reports the magnitudes and relative importance of differences in the account totals for both SAMs, which are very minor, and then reports the steps taken to ensure balanced SAMs are produced. The maximum and minimum transactions values for each version of the database by region are reported in section 6, while section 7 offers some concluding comments.

---

[4]   If errors are uncovered please let the authors know; preferably with a copy of the code used to generate and detect the errors.

[5]   The GAMS programmes for GTAP 9 and 10 have been revised to use the SAM produced by GTAPAgg.

2.   Price Definitions: GTAP & System of National Accounts

An important feature of all SAMs is the explicit identification of the system of prices. An understanding of the system of prices in a SAM is important if the data are to be used appropriately in any economic model, and is critical to the development of the price system in any CGE model. The price system used in the GTAP database is not consistent with the price system used in the System of National Accounts (SNA). Although the price systems have a degree of commonality, it is important to appreciate the differences between the two systems when interpreting the GTAP data and relating the reported values to (most) published national accounts. It is convenient to start with a brief description of the system of prices in the SNA and then review the system used in the GTAP database; this order allows easy explanation of the differences.

SNA Price System

This section explains the price system in the System of National Accounts and its implications for the interpretation of a SAM consistent with the price system in the SNA. As part of this explanation it is important to understand the distinction between prices in input-out tables (IOT) and Supply and Use tables (SUT).

Prices in the SNA

The SNA defines three key prices (see SNA, 2008, 6.49 to 6.69):

1. Purchaser Prices - the prices paid by purchasers, which include transport and distribution margins and any VAT payable.

2. Producer Prices - the price paid to the producer by the purchaser less any VAT or other deductible tax and any transport or distribution charges invoiced separately.

3. Basic Prices - the price paid to the producer by the purchaser less any tax payable plus any subsidy receivable but excluding any transport or distribution charges invoiced separately.

These relationships between these prices are summarised in Figure 1.

Figure 1   **SNA Price Relationships**

<div align="center">

Basic prices

+

Taxes on products excluding invoiced VAT

-

Subsidies on products

=

Producers' prices

+

VAT not deductible by the purchaser

+

Separately invoiced transport charges

+

Wholesalers' and retailers' margins

=

Purchasers' prices

</div>

Source: SNA (2008), Figure 6.1, p 103.

The key prices for a CGE model are the basic and the purchaser prices. Typically, in a SAM consistent with the SNA, commodities supplied to an economy by domestic activities and the rest of the world will be valued in basic prices, while commodities used by an economy will be valued in purchaser prices. It is useful to see how these prices are recorded in a SAM, but to do so requires determining the prices paid by agents when purchasing commodities.

If the SAM is constructed around Supply and Use tables (SUT), i.e., the inter-industry Use matrices are asymmetric[6], then prices for each commodity will be **purchaser** prices and identical for all agents. Domestic production and imports, *cif* paid plus import duties/subsidies, are valued at basic prices.

If the SAM is constructed around input-output tables (IOT), i.e., the inter-industry Use matrices are symmetric, and the SAM follows SNA principles, then, in one possible arrangement, prices in the rows for each commodity will be **basic** prices and identical for all agents. The derivation of purchaser prices for each agent then requires the existence of two sets of additional matrices in which each row corresponds to a commodity and each column to an agent that purchases commodities;

---

[6]   Commodities and activities are separately identified, even if a principle product classification scheme is used for activities, and the Supply matrix is/may be non-square and/or will have entries off the diagonal.

1. the columns of the first set of matrices should contain all the agent specific taxes and subsidies paid by the agent for purchases of each commodity;[7] and

2. the columns of the second set of matrices should contain all the agent specific transport charges and wholesaler and retailer margins paid by the agent for purchases of each commodity.[8]

This IOT structure means that **purchaser** prices are agent and commodity specific, but this requires that adequate data exist to define agent specific taxes and subsidies paid and agent specific transport charges and wholesaler and retailer margins for purchases of each commodity.

GTAP Price System

The GTAP price system has (implicit) definitions for three prices: Agents prices, Market prices and World prices. These definitions can be defined as

1. Agents prices – the agent specific prices paid by agents for commodities;

2. Market prices – the prices of imported commodities delivered to a region inclusive of international trade and transport costs and any import duties/subsidies, **and** the prices received by agents when selling commodities on the domestic market or as exports; and

3. World prices – the prices paid for exports after any export taxes/subsidies, and the prices received for imports after the payment of international trade and transport costs but before the payment of import duties/subsidies.

These relationships between these prices are summarised in Figure 2.

---

[7] Each matrix in the set would relate to a specific tax instrument, e.g., General Sales Taxes (TST), VAT, excise taxes, etc.

[8] Each matrix in the set would relate to a specific trade and transport margin, e.g. transport charges and wholesaler margins, etc.

Figure 2   GTAP Price Relationships

<table>
<tr><td align="center">**Traded G&S**</td><td align="center">**Non-Traded G&S**</td></tr>
<tr><td align="center">Market prices (outputs)</td><td align="center">Market prices (outputs)</td></tr>
<tr><td align="center">-</td><td></td></tr>
<tr><td align="center">Export taxes</td><td></td></tr>
<tr><td align="center">=</td><td></td></tr>
<tr><td align="center">World prices (exports)</td><td></td></tr>
<tr><td align="center">+</td><td></td></tr>
<tr><td align="center">International trade margins</td><td></td></tr>
<tr><td align="center">=</td><td></td></tr>
<tr><td align="center">World prices (imports)</td><td></td></tr>
<tr><td align="center">+</td><td></td></tr>
<tr><td align="center">Import duties/subsidies</td><td></td></tr>
<tr><td align="center">=</td><td></td></tr>
<tr><td align="center">Market prices (imports)</td><td></td></tr>
<tr><td align="center">+</td><td align="center">+</td></tr>
<tr><td align="center">Taxes/subsidies on products</td><td align="center">Taxes/subsidies on products</td></tr>
<tr><td align="center">=</td><td align="center">=</td></tr>
<tr><td align="center">Agents prices imported G&S</td><td align="center">Agents prices domestic G&S</td></tr>
</table>

The documentation for the submission of IO data for regions to GTAP requests the submission of two matrices: matrices UP are valued "commodity taxes inclusive" while the matrices UF are valued "commodity taxes excluded" (see Huff et al., (2000), Section 3 and Table 1). But trade and transport costs are not mentioned. It is tempting to conclude that Agents prices for commodities (matrix UP) are equivalent to purchaser prices in the SNA, but that means that the matrix UF must be valued at basic prices plus trade and transport margins. But the imports matrix MF, presumably *cif* paid, is valued "import duties excluded", which implies that the market prices of imports are equivalent to basic prices.

It does not appear possible to reconcile the SNA and GTAP price systems, and no attempt is made to do in this paper. Rather the SAM, from the GTAP database, is constructed solely in accordance with the GTAP price system.

3.   Social Accounting Matrix

The structure of the Global SAM is illustrated by means of a representative SAM for one region – the structure of the SAM for each and every region is identical. The description of the SAM structure is limited, since it is assumed that the reader is familiar with the basic principles of a SAM (see King, 1985; Sadoulet and de Janvry, 1995; Pyatt and Round, 1977;

Pyatt, 1991; and Reinert and Roland-Holst, 1997); those interested in practical introduction to SAMs can follow the (free) online course at www.cgemod.org.uk/int_sam.html . In general terms the SAM structure adopted follows the conventions of the System of National Accounts for 1993 (UN, 1993), with adjustments for the limited data on intra-institutional accounts and GTAP's price system. The basic structure is identified in Table 1, wherein the transactions captured by each of the active sub matrices of the SAM are defined/described. The SAM layout has an implicit number code for each sub-matrix, given by the row and column numbers; this system of code numbers is used throughout the text to identify the sub-matrix currently under consideration.

SAM Transactions

The SAM reports six groups of agents that demand commodities (reading across the rows of the commodity accounts); activities, private households, government, investment, global transport services and other regions. These transactions take place at GTAP market prices. However, associated with each purchase by an agent there is an additional payment to the government that represents sales taxes. The SAM demonstrates that for imported commodities market prices are the price received by the exporting region plus the per unit transport cost plus the per unit tariff rate (reading down the columns of the commodity accounts). For domestically produced commodities market prices are the prices received by domestic activities, and hence export taxes are recorded as expenditures by the domestic commodity accounts. Domestic market prices are derived from the production costs, which are made up of the costs of intermediate inputs valued at market prices plus the sales taxes paid on intermediate inputs, plus the expenditure on factor use and production taxes and payments to primary inputs.

Table 1 **Transactions in the Social Accounting Matrix for a Representative Region**

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
		Imported Commod	Domestic Commod	Activities	Factors	Regional Household	Private Household	Trade Taxes	Import Sales Taxes	Domestic Sales Tax	Factor Taxes	Production Taxes	Direct Taxes	Govt	Capital	Import Margins	Export Margins	Rest of World
1	Imported Commodities			Intermediate Inputs			Private demand							Government demand	Investment demand			
2	Domestic Commodities			Intermediate inputs			Private demand							Government demand	Investment demand		Exports of Transport Services	Exports of Goods and Services
3	Activities		Supply matrix															
4	Factors			Payments to Factors														
5	Regional Household				Distributed Factor Incomes			Import Duties and Export duties	Import Sales Taxes	Domestic Sales Taxes	Factor Use Taxes	Production Taxes	Income Taxes					
6	Private Household					Private Household income												
7	Trade Taxes	Import Duties	Export duties															
8	Import Sales Taxes			Sales Taxes - Imports			Sales Taxes - Imports							Sales Taxes - Imports	Sales Taxes - Imports			
9	Domestic Sales Tax			Sales Taxes - Domestic			Sales Taxes - Domestic							Sales Taxes - Domestic	Sales Taxes - Domestic			
10	Factor Taxes			Factor Use Taxes														
11	Production Taxes			Production Taxes														
12	Direct Taxes					Income Taxes												
13	Government					Government income												
14	Capital				Depreciation	Savings											Trade balance	Trade balance
15	Import Margins	Transport margins on imports																
16	Export Margins														Imports of transport services			
17	Rest of World	Imports of Goods and Services																

All income initially accrues to the regional household before being distributed to the private household, government and the capital account. Regional income is made up of payments to factors, net of factor income taxes[9] and depreciation, plus indirect taxes, which are made up of trade, sales, factor use and production taxes, and income taxes. The artifice of a regional household means that the private household does not pay income taxes nor does it save, and the government neither saves nor borrows, i.e., the implicit balance on the government budget is zero.

Trade consists of two elements; expenditures on commodities and expenditures on transport margins. Exports are valued *fob* (free on board) and after the payment of any export duties. Exports of transport services to the global transport pool are recorded separately. Imports of commodities are also valued *fob*, with transport services recorded separately. The sum of the two represents expenditure on imports inclusive of carriage, insurance and freight (*cif*). Consequently, there are two sets of trade balances: the first, represents the trade balance with respect to each of the transport services while the second are the trade balances with all other regions.

SAM Dimensions

The dimensions of the SAM are determined by accounts that are identified. Since these change over time it is useful to have a generalised representation of the dimensions of the global SAM. Specifically the dimensions of the SAM can be defined in general terms by using sets for the number of commodities, c; activities, a; factors, f; and the number of regions, k. The numbers of members in each set for several versions of the GTAP database are defined in Table 2.

In the GTAP database there is no distinction between commodities and activities; rather GTAP identifies sectors, where the term sector is used interchangeably to define imported and domestic produced commodities and production activities. This is possible for two reasons:

      i)   underlying the data for each region is a symmetric input-output table that means the Supply Matrix – sub matrix 3:2 in Table 1 – is strictly diagonal; and

---

9   Since version 7 of the GTAP database, factor income taxes are defined as taxes paid directly by factors, e.g., social security payments, plus income taxes paid by households and incorporated business

ii)   the GTAP database is not organised in a single matrix structure.

In single matrix (SAM) structure it is necessary to define 2*c commodity accounts (c for imported and c for domestically produced commodities) and a activity accounts. The database allows each domestic agent to be charged different purchase, or sales, tax rates[10] for both imported and domestic commodities; hence it is necessary to have 2*c accounts for sales taxes. Similarly, each production activity can be charged an activity specific indirect/production tax, which necessitates another account. Production also involves the use of primary inputs, of which f are identified in GTAP; so, there are f accounts for factor purchases and f accounts for taxes paid on factor use by activities.

Table 2   SAM Dimensions

Description	Code	Version 5.4	Version 6.0	Version 7.1	Version 8.1	Version 9	Verion 10
Commodities	c	57	57	57	57	57	65
Activities	a	57	57	57	57	57	65
Factors	f	5	5	5	5	8	8
Regions	k	78	87	112	134	140	141
Margin Commodities	m	3	3	3	3	3	3
Domestic Institutions	none	4	4	4	4	4	4

For trade relationship, each region **can** import from and export to all other regions, hence for each region there needs to be k accounts to record trade transactions. In addition, there needs to be 2*k accounts for each region to capture trade taxes – import duties and export taxes – upto and including version 7.1 - and 3*k accounts for each region since version 8 to capture trade taxes – *ad valorem* and specific (quantity) import duties and export taxes.[11] And since there are three transport margins associated with each trade transaction there needs to be m*k accounts for each region to capture trade and transport costs; with a further m accounts to capture exports by each region to the global trade and transport pool. Finally, there are four domestic institutional accounts: the regional household that collects all regional income, either from factor sales or taxes (indirect and direct). This income is dispersed to the

---

enterprises. Household income taxes are assigned to labour and corporation taxes are assigned to capital (see McDougall and Hagemjer, 2016).

[10]   The database does not distinguish between different types of tax instruments.

[11]   The SAM generated by the GTAPAgg programme does not separate out *ad valorem* and specific (quantity) import duties.

other three domestic institutions; the private household, government and capital (investment/savings). Table 3 provides a plan of the dimensions of the resulting SAM matrices.

The most immediately obvious points about the SAM are the large number of accounts and the relative scarcity of entries in the SAM. The large size of the SAM is primarily a consequence of the number of regions, k, in the database (k trade accounts plus $2k$ for trade taxes) and the fact that m types of transport margins on imports are identified with imports from each region ($m*k$ accounts), while the second main reason is the series of .commodity/user specific sales taxes on imports (c) and domestic commodities (c)

The resultant SAM is of the dimensions $(4c + a + 2f + 7 + mk + 3/4k)* (4c + a + 2f + 7 + mk + 3/4k)*(k)$, where c and a are the number of sectors in the database, f the number of factors and k the number of regions. The dimensions of the SAM matrices for different versions of the GTAP database are

version 5.4 – 770*770*78 (46,246,200 cells),

version 6.0 – 824*824*87 (59,070,912 cells),

version 7.1 – 974*974*112 (106,251,712 cells),

version 8.1 – 1,240*1,240*134 (206,038,400 cells),

version 9.1 – 1,288*1,288*140 (232,252,160 cells).

Version 10 - 1,335*1,335*141 (251,293,725).

Table 3 also highlights several important features of the GTAP database.

- Information is concentrated in the trade accounts.
- The SAM matrices are sparse, e.g., there are 4,580,307 non-zero entries in the GTAP SAM for version 9.1.
- The within regional information emphasises inter-industry and final demand transactions.
- The tax information relates overwhelmingly to indirect taxes on sales and purchases.

- There is very little information about domestic institutions other than for consumption, e.g., savings are not separately identified for private households and governments.

.

Table 3 — Dimensions of the Social Accounting Matrix for a Representative Region

		1 Imported Commod	2 Domestic Commod	3 Activities	4 Factors	5 Regional Household	6 Private Household	7 Trade Taxes	8 Import Sales Taxes	9 Domestic Sales Tax	10 Factor Taxes	11 Production Taxes	12 Direct Taxes	13 Govt	14 Capital	15 Import Margins	16 Export Margins	17 Rest of World	18 Totals
1	Imported Commodities			$c*a$			$c*1$							$c*1$	$c*1$				$c*1$
2	Domestic Commodities			$c*a$			$c*1$	$c*c$						$c*1$	$c*1$		$c*m$	$c*k$	$c*1$
3	Activities		$a*c$																$c*1$
4	Factors			$f*a$															$f*1$
5	Regional Household				$1*f$			$1*c$	$1*c$	$1*c$	$1*f$	$1*1$	$1*1$						$1*1$
6	Private Household					$1*1$													$1*1$
7	Trade Taxes	$1/2k*c$	$k*c$																$2k*1$
8	Import Sales Taxes			$c*a$			$c*1$							$c*1$	$c*1$				$c*1$
9	Domestic Sales Tax			$c*a$			$c*1$							$c*1$	$c*1$				$c*1$
10	Factor Taxes			$f*a$															$f*1$
11	Production Taxes			$1*a$															$1*1$
12	Direct Taxes				$1*f$														$1*1$
13	Government					$1*1$													$1*1$
14	Capital				$1*f$	$1*1$											$1*m$	$1*k$	$1*1$
15	Import Margins	$m*k*c$																	$m*k*1$
16	Export Margins															$m*m*k$			$m*1$
17	Rest of World	$k*c$																	$k*1$
18	Totals	$1*c$	$1*c$	$1*a$	$1*f$	$1*1$	$1*1$	$1*2k$	$1*c$	$1*c$	$1*f$	$1*1$	$1*1$	$1*1$	$1*1$	$1*m*k$	$1*m$	$1*k$	

3   The GTAP Data

It is useful to start with some brief notes about the organization of the GTAP databases and some aspects of GEMPACK, since recourse will be made to extracts of GEMPACK code. These notes fall into 5 categories; modeling in terms of rates of change, sets and the local control of sets, the regional household, trade and transport margins and the global bank. As a matter of convention, in the GTAP model all variables, i.e., the rates of change, are in lower case and the coefficients/parameters are in upper case. This raises an important terminological difference: GEMPACK uses the term coefficients to describe elements that are known as parameters in GAMS – in this paper the term parameter is adopted although in this context the terms are interchangeable. Subsequently, this section goes on to identify the data from the GTAP database that are used to define transactions.

GTAP and GEMPACK

The solution algorithms used by GEMPACK solve for rates of change; this requires that all cells in any sub matrix of the SAM, for which the model has behavioural relationships, must have non-zero values. In terms of single precision this requires that cells for which there are no real-world transactions report values of 0.000001 (ONE US Dollar) or greater; it is apparently not possible to identify which of these small non-zero transactions represent real world transactions.

   This has implications that should be noted:

1. The GTAP database, and SAM, contains non-zero zeros, i.e., the number of real-world transactions are less than the number of active cells in the database;

2. the row and column totals for a GTAP SAM are not identical, which limits the extent to which the accounting identities can be used to verify the data are consistent[12]; and

3. any elimination (or truncating) of very small transactions, say less than $(US)100, will result in differences in the row and column totals of the SAM.

---

[12]   The equality of row and column totals in a SAM can be used to identify consistency, the reported transactions are consistent, but NOT that they are complete.

One feature of GEMPACK that might cause discomfiture to GAMS modelers is the solution of models in terms of rates of changes – in essence the models are written in terms of log differentials. One area in which this can cause some confusion is the treatment of taxes. Typically, 'levels modelers' might be inclined to talk about the *tax rates*, whereas 'rates of changes modelers' might talk about the *powers of taxes*. The explanation is straightforward. Consider the level relationship for an *ad valorem* tax

$$P_1 = P_2 * (1 + T)$$

where *T* is the *ad valorem* tax rate and (1 + *T*) is defined as the **power of the tax**. In terms of rates of change this can be written as

$$d \ln (P_1) = d \ln (P_2) + d \ln (1 + T)$$
$$p_1 = p_2 + t$$

such that *t* is the proportionate change in the power of the tax. It is useful to note this when reading the GTAP code for the tax system.

SETS and Control of SETS

The GTAP model – gtap62.tab (GTAP, 2002; Hertel, 1997) - defines the following sets that are relevant when transforming GTAP data into a SAM format.

```
REG # regions in the model #
TRAD_COMM # traded commodities #
MARG_COMM # margin commodities #
NMRG_COMM # non-margin commodities # = TRAD_COMM - MARG_COMM;';
CGDS_COMM # capital goods commodities #
ENDW_COMM # endowment commodities #
PROD_COMM # produced commodities # = TRAD_COMM union CGDS_COMM;
DEMD_COMM # demanded commodities # = ENDW_COMM union TRAD_COMM;
NSAV_COMM # non-savings commodities # = DEMD_COMM union CGDS_COMM;
ENDWS_COMM
# sluggish endowment commodities # = (all,i,ENDW_COMM: SLUG(i)>0);
ENDWM_COMM
# mobile endowment commodities # = ENDW_COMM - ENDWS_COMM;
ENDWC_COMM # capital endowment commodity # (capital);
```

The GTAP model also contains a useful illustration of the relationships between the sets and subsets, an adaptation of which is presented below.

NSAV_COMM				
DEMD_COMM				
	PROD_COMM			
ENDW_COMM	TRAD_COMM			CGDS_COMM
	MARG_COMM	NMARG_COMM		
ENDWM_COMM	ENDWS_COMM			

It is important, for GAMS users, to note that when reading the extracts of the GTAP model code reproduced below, that the model is coded in GEMPACK where the sets are local – hence the inclusion of local definitions for variables and parameters. To understand the local control of sets consider the following extract of code from the GTAP tablo[13] file

```
Coefficient (all,i,DEMD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFA(i,j,r) # producer expenditure on i by j in r valued at agents' prices #;
Formula (all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFA(i,j,r) = EVFA(i,j,r);
Formula (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,s,REG)
    VFA(i,j,s) = VDFA(i,j,s) + VIFA(i,j,s);
```

This first defines the coefficient (parameter in GAMS) VFA over *i*, *j* and *r* with members of *i* defined as DEMD_COM, of *j* as PROD_COM and of *r* as REG. The subsequent two formulae are then constrained to operate over subsets of *i*, i.e., ENDW_COM and TRAD_COM (see the schematic above), but they can reuse the character *i* because of the system of local control of sets. Similarly, it is perfectly acceptable to redefine *i* for subsequent use.

A major consequence of this feature of GEMPACK is the necessity to explicitly identify the set definitions that apply to each block of GEMPACK code[14].

Regional Household

The GTAP database and model use the concept of a Regional Household. Since this is crucial to an understanding of the relationships between agents it is useful to provide a brief description of the Regional Household concept. In the GTAP database all incomes from taxes (direct and indirect), all factor incomes, depreciation and the trade balance accrue to the Regional Household. These incomes are then distributed between the Private Household[15], the Government and the Capital (saving-investment) accounts, which are the only domestic

---

[13]   In GEMPACK a tablo file is the name given to a programme file – the programme for editing these files is known as TABMATE.

[14]   This may appear burdensome to some GAMS users, but it does avoid the potential problems with set ordering in GAMS. There are features of GEMPACK that are arguably simpler than the equivalent in GAMS and vice versa.

[15]   Implicitly the Private household account encompasses households, incorporated business enterprises and NPISH.

institutions responsible for the final demand of commodities. Hence. the Regional Household is the sole source of income to the private household, the government and the capital accounts. Consequently, the private household does not record transactions directly with either the government or the capital accounts and the government account does not record transactions with the capital or private household accounts; hence in the database the private household neither saves nor pays income taxes and the government does not run a surplus of deficit on its current account.

The Regional Household provides a method by which the limitations imposed by the absence of data on intra-institutional transactions can be partially circumvented.[16] This method does however have a cost in modeling terms; it involves the imposition of a specific distribution of income between domestic institutions.

Trade and Transport Margins

The GTAP database records the transactions with respect to three types of international trade and transport margin for each commodity imported by each destination region from each source region. The sum of the value of these margins for each commodity imported from each source are defined as the value of the difference between the *fob* valuation of the commodity exported from source region *s* to destination region *y* and the *cif* valuation of each commodity imported by destination region *d* from source region *s*. But, the database does not record data on the regions that supply each trade and transport margins transaction, i.e., there is not full information on bilateral trade in trade and transport margins. Instead for the export of trade and transport margins by each region the database only records the total value, whereas for margin imports the database records, separately, the value of the trade and transport cost associated with every import transaction valued at *fob* prices. The global sum of the values of the exports of each trade and transport margin must equal the global sum of the values of the respective imports of trade and transport margins. But it is not necessary that the values of the exports of trade and transport margins by each region must equal the values of the imports of trade and transport margins by each region.

In the GTAP model this is handled by the concept of a global pool for trade and transport margins. This pool supplies all the demands for (the import of) trade and transport

margins and then balances the global market for trade and transport margins by purchasing the margins from the regions that export trade and transport margins. Hence, each region runs a trade account with the global pool for trade and transport margins; *ex post* the sum of the balances on these trade accounts must be zero. Since the total values of margin exports and imports for each region and each margin goods are known, the balance in trade and transport margins trade can be readily calculated.

The Global Bank and Trade Balance

A related issue exists with respect to global savings. The values of total savings and investment within each region must be equal for all regions. But total savings by domestic institutions within each region need not equal the total value of investment within each region because of inter-regional capital account transactions. Furthermore, the *ex post* the balance on the capital accounts must be equal and opposite to the balances on the current account for each region and the global sum of the balances of trade in goods and services – current account - must equal zero, as must the global sum of inter-regional capital account transactions.

This presents economic logic problems for experiments since the logic of a price driven model suggests that differences in investment expenditures across regions would be influenced by differences in returns to capital in different regions. Nevertheless, this is not a substantive issue for the reporting of transactions in the database since it simply records transactions in the base period and by definition these must satisfy the requisite accounting conditions. Hence, for database purposes it is sufficient to record the calculated balance of trade in goods and services for each region with every other region, although issues may remain about the appropriate modeling approach.

Data Read in by GTAP Model

The GTAP model only uses a subset of the GTAP database to define the transactions within and between regions in the base period. Consequently, it is only necessary to use that subset in the derivation of the global SAM. Details of the data read in by the GTAP model are reported, in terms used in the GTAP model, below.

---

[16]   McDonald and Sonmez (2004) demonstrate that it is relatively straightforward to compile adequate additional data to at least partially address the absence of data on intra-institutional transactions. McDonald *et al*., (2016) extend this process and identify sources of data.

GTAP has a generally consistent convention for labeling data parameters/coefficients, which, while not applying to all data matrices, is extremely useful in reading the code.

V	Value;
D	Domestic;
I	Imported;
P	Private household;
G	Government;
F	Firms;
A	Agents prices;
M	Market prices;
E	Endowment;
X	Exports;
W	World.

Hence the data matrix VIGA is the **V**alue of **I**mports for **G**overnment consumption at **A**gents prices.

The GTAP data used to construct the SAM are:

Variables and base flows associated with saving

```
Coefficient (all,r,REG)
SAVE(r) # expenditure on NET savings in region r valued at agent's prices
```

Variables and base flows associated with government consumption

```
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VDGA(i,r) # government consumption expenditure on domestic i in r #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VDGM(i,r) # government consumption expenditure on domestic i in r #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VIGA(i,r) # government consumption expenditure on imported i #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VIGM(i,r) # gov't consumption expenditure on i in r #;
```

Variables and base flows associated with private consumption

```
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VDPA(i,r) # private consumption expenditure on domestic i in r #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VDPM(i,r) # private consumption expenditure on domestic i in r #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VIPA(i,r) # private consumption expenditure on imported i in r#;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VIPM(i,r) # private consumption expenditure on i in r #;
```

Variables and base flows associated with the firms

```
Coefficient (ge 0)(all,i,ENDW_COMM)(all,r,REG)
```

```
    EVOA(i,r) # value of commodity i output in region r #;
Coefficient (ge 0)(all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    EVFA(i,j,r) # producer expenditure on i by j in r valued at agent's prices#;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VDFA(i,j,r) # purchases of domestic i for use by j in region r #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VIFA(i,j,r) # purchases of imported i for use by j in region r #;
Coefficient (ge 0)(all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFM(i,j,r) # producer expenditure on i by j in r valued at mkt prices #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VIFM(i,j,r) # purchases of imports i for use by j in region r #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VDFM(i,j,r) # purchases of domestic i for use by j in region r #;
```

Variables and base flows associated with the global bank

```
Coefficient (ge 0)(all, r, REG)
    VDEP(r) # value of capital depeciation in r (exogenous) #;
```

Variables and base flows associated with the international trade and transport

```
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VIMS(i,r,s) # imports of i from r to s valued at domestic mkt prices #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VIWS(i,r,s) # imports of i from r to s valued cif (tradeables only) #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VXWD(i,r,s) # exports of i from r to s valued fob (tradeables only) #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VXMD(i,r,s) # exports of i from r to s valued at mkt prices (tradeables only) #;
Coefficient (ge 0)(all,m,MARG_COMM)(all,r,REG)
    VST(m,r) # exports of m from r for int'l trnsport valued at mkt p (tradeables
only) #;
```

The entire contents of the file basedata.har produced by the GTAPAgg programme are read into the GAMS programme using the same names for the coefficients/parameters.

4 Constructing the Social Accounting Matrix

This section describes how the various sub matrices of the SAM are assigned. Table 4 presents a map of the assignments where, in the main, the notation is derived from the GTAP data labels used the HAR file that contains the data. The major issues all relate to the treatment of taxes and hence to the systems by which prices are formed. Part of the GAMS programme that carries out these actions is listed in Appendix 2.

The description of the assignment statements starts with the components of domestic demand, and domestic production, and then turns to government income, which involves specifying the tax transactions. The discussion then moves onto the purchases of inputs by activities and to the components of regional income. Finally, the discussion turns to the treatment of trade and the associated trade and transport margins and thence the income to the capital account, i.e., the balances on the capital account with respect to the rest of the world.

GAMS Sets

The following extract from the GAMS code identifies the sets used in the programme.

```
WITHIN REGION SETS
        SS                  SAM accounts

Commodity sets
        MCC(SS)             Imported Commodities
        DCC(SS)             Domestic Commodities

Activity sets
        AA(SS)              Activities
        A2                  Production commodities Activities in GTAP

Factor sets
        FF(SS)              Factors

Tax sets
        TMMTAX(SS)          Import duties
        TEETAX(SS)          Export taxes
        SMMTAX(SS)          Sales Taxes on imported goods and services
        SDDTAX(SS)          Sales Taxes on domestic goods and services
        FFTAX(SS)           Taxes on Factors

Investment set
        II(SS)              Investment

Trade sets
        OTPPMARG(SS)        Other trade and transport costs
        WTPPMARG(SS)        Water trade and transport costs
        ATPPMARG(SS)        Air trade and transport costs

        VSTTMARG(SS)        Exports of trade and transport services

        WW(SS)              Rest of the World
```

```
REGION SET
         KK              Regions
```

The outer set SS contains all the within region accounts with a series of sub sets, e.g., MCC(SS) indicates that MCC is a subset of SS.

In addition two set that are not subsets of SS are declared. The set A2 corresponds to the set PROD_COMM in the GTAP model since it includes all commodities that are potentially traded AND the set CGDS_COM, and is only used to partition the demand for intermediate goods from the demand for investment goods. The set KK is the set of regions in the GTAP database.

A correspondence to GTAP sets for any region is provided in the table below

GAMS SAM Account Set	GAMS Subsets	GTAP Sets			
SAC	FFTAX	ENDW_COMM			
	FF				
	MCC	TRAD_COMM	PROD_COMM	DEMD_COMM	NSAV_COMM
	DCC				
	AA				
	SMMTAX				
	SDDTAX				
	II	CGDS_COMM			
	VSTTMARG	MARG_COMM			
	TMMTAX	REG			
	TEETAX				
	OTPPMARG				
	WTPPMARG				
	ATPPMARG				

where REG refers to the set of regions and the third column identifies those GTAP sets that define the membership of the GAMS subsets in the second column.

Table 4 — Mapping Transactions: Social Accounting Matrix for a Representative Country

		1 Imported Commod	2 Domestic Commod	3 Activities	4 Factors	5 Regional Household	6 Private Household	7 Trade Taxes	8 Import Sales Taxes	9 Domestic Sales Tax	10 Factor Taxes	11 Production Taxes	12 Direct Taxes	13 Govt	14 Capital	15 Import Margins	16 Export Margins	17 Rest of World	18 Totals
1	Imported Commodities			VIAM			VIPM							VIGM	VIIM				1:18
2	Domestic Commodities			VDAM			VDPM							VDGM	VDIM		VST	VXWD	2:18
3	Activities		VOM																3:18
4	Factors			VFM															4:18
5	Regional Household				EVOA-VDEP			TMTAX + TETAX	ISTAX	DSTAX	FTAX	PTAX	DTAX						5:18
6	Private Household					YH													6:18
7	Trade Taxes	VIMS-VIWS	VXWD - VXMD																TMTAX TETAX
8	Import Sales Taxes			VIAA-VIAM			VIPA-VIPM							VIGM-VIGA	VIIA-VIIM				ISTAX
9	Domestic Sales Tax			VDAA-VDAM			VDPA-VDPM							VDGM-VDGA	VDIA-VDIM				DSTAX
10	Factor Taxes			EVFA-VFM															FTAX
11	Production Taxes			PRODTAX											*PTAXINV*				PTAX
12	Direct Taxes				PTAXFACT														DTAX
13	Government					YG													13:18
14	Capital				VDEP	SAVE											VTWR-VST	VIWS-VTWR-VXWD	14:18
15	Import Margins	VTWR																	15:18
16	Export Margins														VTWR				16:18
17	Rest of World	VIWS-VTWR																	17:18
18	Totals	18:1	18:2	18:3	18:4	18:5	18:6	18:7	18:8	18:9	18:10	18:11	18:12	18:13	18:14	18:15	18:16	18:17	

A naming convention is adopted in the GAMS code for each GTAP parameter/coefficient used in the programme. All the GTAP parameters are read in using the names and sets assigned to them in the HAR file. But the local control of sets within GEMPACK means that it is possible for a set to be used with multiple definitions, which goes against the logic of the SAM account structure and the global control of sets in GAMS. Hence as part of the transition process it is necessary to re label the accounts for each parameter. Overwhelmingly this is carried out by the application of mapping sets that allow the relabelling of set elements through the use of conditional statements. For instance in the following block of GAMS code

```
SAMG(mcc,aa,kk)        = SUM((trad_comm,prod_comm)
                            $(MAPIMP(mcc,trad_comm)
                            $MAPACT(aa,prod_comm)),
                            VIFM(trad_comm,prod_comm,kk)) ;
```

the elements *trad_comm* and *prod_comm* in the parameter *VIFM* are relabeled as elements *mcc* and *aa* in the parameter *SAMG* via the mapping sets *MAPIMP* and *MAPACT*.[17]

Commodity Demand by Domestic Agents

Commodity demand by domestic agents consists of final demand by domestic institutions – private household, government and investment – and intermediate demand by activities. Domestic demand is further disaggregated by distinguishing between imported and domestically produced commodities. If standard national accounting conventions for a SAM are followed, purchases by domestic agents would be valued at purchaser prices (agents prices in the GTAP database), but, for reasons to do with the indirect tax instruments that are discussed below, these transactions needed to be recorded at sellers prices (market prices in the GTAP database).

There are two key prices for the commodity demand system. The market price of the domestic commodity i in region r, $pm_{i,r}$, and market price of the imported commodity i in region r, $pim_{i,r}$. The relevant GTAP code is given below.

```
Variable (orig_level = 1.0)(all,i,NSAV_COMM)(all,r,REG)
   pm(i,r) # market price of commodity i in region r #;
Variable (orig_level = 1.0)(all,i,TRAD_COMM)(all,r,REG)
   pim(i,r) # market price of composite import i in region r #;
```

---

[17] Some early versions of these programmes did not use this method. Versions to work with GTAP data after version 5.3 have been recoded while earlier versions have not been updated since it is assumed that only versions of GTAP after 5.3 remain relevant.

The market price is defined as the price received by the seller of the commodity. In a SAM that follows the SNA conventions for supply and use tables the commodity demands by domestic agents would be valued in purchaser prices; however that convention implicitly presume that a 'law' of one price holds. Hence for this SAM the domestic commodity demands by domestic agents are valued in seller prices, which are uniform for all purchasers. (See the notes on the tax system below for an explanation of how this affects the recording of taxes.)

Intermediate Demand and Investment (1:3, 2:3, 1:14, 2:14)[18]

Intermediate demands for imported and domestic commodities are components of purchases by firms. In market price terms these are identified in the GTAP tablo file as follows

```
Coefficient (ge 0)(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VIFM(i,j,r) # purchases of imports i for use by j in region r #;
Update (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VIFM(i,j,r) = pim(i,r) * qfm(i,j,r);
Read
    VIFM from file GTAPDATA header "VIFM";

Coefficient (ge 0)(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VDFM(i,j,r) # purchases of domestic i for use by j in region r #;
Update (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VDFM(i,j,r) = pm(i,r) * qfd(i,j,r);
Read
    VDFM from file GTAPDATA header "VDFM";
```

where

```
Variable (orig_level=VDFM)(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,s,REG)
    qfd(i,j,s)  # domestic good i demanded by industry j in region s #;
Variable (orig_level=VIFM)(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,s,REG)
    qfm(i,j,s) # demand for i by industry j in region s #;
```

Note that the set definitions for these parameters, excluding EVOA, are defined over *PROD_COMM*. This is because investment expenditure is treated as a demand by firms for capital commodities (*CGDS*). This raises a minor complication through the need to separate *CGDS* out to derive investment demand and results in the declaration of a tax parameter that is **always** assigned a zero value. Since the SAM separately records demands by activities (intermediate demand) and investment demand the GAMS code separates out the intermediate and investment demand data by the use of conditional mapping sets.

Private Household Demand (1:6, 2:6)

The demand for domestic and imported commodities by the (domestic) private households, one per region, are identified in the GTAP tablo file in market price values as

---

[18]   The number in parentheses refer to the row and column numbers used in Tables 1, 3 and 4.

```
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VDPM(i,r) # private consumption expenditure on domestic i in r #;
Update (all,i,TRAD_COMM)(all,r,REG)
    VDPM(i,r) = pm(i,r) * qpd(i,r);
Read
    VDPM from file GTAPDATA header "VDPM";

Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VIPM(i,r) # private consumption expenditure on i in r #;
Update (all,i,TRAD_COMM)(all,r,REG)
    VIPM(i,r)= pim(i,r) * qpm(i,r);
Read
    VIPM from file GTAPDATA header "VIPM";
```

where

```
Variable (orig_level=VDPM)(all,i,TRAD_COMM)(all,s,REG)
    qpd(i,s) # private hhld demand for domestic i in region s #;
Variable (orig_level=VIPM)(all,i,TRAD_COMM)(all,s,REG)
    qpm(i,s) # private hhld demand for imports of i in region s #;
```

Government Demand (1:13, 2:13)

The demand for domestic and imported commodities by the regional governments are

identified in the GTAP tablo file in market price values as

```
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VDGM(i,r) # government consumption expenditure on domestic i in r #;
Update (all,i,TRAD_COMM)(all,r,REG)
    VDGM(i,r) = pm(i,r) * qgd(i,r);
Read
    VDGM from file GTAPDATA header "VDGM";

Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)
    VIGM(i,r) # gov't consumption expenditure on i in r #;
Update (all,i,TRAD_COMM)(all,r,REG)
    VIGM(i,r) = pim(i,r) * qgm(i,r);
Read
    VIGM from file GTAPDATA header "VIGM";
```

where

```
Variable (orig_level=VDGM)(all,i,TRAD_COMM)(all,s,REG)
    qgd(i,s) # government hhld demand for domestic i in region s #;
Variable (orig_level=VIGM)(all,i,TRAD_COMM)(all,s,REG)
    qgm(i,s) # government hhld demand for imports of i in region s #;
```

Domestic Commodity Output (3:2)

Domestic output for the domestic market (*VDM*) is defined as the sales on the domestic

market of commodities produced by domestic activities. In the GTAP tablo file this is given as

```
Coefficient (all,i,TRAD_COMM)(all,r,REG)
    VDM(i,r) # domestic sales of i in r at mkt prices (tradeables only) #;
Formula (all,i,TRAD_COMM)(all,r,REG)
    VDM(i,r) = VDPM(i,r) + VDGM(i,r) + sum(j,PROD_COMM, VDFM(i,j,r));
```

Hence

$$VDM_{dcc,kk} = VDPM_{dcc,kk} + VDGM_{dcc,kk} + VDAM_{dcc,kk} + VDIM_{dcc,kk}$$

where the definitions of the right hand side terms are the market priced values of private demand (*VDPM*), government demand (*VDGM*), intermediate (activity) demand (*VDAM*) and investment demand (*VDIM*).

But the convention adopted for this SAM is that exports are from the domestic commodity accounts.[19] Hence the supply matrix is defined as the total value of domestic production destined for domestic and export (foreign) markets valued at producer prices. Hence total domestic supply (*VOM*) is defined as *VDM* plus commodity exports at market prices (*VXMD*), and transport services (*VST*), where commodity exports at market prices are equal to commodity exports at world prices (*VXWD*) minus export taxes (*ETAX*). This is defined as part of *VOM* in the GTAP tablo file, i.e.,

```
Coefficient (all,i,NSAV_COMM)(all,r,REG)
    VOM(i,r) # value of commodity i output in region r at market prices #;
Formula (all,i,ENDW_COMM)(all,r,REG)
    VOM(i,r) = sum(j,PROD_COMM, VFM(i,j,r));
Formula (all,m,MARG_COMM)(all,r,REG)
    VOM(m,r) = VDM(m,r) + sum(s,REG, VXMD(m,r,s)) + VST(m,r);
Formula (all,i,NMRG_COMM)(all,r,REG)
    VOM(i,r) = VDM(i,r) + sum(s,REG, VXMD(i,r,s));
Formula (all,h,CGDS_COMM)(all,r,REG)
    VOM(h,r) = VOA(h,r);
```

although the presentation is complicated by the definition of *VOM* to include the values of endowments and capital goods at market prices.

Hence

$$VOM_{dcc,kk} = VDM_{dcc,kk} + VXMD_{dcc,kk} + VST_{m,kk}$$
$$= VDM_{dcc,kk} + \left(VXWD_{dcc,kk} - ETAX_{dcc,kk}\right) + VST_{m,kk}$$

where *m* is a subset of *dcc* that identifies the margin commodities.

Expenditures by Activities/Incomes to Factors

Expenditures by activities fall under three headings: purchases of intermediate inputs, primary inputs and payment of indirect taxes. Purchases of intermediate inputs are recorded as components of commodity demand. These are valued at market prices with the associated 'sales' taxes being treated as separate expenditures by the activities – see the section on indirect taxes. The treatment of production taxes (*PRODTAX*) is also discussed in the section on indirect taxes.

---

[19]   This contrasts with the SAM layout implicit to the GTAP model, which follows the practice of treating exports as originating from the activity accounts (for an early example of this layout see Dervis *et al*.,

This leaves the payments by activities for factor services (4:3). The GTAP tablo file indicates that the payments to factor accounts for factor services are defined as

```
Coefficient (all,i,NSAV_COMM)(all,r,REG)
    VOM(i,r) # value of commodity i output in region r at market prices #;
Formula (all,i,ENDW_COMM)(all,r,REG)
    VOM(i,r) = sum(j,PROD_COMM, VFM(i,j,r));
```

where[20]

```
Variable (all,i,ENDWM_COMM)(all,j,PROD_COMM)(all,r,REG)
    pm(i,j,r) # market price of mobile endowment i used by j, in r #;
```

and

```
Variable (all,i,ENDWS_COMM)(all,j,PROD_COMM)(all,r,REG)
    pmes(i,j,r) # market price of sluggish endowment i used by j, in r #;
Coefficient (ge 0)(all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFM(i,j,r) # producer expenditure on i by j in r valued at mkt prices #;
Update (all,i,ENDWM_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFM(i,j,r) = pm(i,r) * qfe(i,j,r);
Update (all,i,ENDWS_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFM(i,j,r) = pmes(i,j,r) * qfe(i,j,r);
Read
    VFM from file GTAPDATA header "VFM";
```

Hence they are valued in market price terms. This means that there is a tax on the use of factors that is activity specific – this is explained below in the section on indirect taxes.

Sales of factor services to domestic activities are the sole sources of factor incomes in the database.

Factor Expenditures

Factor incomes are distributed to domestic agents in three sub matrices. The first is the allowance for depreciation (14:4)

```
Variable (all, r, REG)
    kb(r) # beginning-of-period capital stock, in r #;
Variable (all, r, REG)
    pcgds(r) # price of investment goods = ps("cgds",r) #;
Coefficient (ge 0)(all, r, REG)
    VDEP(r) # value of capital depeciation in r (exogenous) #;
Update (all, r, REG)
    VDEP(r) = kb(r) * pcgds(r);
Read
    VDEP from file GTAPDATA header "VDEP";
```

which is recorded as an expenditure by the factors accounts to the domestic capital account, i.e., a form of saving.

1984). Either arrangement is acceptable with the reduced form input-output table used in the GTAP database.

[20] Note that the term for *VFM* below is defined according to whether an endowment is defined as MOBILE or SLUGGISH.

Second the factor account must pay factor income taxes, which are defined as the totals of the payments to each factor by all activities less the 'endowments output at agents' prices (12:4). These are explained below in the section on 'Factor Income Taxes'.

And the third component is the distribution of factor income after allowance for depreciation and the payment of factor taxes to the regional household (5:4). Hence the factor incomes distributed to the regional household are the 'endowments output at agents' prices'

```
Variable (all,i,NSAV_COMM)(all,r,REG)
    ps(i,r) # supply price of commodity i in region r #;
Variable (orig_level=VOM)(all,i,NSAV_COMM)(all,r,REG)
    qo(i,r) # industry output of commodity i in region r #;
Coefficient (ge 0)(all,i,ENDW_COMM)(all,r,REG)
    EVOA(i,r) # value of commodity i output in region r #;
Update (all,i,ENDW_COMM)(all,r,REG)
    EVOA(i,r) = ps(i,r) * qo(i,r);
Read
    EVOA from file GTAPDATA header "EVOA";
```

less the allowance for depreciation.

Regional Household Income

Regional household income consists of distributed factor incomes, which were explained above, and total tax revenues, which are defined as indirect tax income plus direct tax income. Because of the Regional Household system, which means that private households pay no direct/income taxes and the private household and government do not save, the tax revenues in GTAP are dominated by indirect taxes.

The definitions for these taxes are crucial to an accurate accounting of transactions in the GTAP database. The definitions for the tax receipts in the GTAP database are taken from version 6.2 of the GTAP model tablo file, as set out in the section 'Computation of Indirect Tax Receipts by the Regional Household'.[21]

Indirect Tax Revenues

Indirect tax receipts are defined in the GTAP model tablo file (gtab.tab v6.1) as

```
INDTAX(r) # indirect tax receipts in r #;
    INDTAX(r) =
    TPC(r) + TGC(r) + TIU(r) + TFU(r) + TOUT(r) + TEX(r) + TIM(r);
where

        TIM(r) # import tax payments in r #;
        TEX(r) # export tax payments in r #;
        TIU(r) # firms' tax payments on intermediate goods usage in r #;
        TGC(r) # government consumption tax payments in r #;
        TPC(r) # private consumption tax payments in r #;
        TOUT(r) # production tax payments in r #;
```

---

[21]   This starts at line 747 in gtap61.tab.

```
      TFU(r) # firms' tax payments on primary factor usage in r #;
```

The definitions of each of these indirect tax instruments in the GTAP database and how they are treated in the global SAM are explained below. When considering the GTAP tax instruments it is important to note the convention in the GTAP database and model by which the tax revenue is defined; the right hand side elements of the tax revenue equations are ordered A(gents), M(arket), W(orld), and hence it is also important to note how prices are formed in the GTAP model.

Trade Tax Revenues

    Import Duties (7:1)

The GTAP model defines the revenues from import duties (*MTAX*) on commodity *i* imported from region *r* by region *s* as

```
Coefficient (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    MTAX(i,r,s) # tax on imports of good i from source r in destination s #;
Formula (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    MTAX(i,r,s) = VIMS(i,r,s) - VIWS(i,r,s);
Coefficient (all,r,REG)
    TIM(r) # import tax payments in r #;
Formula (all,r,REG)
    TIM(r) = sum(i,TRAD_COMM,sum(s,REG, MTAX(i,s,r)));
```

    Further the (rates of change in) market prices (*pms*) are defined as

```
Equation MKTPRICES
# eq'n links domestic and world prices (HT 24) #
(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    pms(i,r,s) = tm(i,s) + tms(i,r,s) + pcif(i,r,s);
```

where *tms* is the percentage change in the power of the import duty rate and *pcif* is the *cif* paid import price (*tm* is a variable included for purposes of simulations and in the current context is not relevant). Hence import duty revenues must be added to the payments to the source region to identify the payments for imports made by domestic consumers.[22]

    Export Taxes (7:2)

Export taxes are defined as the revenues from taxes charged on exports (*ETAX*) on commodity *i* exported from region *r* to region *s*

```
Coefficient (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    XTAXD(i,r,s) # tax on exports of good i from source r to destination s #;
Formula (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    XTAXD(i,r,s) = VXWD(i,r,s) - VXMD(i,r,s);
Coefficient (all,r,REG)
```

---

[22]   Supplementary data in the GTAP database allow for the identification of import duties that are levied *ad valorem* and those levied on quantities. The SAM produced by GTAPAgg does not include these data. The GAMS codes (http://cgemod.org.uk/gtpsam.html) for versions 8 and 9 create SAMs that include the supplementary import duty data in the SAM.

```
    TEX(r) # export tax payments in r #;
Formula (all,r,REG)
    TEX(r) = sum(i,TRAD_COMM, sum(s,REG, XTAXD(i,r,s)));
```

The (rates of change in) export prices (*pfob*) are defined as

```
Equation EXPRICES
# eq'n links agent's and world prices (HT 27) #
(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    pfob(i,r,s) = pm(i,r) - tx(i,r) - txs(i,r,s);
```

where *txs* is the percentage change in the power of the export tax rate, *pfob* the *fob* export price and *pm* the market price of commodity *i* in region *r*. Hence export tax revenues must be subtracted from the payments by the purchasing region to identify the revenues from exports received by domestic producers.

This is confirmed by the definitions of *VXMD*

```
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VXMD(i,r,s)
    # exports of i from r to s valued at mkt prices (tradeables only) #;
Update (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VXMD(i,r,s) = pm(i,r) * qxs(i,r,s);
```

and *VXWD*

```
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VXWD(i,r,s) # exports of i from r to s valued fob (tradeables only) #;
Update (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VXWD(i,r,s) = pfob(i,r,s) * qxs(i,r,s);
Read
    VXWD from file GTAPDATA header "VXWD";
```

Hence export taxes are defined by *VXWD - VXMD*.[23]

Taxes on Demand/Consumption

There are 6/8 indirect taxes on demand – indirect taxes on demand for domestic and imported commodities by the private household, government, investment and activities, with the investment and activity demands being 'lumped' together under the label of firms' demand.

Transactions between activities, the private household, government and investment and the commodity accounts are valued at the sellers' prices – which for these transactions are termed market prices by GTAP. Then transactions between the institutions and the sales tax accounts are recorded that identify the 'sales' taxes paid by each institution with respect to purchases of each commodity. Consequently sales tax transactions do not appear as row entries in the columns for the commodity accounts. While the purchasing institutions pay different prices the sellers receive a common price (*pm* or *pim*) irrespective of the institution

---

[23]   The convention followed in this SAM is that the definitions for prices are built up as accounting identities through the columns of the commodity accounts. Hence export taxes are accounted for as the difference

that makes the purchase; hence the domestic supply matrix is recorded in market price terms (broadly equivalent to basic prices), while the supply of imported commodities is valued inclusive of trade taxes and transport margins for imported commodities, but excluding any domestic sales taxes.

The allocation method described below retains all the information in the GTAP database. Alternative representations of the GTAP database are consequently reduced forms.

Taxes on Intermediate and Investment Demand (8:3, 9:3, 8:14, 9:14)

Taxes on purchases of intermediate and investment commodity differ according to whether the commodity is domestically produced or imported, i.e.,

```
Variable (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    tfd(i,j,r) # tax on domestic i purchased by j in r #;
Variable (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    tfm(i,j,r) # tax on imported i purchased by j in r #;
```

hence there are separate tax revenues for domestic and imported commodities, i.e.,

```
Coefficient (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    DFTAX(i,j,r) # tax on use of domestic intermediate good i by j in r #;
Formula (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    DFTAX(i,j,r) = VDFA(i,j,r) - VDFM(i,j,r);
Coefficient (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    IFTAX(i,j,r) # tax on use of imported intermediate good i by j in r #;
Formula (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    IFTAX(i,j,r) = VIFA(i,j,r) - VIFM(i,j,r);
Coefficient (all,r,REG)
    TIU(r) # firms' tax payments on intermediate goods usage in r #;
Formula (all,r,REG)
    TIU(r) = sum(j,PROD_COMM, sum(i,TRAD_COMM, DFTAX(i,j,r) + IFTAX(i,j,r)));
```

But taxes are defined by

```
Equation DMNDDPRICE
# eq'n links domestic market and firm prices (HT 20) #
(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    pfd(i,j,r) = tfd(i,j,r) + pm(i,r);
Equation DMNDIPRICES
# eq'n links domestic market and firm prices (HT 23) #
(all,i,TRAD_COMM)(all,j,PROD_COMM)(all,r,REG)
    pfm(i,j,r) = tfm(i,j,r) + pim(i,r);
```

where *pfd* and *pfm* define the prices paid by the 'activities' while *pm* and *pim* define the price received by the suppliers. Hence *VDFM* and *VIFM* record the transactions between the commodity and 'activity' accounts and *DFTAX* and *IFTAX* define the associated tax payments by 'activities' to the government.

Note however that *VDFM/VFDA* and *VIFM/VIFA* are defined over PROD_COMM and hence 'activities' include the investment account. In the GAMS programme the allocations of

---

between the price paid by the destination region and the price received by the domestic activity. The notable exception is the taxes on consumption.

these taxes to their respective sub matrices of the global SAM are controlled by (conditional)

mapping sets.

Taxes on Private Household Demand (8:6, 9:6)

Taxes on purchases of commodities by the private households differ according to whether the

commodities are domestically produced or imported, i.e.,

```
Variable (all,i,TRAD_COMM)(all,r,REG)
    tpd(i,r) # shock to tax on domestic i purchased by private hhld in r #;
Variable (all,i,TRAD_COMM)(all,r,REG)
    tpm(i,r) # shock to tax on imported i purchased by private hhld in r #;
```

hence there are separate tax revenues for domestic and imported commodities, i.e.,

```
Coefficient (all,i,TRAD_COMM)(all,r,REG)
    DPTAX(i,r) # tax on private consumption of domestic good i in region r #;
Formula (all,i,TRAD_COMM)(all,r,REG)
    DPTAX(i,r) = VDPA(i,r) - VDPM(i,r);
Coefficient (all,i,TRAD_COMM)(all,r,REG)
    IPTAX(i,r) # tax on private consumption of imported good i in region r #;
Formula (all,i,TRAD_COMM)(all,r,REG)
    IPTAX(i,r) = VIPA(i,r) - VIPM(i,r);
Coefficient (all,r,REG)
    TPC(r) # private consumption tax payments in r #;
Formula (all,r,REG)
    TPC(r) = sum(i,TRAD_COMM, DPTAX(i,r) + IPTAX(i,r));
```

But taxes are defined by

```
Equation PHHDPRICE
# eq'n links domestic market and private consumption prices (HT 18) #
(all,i,TRAD_COMM)(all,r,REG)
    ppd(i,r) = atpd(i,r) + pm(i,r);
Equation PHHIPRICES
# eq'n links domestic market and private consumption prices (HT 21) #
(all,i,TRAD_COMM)(all,r,REG)
    ppm(i,r) = atpm(i,r) + pim(i,r);
```

where

```
Variable (all,r,REG)
    tp(r) # comm.-, source-gen. shift in tax on private cons. #;
Variable (all,i,TRAD_COMM)(all,r,REG)
    tpd(i,r) # comm.-, source-spec. shift in tax on private cons. of dom. #;
Variable (all,i,TRAD_COMM)(all,r,REG)
    atpd(i,r) # power of tax on domestic i purchased by private hhld in r #;
Equation TPDSHIFT
# permits uniform consumption tax change #
(all,i,TRAD_COMM)(all,r,REG)
    atpd(i,r) = tpd(i,r) + tp(r);
Variable (all,i,TRAD_COMM)(all,r,REG)
    tpm(i,r) # comm.-, source-spec. shift in tax on private cons. of imp. #;
Variable (all,i,TRAD_COMM)(all,r,REG)
    atpm(i,r) # power of tax on imported i purchased by private hhld in r #;
Equation TPMSHIFT
# permits uniform consumption tax change #
(all,i,TRAD_COMM)(all,r,REG)
    atpm(i,r) = tpm(i,r) + tp(r);
```

where *ppd* and *ppm* define the prices paid by the private households while *pm* and *pim* define

the price received by the suppliers. Hence *VDPM* and *VIPM* – the values of private demand at

market prices for domestic and imported commodities - record the transactions between the

commodity and private household accounts and *DPTAX* and *IPTAX* define the associated tax

payments by private households to the government.

Taxes on Government Demand (8:13, 9:13)

Taxes on purchases of commodities by the government differ according to whether the

commodities are domestically produced or imported, i.e.,

```
Variable (all,i,TRAD_COMM)(all,r,REG)
    tgd(i,r) # tax on domestic i purchased by government hhlds in r #;
Variable (all,i,TRAD_COMM)(all,r,REG)
    tgm(i,r) # tax on imported i purchased by gov't hhld in r #;
```

hence there are separate tax revenues for domestic (DGTAX) and imported (IGTAX)

commodities, i.e.,

```
Coefficient (all,i,TRAD_COMM)(all,r,REG)
    DGTAX(i,r) # tax on government consumption of domestic good i in region r #;
Formula (all,i,TRAD_COMM)(all,r,REG)
    DGTAX(i,r) = VDGA(i,r) - VDGM(i,r);
Coefficient (all,i,TRAD_COMM)(all,r,REG)
    IGTAX(i,r) # tax on government consumption of imported good i in region r #;
Formula (all,i,TRAD_COMM)(all,r,REG)
    IGTAX(i,r) = VIGA(i,r) - VIGM(i,r);
Coefficient (all,r,REG)
    TGC(r) # government consumption tax payments in r #;
Formula (all,r,REG)
    TGC(r) = sum(i,TRAD_COMM, DGTAX(i,r) + IGTAX(i,r));
```

But taxes are defined by

```
Equation GHHDPRICE
# eq'n links domestic market and government consumption prices (HT 19) #
(all,i,TRAD_COMM)(all,r,REG)
    pgd(i,r) = tgd(i,r) + pm(i,r);
Equation GHHIPRICES
# eq'n links domestic market and government consumption prices (HT 22) #
(all,i,TRAD_COMM)(all,r,REG)
    pgm(i,r) = tgm(i,r) + pim(i,r);
```

where *pgd* and *pgm* define the prices paid by the government while *pm* and *pim* define the

price received by the suppliers. Hence *VDGM* and *VIGM* – the values of government demand

at market prices for domestic and imported commodities - record the transactions between the

commodity and private household accounts and *DGTAX* and *IGTAX* define the associated tax

payments by government to the government.

Factor Use Taxes (10:3)

These are taxes on factor use paid by the activities. The GTAP tablo file defines these as

follows

```
Coefficient (all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    ETAX(i,j,r) # tax on use of endowment good i by industry j in region r #;
Formula (all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    ETAX(i,j,r) = VFA(i,j,r) - VFM(i,j,r);
Coefficient (all,r,REG)
    TFU(r) # firms' tax payments on primary factor usage in r #;
```

```
Formula (all,r,REG)
    TFU(r) = sum(i,ENDW_COMM, sum(j,PROD_COMM, ETAX(i,j,r)));
```
where
```
Coefficient (all,i,DEMD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFA(i,j,r) # producer expenditure on i by j in r valued at agents' prices #;
Formula (all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFA(i,j,r) = EVFA(i,j,r);
Formula (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,s,REG)
    VFA(i,j,s) = VDFA(i,j,s) + VIFA(i,j,s);
Coefficient (ge 0)(all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    EVFA(i,j,r) # producer expenditure on i by j in r at agent's prices #;
```

The coefficient *ETAX* is defined over three sets

    i)   ENDW_COMM – the factors of production (*f* in GAMS);

    ii)   PROD_COMM – the activities (*a* in GAMS) and investment goods ("*kap*" in GAMS);

    iii)   REG – the regions (*k* in GAMS).

However note that while *VFM* is defined over the same three sets, *VFA* is defined over three sets, one of which is different, i.e.,

    i)   DEMD_COMM – the factors of production (*f* in GAMS) plus the activities (*a* in GAMS), i.e., ENDW_COMM is a subset of DEMD_COMM;

    ii)   PROD_COMM – the activities (*a* in GAMS) and investment goods ("*kap*" in GAMS);

    iii)   REG – the regions (*k* in GAMS).

Hence only part of *VFA* is used in the determination of the factor use taxes, and in the GTAP model *VFA* is derived from the read in data. Furthermore, note that there are no entries for *CGDS* in either *VFM* or *EVFA*. Consequently when assigning the transaction values for factor use taxes it is necessary exclude the elements of *VFM* and *EVFA* that refer to *s* to capture the taxes paid by activities for the use of factors – this is achieved by the use of a series of (conditional) mapping sets.

'Production' Taxes (11:3, 12:4 and 11:14)

These are the taxes identified by the parameter *PTAX* in the GTAP model. *PTAX* is defined over two sets

    1.   NSAV_COMM – the factors of production (*f* in GAMS), the activities (*a* in GAMS) and investment goods ("*kap*" in GAMS);

    2.   REG – the regions (*k* in GAMS).

The GTAP databases defines production tax payments as

```
Coefficient (all,i,NSAV_COMM)(all,r,REG)
    PTAX(i,r) # output tax on good i in region r #;
Formula (all,i,NSAV_COMM)(all,r,REG)
    PTAX(i,r) = VOM(i,r) - VOA(i,r);
Coefficient (all,r,REG)
    TOUT(r) # production tax payments in r #;
Formula (all,r,REG)
    TOUT(r) = sum(i,PROD_COMM, PTAX(i,r));
```

The value of output at agents prices (*VOA*) is relatively straightforward since it only uses 2 subsets of NSAV_COMM.

```
Coefficient (all,i,NSAV_COMM)(all,r,REG)
    VOA(i,r) # value of commodity i output in region r at agents' prices #;
Formula (all,i,ENDW_COMM)(all,r,REG)
    VOA(i,r) = EVOA(i,r);
Formula (all,i,PROD_COMM)(all,r,REG)
    VOA(i,r) = sum(j,DEMD_COMM, VFA(j,i,r));
```

with

```
Coefficient (all,i,DEMD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFA(i,j,r) # producer expenditure on i by j in r valued at agents' prices #;
Formula (all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFA(i,j,r) = EVFA(i,j,r);
Formula (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,s,REG)
    VFA(i,j,s) = VDFA(i,j,s) + VIFA(i,j,s);
```

while the value of output at market prices (*VOM*) draws on more subsets of NSAV_COMM. i.e.,

```
Coefficient (all,i,NSAV_COMM)(all,r,REG)
    VOM(i,r) # value of commodity i output in region r at market prices #;
Formula (all,i,ENDW_COMM)(all,r,REG)
    VOM(i,r) = sum(j,PROD_COMM, VFM(i,j,r));
Formula (all,m,MARG_COMM)(all,r,REG)
    VOM(m,r) = VDM(m,r) + sum(s,REG, VXMD(m,r,s)) + VST(m,r);
Formula (all,i,NMRG_COMM)(all,r,REG)
    VOM(i,r) = VDM(i,r) + sum(s,REG, VXMD(i,r,s));
Formula (all,h,CGDS_COMM)(all,r,REG)
    VOM(h,r) = VOA(h,r);
```

with

```
Coefficient (all,i,TRAD_COMM)(all,r,REG)
    VDM(i,r) # domestic sales of i in r at mkt prices (tradeables only) #;
Formula (all,i,TRAD_COMM)(all,r,REG)
    VDM(i,r) = VDPM(i,r) + VDGM(i,r) + sum(j,PROD_COMM, VDFM(i,j,r));
```

Finally the general equation for the *PTAX* is

```
Equation OUTPUTPRICES
# eq'n links pre- and post-tax supply prices for all industries  (HT 15) #
(all,i,PROD_COMM)(all,r,REG)
    ps(i,r) = to(i,r) + pm(i,r);
!<
    This equation links pre- and post-tax supply prices for all industries.
    This captures the effect of output taxes. TO(i,r) < 1 in the case of a
    tax. (HT 15)
>!
```

where

```
Variable (all,i,NSAV_COMM)(all,r,REG)
    to(i,r) # output (or income) tax in region r #;
```

While the parameters/coefficients *VOM* and *VOA* are defined over 2 sets they are assigned values from a number of sources whose values are defined over different sets. Since these different sets ultimately refer to different sub matrices of the SAM it will prove convenient to decompose the coefficient *PTAX* and therefore the coefficients *VOA* and *VOM*. The set for *i* in *PTAX* suggests that *PTAX* will be paid by three types of agents in the SAM – activities, factors and the investment account – hence the SAM should allow for non-zero entries in the sub matrices 11:3, 12:4 and 11:14.[24]

Hence all transactions that go into *PTAX* should be valued in terms of *ps* and *pm*.

It is convenient to deal with *PTAX* in three parts; one each for activities, investment goods and factors. The endowment/factor component is most straightforward.

Factor Income Taxes (12:4)

PTAX for factors – *PTAXFACT* – is defined over the sets for factors (*f*) and regions (*r*) as

```
Formula (all,i,NSAV_COMM)(all,r,REG)
    PTAX(i,r) = VOM(i,r) - VOA(i,r);
Formula (all,i,ENDW_COMM)(all,r,REG)
    VOA(i,r) = EVOA(i,r);
Formula (all,i,ENDW_COMM)(all,r,REG)
    VOM(i,r) = sum(j,PROD_COMM, VFM(i,j,r));
```

where

```
Coefficient (ge 0)(all,i,ENDW_COMM)(all,r,REG)
    EVOA(i,r) # value of commodity i output in region r #;
Update (all,i,ENDW_COMM)(all,r,REG)
    EVOA(i,r) = ps(i,r) * qo(i,r);
Read
    EVOA from file GTAPDATA header "EVOA";
```

and

```
Coefficient (ge 0)(all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFM(i,j,r) # producer expenditure on i by j in r valued at mkt prices #;
Update (all,i,ENDWM_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFM(i,j,r) = pm(i,r) * qfe(i,j,r);
Update (all,i,ENDWS_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFM(i,j,r) = pmes(i,j,r) * qfe(i,j,r);
Read
    VFM from file GTAPDATA header "VFM";

Equation FACTORINCPRICES
# eq'n links pre- and post-tax endowment supply prices (HT 15) #
(all,i,ENDW_COMM)(all,r,REG)
    ps(i,r) = to(i,r) + pm(i,r);
```

Hence the factor income tax transactions can be defined as

$$PTAXFACT_{ff,kk} = \sum_{prod\_comm} VFM_{endwm\_comm,prod\_comm,reg} - EVOA_{endw\_comm,reg}$$

$$= \sum_{aa} VFM2_{ff,aa,kk} - EVOA_{ff,kk}$$

where *VFM2* is a defined as *VFM* excluding the entries for '*CGDS*'. Hence these are taxes paid by factors direct to the government, and, in terms of the SAM relationships, are defined as the difference between payments to factors and the disposable income of factors after payment of depreciation and taxes, i.e., factor income taxes.

Since version 7 of the GTAP database Factor Income taxes have been defined to include direct taxes paid by households and incorporated business enterprises (see McDougall and Hagemjer, 2016).

Taxes on Production (11:3)

PTAX for activities – *PRODTAX* – is defined over the sets for activities (*a*) and regions (*k*) as

```
Formula (all,i,NSAV_COMM)(all,r,REG)
    PTAX(i,r) = VOM(i,r) - VOA(i,r);
Formula (all,i,PROD_COMM)(all,r,REG)
    VOA(i,r) = sum(j,DEMD_COMM, VFA(j,i,r));
Formula (all,m,MARG_COMM)(all,r,REG)
    VOM(m,r) = VDM(m,r) + sum(s,REG, VXMD(m,r,s)) + VST(m,r);
Formula (all,i,NMRG_COMM)(all,r,REG)
    VOM(i,r) = VDM(i,r) + sum(s,REG, VXMD(i,r,s));
```

with

```
Coefficient (all,i,TRAD_COMM)(all,r,REG)
    VDM(i,r) # domestic sales of i in r at mkt prices (tradeables only) #;
Formula (all,i,TRAD_COMM)(all,r,REG)
    VDM(i,r) = VDPM(i,r) + VDGM(i,r) + sum(j,PROD_COMM, VDFM(i,j,r));
Coefficient (all,i,DEMD_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFA(i,j,r) # producer expenditure on i by j in r valued at agents' prices #;
Formula (all,i,ENDW_COMM)(all,j,PROD_COMM)(all,r,REG)
    VFA(i,j,r) = EVFA(i,j,r);
Formula (all,i,TRAD_COMM)(all,j,PROD_COMM)(all,s,REG)
    VFA(i,j,s) = VDFA(i,j,s) + VIFA(i,j,s);
```

VOM has already been defined over the TRAD_COMM set, but it remains to define *VOA* over this set. Working with PROD_COMM for the moment, the formula above indicates that *VOA* is derived from *VFA*, where *VFA* defines matrices of dimension DEMD_COMM and PROD_COMM for each region *r*. The component of *VFA* referring to ENDW_COMM comes from *EVFA* while the components referring to PROD_COMM come from *VDFA* and *VIFA*. Hence

$$VOA_{aa,kk} = \sum_{ff} EVFA_{ff,aa,kk} + \sum_{dcc} VDFA_{dcc,aa,kk} + \sum_{mcc} VIFA_{mcc,aa,kk}$$

which simply defines the value of output by activity *a* in region *r* as the sum of the payments to primary inputs at agents prices (*EVFA*) plus the sum of domestic (*VDFA*) and imported (*VIFA*) inputs valued at agents prices. But *VOM* is defined as

$$VOM_{dcc,kk} = VDM_{dcc,kk} + VXMD_{dcc,kk} + VST_{m,kk}$$

which suggests that the dimension of *VOA* and *VOM* are non-commensurate. However, the elements of *VOM* are the entries on the diagonal supply matrix and hence, since there is a one to one correspondence between the members of the sets *dcc* and *aa*, it is possible to define $VOM_{dcc,kk} = VOM_{aa,kk}$ using a (conditional) mapping set for *dcc* and *aa*.

Taxes on Investment (11:14)

PATX for investment – *PTAXINV* – is straightforward. The expression for the power of the tax comes with a warning

```
Variable (all,i,NSAV_COMM)(all,r,REG)
    to(i,r) # output (or income) tax in region r #;
!<
    Note: it is important that the user NOT shock the tax on capital goods
    output, as this will cause an inconsistency in the update relationship
    for VOM(cgds).
>!
```

This indicates that there is no facility for *PTAXINV* in the GTAP model. A conclusion that is confirmed by the formula for the output valuations at market and agents prices for capital goods, i.e.,

```
Formula (all,h,CGDS_COMM)(all,r,REG)
    VOM(h,r) = VOA(h̄,r);
```

Hence by definition *PTAXINV* is zero.

Taxes and Regional Household Income (5:7, 5:8, 5:8, 5:10, 5:11, 5:12)

All the tax instruments discussed above are expressed in terms of the expenditures on taxes by different agents in different contexts, and in this SAM all these taxes are paid to accounts relating to particular tax instruments. To convert these into incomes to the regional household it is only necessary to compute the row total for each tax instrument and transpose these as row entries into the regional household account.

Trade Accounts

Trade transactions are at the heart of the GTAP database. A distinctive feature of versions 5 and 6 of the GTAP database is the recording of trade margins; in essence this allows the identification of origin specific trade and transport costs associated with imports of

commodities by each region with each region contributing to a global trade and transport 'pool' from which trade and transport services are drawn. Consequently, while the use of trade and transport services associates the use of services with specific imports, the specific suppliers of these services are not identified.

Imports (15:1, 16:5)

Imports at world prices (*VIWS*) are valued carriage, insurance and freight (cif) paid, i.e.,

```
Variable (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    pcif(i,r,s) # CIF world price of commodity i supplied from r to s #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VIWS(i,r,s) # imports of i from r to s valued cif (tradeables only) #;
Update (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VIWS(i,r,s) = pcif(i,r,s) * qxs(i,r,s);
Read
    VIWS from file GTAPDATA header "VIWS";
```

where

```
Variable (orig_level=VXMD)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    qxs(i,r,s) # export sales of commodity i from r to region s #;
```

while imports at market prices (*VIMS*) are valued inclusive of import duties (see the discussion of import duties in the section on indirect taxes)

```
Variable (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    pms(i,r,s) # domestic price for good i supplied from r to region s #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VIMS(i,r,s) # imports of i from r to s valued at domestic mkt prices #;
Update (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VIMS(i,r,s) = pms(i,r,s) * qxs(i,r,s);
Read
    VIMS from file GTAPDATA header "VIMS";
```

Hence VIWS defines the import transactions inclusive of trade and transport costs. The GTAP model defines

```
Coefficient (ge 0) (all,m,MARG_COMM)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VTMFSD(m,i,r,s)
    # int'l margin usage, by margin, freight, source, and destination #;
Update (all,m,MARG_COMM)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VTMFSD(m,i,r,s) = pt(m) * qtmfsd(m,i,r,s);
Read
    VTMFSD from file GTAPDATA header "VTWR";
```

where

```
Variable (all,m,MARG_COMM)
    pt(m) # price of composite margins services, type #;
Variable (all,m,MARG_COMM)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    qtmfsd(m,i,r,s) # international usage margin m on i from r to s #;
```

such that *VTMFSD*/*VTWR*[25] records the transactions on margin services (*m*) associated with commodity *i* exported by region *r* to region *s*.

---

[25]   In the GEMPACK model code the coefficient is called *VTMFSD* while in the HAR file the array is labeled *VTWR*. The HAR file label is carried over to the GAMS programme by the HAR2GDX utility.

Hence transactions with other regions for imported commodities are defined as (*VIWS* – *VTMFSD*) and transactions with the global pool of transport services associated with these imports are defined by *VTMFSD*.

However, this direct method of calculating the value of import transactions (*fob*) does encounter a problem for a programme in GAMS. Specifically imports (*fob*) of commodity *c* from region *r* by region *s* are not identical to the exports (*fob*) of commodity *c* to region *r* from region *s*. The differences are minimal – most only appear at the fourth or more decimal places – but they do create some problems when verifying that the global SAM is complete and consistent and they impede subsequent transformations of the global SAM. To avoid these potential problems the values of import transaction (*fob*) are actually calculated indirectly; specifically imports (*fob*) are computed by transposing the 3-dimensional matrix of exports (*fob*) (for details see below). This ensures that the values of imports and exports are identical, which avoids any subsequent problems.

Exports (2:16, 2:17)

Export transactions are recorded at free on board (*fob*) prices, i.e., at the price paid by the importing region before the payment of trade and transports costs. The GTAP databases records this as *VXWD*

```
Variable (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    pfob(i,r,s) # FOB world price of commodity i supplied from r to s #;
Coefficient (ge 0)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VXWD(i,r,s) # exports of i from r to s valued fob (tradeables only) #;
Update (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VXWD(i,r,s) = pfob(i,r,s) * qxs(i,r,s);
Read
    VXWD from file GTAPDATA header "VXWD";
```

where

```
Equation EXPRICES
# eq'n links agent's and world prices (HT 27) #
(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    pfob(i,r,s) = pm(i,r) - tx(i,r) - txs(i,r,s);
```

and hence the world price (*pfob*) differs from the market price (*pm*) by the export taxes (*txs*).[26]

But the database records separately the exports of transport services as

```
Variable (orig_level=VST)(all,m,MARG_COMM)(all,r,REG)
    qst(m,r) # sales of m from r to international transport #;
Coefficient (ge 0)(all,m,MARG_COMM)(all,r,REG)
    VST(m,r)
    # exprts of m from r for int'l trnsport valued at mkt p (tradeables only) #;
Update (all,m,MARG_COMM)(all,r,REG)
    VST(m,r) = pm(m,r) * qst(m,r);
Read
    VST from file GTAPDATA header "VST";
```

---

[26]   *tx* is included for simulation purposes only and hence is redundant in the current context.

note that the price (*pm*) on exports of transport services is a market price, i.e., there is no allowance for export taxes on transport services.

Margins

The transport margins have been identified above but it also worth noting that the relationships between imports and exports also require an explicit balancing of the *fob* exports from a source region with the *cif* imports and margins by a destination region, i.e.,

```
In a balanced data base, (all,i,TRAD_COMM)(all,r,REG)(all,s,REG),
VIWS(i,r,s) = VXWD(i,r,s) + VTFSD(i,r,s)
```

where

```
Coefficient (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VTFSD(i,r,s) # aggregate value of svces in the shipment of i from r to s #;
Formula (all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VTFSD(i,r,s) = sum(m,MARG_COMM, VTMFSD(m,i,r,s));
Coefficient (ge 0) (all,m,MARG_COMM)(all,i,TRAD_COMM)(all,r,REG)(all,s,REG)
    VTMFSD(m,i,r,s)
        # int'l margin usage, by margin, freight, source, and destination #;
```

Capital Account (14:4, 14:5, 14:16, 14:17)

Income to the capital account consists of four elements, two of which are related to the domestic accounts and two to the trade accounts. Domestic savings consist of the allowance for depreciation (see above) and net domestic saving, which is defined in the GTAP tablo file as

```
Variable (all, r, REG)
    psave(r) # price of savings in region r #;
Variable (all,r,REG)
    qsave(r) # regional demand for NET savings #;
Coefficient (all,r,REG)
    SAVE(r) # expenditure on NET savings in region r valued at agent's prices #;
Update (all,r,REG)
    SAVE(r) = psave(r) * qsave(r);
Read
    SAVE from file GTAPDATA header "SAVE";
```

Following the convention in the GTAP model all savings are assumed to originate from the regional household.

This leaves the balance on the capital account. Strictly this could be defined as the differences in the values of exports *fob* and import *cif*, but the GTAP tablo file treats trade margins through the device of a global transport pool, i.e., while specific transport transactions are associated with imports from specific origins, they are not mapped onto the region that supplies the transport services. Rather transport services are exported to a global transport pool. Hence the balance on the capital account with the rest of the world consists of elements relating to both commodities and transport services. The capital account balance

with each trading partner is defined as the value of imports valued *cif* less the value of transport services exported to the global transport pool less the value of export. The capital account balance for each region with the global transport pool is defined as the values of transport services imported by region less the value of transport services exported to the global transport pool.

5 Reconciling the Global SAMs

To check that the global SAM is consistent, it is important to verify that the row and column totals equate, since this is the most reliable way of checking that the procedure has been implemented 'correctly'. This involves two related sets of checks. It is necessary

1. to check that the row and column totals of each regional SAM are the same, and
2. to verify that each and every component of the transactions between regions is identical to its matching counterpart.

If any differences are identified it is necessary to determine whether these are a consequence of 'errors' in the base data or a result of 'errors' introduced when transforming the base data into the global SAM.

The resultant global SAM is not precisely balanced, i.e., the row and column totals differ, although the extent of the imbalance is small.[27] The number of accounts and cells in the SAMs for GTAP version 5.4 and 9.1 are reported in Table 5; note the threefold increase in the number of accounts (number of rows/columns) and the fivefold increase in the number of cells.

Table 5   Numbers of Accounts and Cells in Versions 5.4 and 9.1 of the Global SAM

Version	Accounts per Region	Regions	Total Number of Accounts	Total Number of Cells
Version 5.4	772	78	60,216	46,486,752
Version 9.1	1,357	140	189,980	257,802,860

Tables 6a and 6b provide summary statistics about the absolute and proportionate differences between the row and column totals for the SAMs derived from versions 5.4 and 9.1 of the GTAP database. These statistics confirm that while there are an appreciable number of cases for which there are absolute and percentage differences in the accounts totals these are for a relative small proportion of the SAM accounts. There is some evidence to suggest that both the frequency and magnitude of differences in the row and column are increasing;

---

[27]   The differences between the row and column totals are small and consequently are unlikely to be relevant when solving a model, especially when solved to single precision as in GEMPACK.

this is not altogether surprising since with the increase in the number of regions there must have been a substantial increase in the number of non-zero zeros in the database.

Overwhelmingly, the computed differences between the row and column totals occur in the imported commodity and institutional accounts, and the imbalances for the Regional Household and Capital accounts are overwhelmingly a consequence of aggregations of the imbalances in the imported commodity accounts. Furthermore since all the checks on the data indicate that the differences in the row and column totals can be traced back to the GTAP database, it can be safely concluded that they are not consequences of errors in the conversion to a SAM representation, rather that they are differences that are within the tolerances set by the compilers of the database and hence are 'rounding errors'.[28] Nevertheless, even minor differences in total incomes and expenditures can cause modelling difficulties; hence routines are included in the programmes to resolve these differences.[29]

Table 6a   **Absolute and Percentage Differences in Row and Column Totals in the SAM - Version 5.4 of the GTAP Database**

Absolute Diffences		Percentage Diffences		Absolute and Percentage Diffences	
Critical Value	**Cases**	**Critical Value**	**Cases**	**Critical Value**	**Cases**
5	0	5	8	5	0
1	0	1	13	1	0
0.5	1	0.5	16	0.5	0
0.1	6	0.1	22	0.1	0
0.01	38	0.01	40	0.01	0
0.001	167	0.001	86	0.001	1
0.0001	634	0.0001	220	0.0001	457
0.00001	1,335	0.00001	1,419	0.00001	1,307
0.000001	2,003	0.000001	2,079	0.000001	1,986

Source: Authors calculations.

---

[28]   One difficulty encountered when seeking to check that 'rounding errors' can be traced back to the database is the fact that some elements of the SAM are calculated as residuals because they are not explicitly reported in the GTAP database.

[29]   If these routines are not wanted then the user can switch off the relevant sections of the programme, i.e., sections 10 and 11.

Table 6b   **Absolute and Percentage Differences in Row and Column Totals in the SAM - Version 9.1 (2011) of the GTAP Database**

Absolute Diffences		Percentage Diffences		Absolute and Percentage Diffences	
Critical Value	Cases	Critical Value	Cases	Critical Value	Cases
5	0	5	1,087	5	0
1	0	1	2,035	1	0
0.5	1	0.5	2,550	0.5	0
0.1	10	0.1	4,206	0.1	0
0.01	148	0.01	7,579	0.01	0
0.001	1,110	0.001	12,532	0.001	20
0.0001	4,072	0.0001	20,485	0.0001	442
0.00001	9,512	0.00001	31,919	0.00001	3,906

Source: Authors calculations.

Reconciling the SAMs from the GTAP Database

There are three distinct categories of 'rounding errors' in the database. These can be resolved sequentially with the differences in the row and column totals for each region being recalculated after each group of 'errors' is corrected and then the global SAM can be rechecked to ensure the final SAM was exactly balanced. The 'errors' and the means of adjusting them used for GTAP 5, 6, 7 and 8 are detailed below.

i)   No private household demands for imports of dwellings (account *mdwe*) were recorded in some regions. The recorded transactions for the other regions (0.000001) appear to have been included to ensure that some transactions were recorded. The *mdwe* transactions in all regions were therefore set equal to 0.000001.

ii)   The trade accounts were left unchanged and the differences resolved by changing Government final demand transactions so as to remove the differences.

iii)   The row and column totals for the regional household accounts differ and these differences are matched by equal and opposite differences in the row and column totals for the capital accounts. Again, these were minor differences, and they were corrected by adjusting the transfers from the regional households to the private households, government and capital accounts.

A different approach was used for GTAP 9.1; this is detailed in McDonald and Thierfelder (2018)

6   Transactions Values in the Global SAMs

The maximum values for transactions within each regional SAM for versions 5.4 and 9.1 of the Database are reported in Tables 7a and 7b. For each and every region in both versions the minimum absolute value of transactions reported was 0.000001 (ONE dollar US). This indicates that the proportionate differences between maximum and minimum transactions are very substantial, which suggests that for use in GAMS based models there is a *prime facie* case for reducing the 'precision' with which entries in the global SAM are recorded.

Table 7a   **Maximum absolute value of transactions in SAM by region – Version 5.4 of the GTAP Database**

Region	Max TV	Region	Max TV	Region	Max TV
aus	240,438.960	bra	508,587.175	mlt	2,121.080
nzl	41,337.539	chl	50,492.246	pol	87,076.341
chn	414,093.414	ury	14,747.592	rom	26,434.672
hkg	103,586.378	xsm	7,989.730	svk	12,072.915
jpn	2,532,619.344	aut	115,852.484	svn	11,048.583
kor	237,332.856	bel	153,784.285	est	2,848.998
twn	172,051.712	dnk	86,348.500	lva	3,100.144
idn	133,053.769	fin	63,420.055	ltu	5,717.126
mys	46,834.000	fra	834,165.609	rus	271,856.158
phl	59,728.724	deu	1,205,838.644	xsu	81,099.662
sgp	67,359.280	gbr	847,540.978	cyp	5,037.342
tha	84,475.216	grc	87,056.374	tur	129,187.751
vnm	15,468.876	irl	37,969.666	xme	274,229.644
bgd	35,002.404	ita	718,484.621	mar	23,011.554
ind	258,449.980	lux	6,853.892	xnf	113,888.273
lka	10,912.525	nld	215,073.382	bwa	2,319.853
xsa	53,097.764	prt	65,164.010	xsc	82,846.206
can	370,405.880	esp	329,737.164	mwi	2,148.373
usa	5,495,105.851	swe	120,956.802	moz	3,217.104
mex	262,957.181	che	149,266.741	tza	5,919.953
xcm	79,049.589	xef	77,843.997	zmb	3,122.862
col	64,847.745	alb	2,239.777	zwe	5,272.511
per	43,070.491	bgr	9,542.330	xsf	4,609.995
ven	58,484.794	hrv	11,774.252	uga	5,507.291
xap	19,256.664	cze	28,090.545	xss	124,257.099
arg	254,987.329	hun	28,459.323	xrw	192,746.335

Source: Authors calculations from GTAP database.

Table 7b   **Maximum absolute value of transactions in SAM by region – Version 9.1 (2011) of the GTAP Database**

Region	Max TV	Region	Max TV	Region	Max TV	Region	Max TV
aus	746,729.9	ecu	47,228.9	nld	380,773.7	are	238,361.4
nzl	96,624.2	pry	19,432.2	pol	334,889.0	xws	194,231.5
xoc	31,099.2	per	102,592.9	prt	158,905.5	egy	188,578.7
chn	2,658,236.5	ury	33,141.1	svk	57,269.2	mar	62,826.9
hkg	201,722.5	ven	188,471.5	svn	31,258.3	tun	30,522.6
jpn	3,523,850.5	xsm	6,950.4	esp	876,699.5	xnf	85,137.2
kor	634,292.8	cri	23,653.2	swe	260,931.5	ben	9,984.1
mng	4,147.0	gtm	39,886.1	gbr	1,672,573.9	bfa	5,940.0
twn	249,715.0	hnd	12,276.2	che	392,262.5	cmr	17,846.5
xea	23,759.9	nic	8,001.9	nor	210,187.1	civ	16,397.8
brn	8,072.7	pan	32,540.1	xef	12,637.0	gha	27,059.9
khm	10,879.4	slv	21,565.7	alb	10,894.8	gin	5,022.0
idn	486,797.2	xca	865.3	bgr	37,001.8	nga	281,544.1
lao	5,966.9	dom	47,439.9	blr	37,104.5	sen	12,210.7
mys	148,338.3	jam	12,284.6	hrv	38,205.6	tgo	4,783.1
phl	175,354.8	pri	78,758.0	rou	135,015.7	xwf	29,446.1
sgp	107,309.6	tto	11,018.5	rus	943,098.6	xcf	33,189.6
tha	197,974.7	xcb	69,329.1	ukr	115,358.4	xac	63,490.0
vnm	108,191.1	aut	235,229.9	xee	7,341.8	eth	24,737.5
xse	35,658.3	bel	297,907.1	xer	74,835.1	ken	28,287.2
bgd	84,650.1	cyp	18,664.9	kaz	90,242.9	mdg	7,669.9
ind	1,171,619.4	cze	109,450.2	kgz	8,118.3	mwi	4,005.7
npl	16,431.9	dnk	163,285.3	xsu	33,479.8	mus	8,738.2
pak	188,064.2	est	13,175.6	arm	8,374.8	moz	8,640.3
lka	43,429.4	fin	150,341.9	aze	31,776.6	rwa	4,757.5
xsa	22,210.9	fra	1,622,232.8	geo	12,606.3	tza	15,729.4
can	980,348.6	deu	2,104,634.0	bhr	14,728.8	uga	11,308.3
usa	10,887,626.0	grc	233,445.7	irn	283,322.9	zmb	11,845.6
mex	762,745.2	hun	90,690.2	isr	149,643.4	zwe	10,394.4
xna	5,414.5	irl	96,306.2	jor	25,380.0	xec	55,087.6
arg	361,820.3	ita	1,377,400.9	kwt	95,660.3	bwa	6,275.4
bol	16,101.3	lva	20,674.2	omn	35,618.8	nam	6,091.0
bra	1,485,872.5	ltu	31,780.3	qat	108,932.2	zaf	244,631.2
chl	152,149.1	lux	35,983.8	sau	359,159.5	xsc	4,016.4
col	204,446.8	mlt	9,417.1	tur	567,664.9	xtw	102.4

<u>Source</u>: Authors calculations from GTAP database.

7.   Concluding Comments

These programmes convert the GTAP database into a global SAM and that, barring a few very small adjustments that have no modeling implications, the resultant SAM conforms to a standard definition of a SAM. Such a result is hardly surprising; since the data for ALL whole economy models can always be represented in a SAM format provided the models are complete and consistent, which is the case for the GTAP model. Nevertheless the exercise has several substantial benefits. It provides

i)   SAM databases that can be used by SAM based CGE modelers whose preferred programming language is GAMS;

ii)   a method by which individual region SAMs can be easily extracted from the GTAP database, which may make it easier for some researchers to evaluate the data for specific regions; and

iii)   an alternative representation of the GTAP database that might be more easily interpreted by some users of the database.

Since this technical paper was first written (late 2003) there have been several developments.

i)   a GEMPACK version of the code was developed (Kharitonov and Walmsley, 2004);

ii)   the SAM representation of the GTAP database was included in GTAPAgg as part of the GTAP 7 database;

iii)   the documentation of the new version of the GTAP model includes a SAM representation of the GTAP database; and

iv)   various methods for augmenting the GTAP database, especially with respect to current account transactions have been developed (see McDonald and Sonmez, 2004; Walmsley and Minor, 2013; McDonald et al., 2016).

References

GTAP (2002). 'GTAP61.TAB',
      http://www.gtap.agecon.purdue.edu/products/models/current.asp

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

Kharitonov, V.V. and Walmsley, T.L., (2004). 'GEMPACK Code for Creating a Globald
      SAM form GTAP Data', *mimeo*.

King, B.B., (1985). 'What is a SAM?', in Pyatt, G. and Round, J.I. (ed), Social Accounting
      Matrices: A Basis for Planning. Washington: World Bank.

McDonald, S., (2004). 'Developing and Aggregating Reduced Form Social Accounting
      Matrices from the GTAP Database', *mimeo*.

McDonald, S., and Sonmez, Y., (2004). 'Extending a Global GTAP Social Accounting Matrix
      to incorporate Inter-institutional and Inter-Regional Transactions', paper presented at
      the *Seventh Annual Conference on Global Economic Analysis*, World Bank,
      Washington, June.

McDonald, S. and van Schoor, M., (2004). 'An Aggregation Program for a Global SAM from
      GTAP Data', *PROVIDE Technical Paper* (forthcoming).

McDougall, R. and Hagemjer, J., (2016). 'Chapter 13: Income and Factor Taxes', (GTAP 9
      Data Base Documentation)
      (www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=4884)

Pyatt, G. and Round, J.I., (1977). 'Social Accounting Matrices for Development Planning',
      Review of Income and Wealth, Vol 23, pp 339-364.

Pyatt, G., (1991). 'Fundamentals of Social Accounting', Economic Systems Research, Vol 3,
      pp 315-341.

Reinert, K.A. and Roland-Holst, D.W., (1977). 'Social Accounting Matrices' in Francois, J.F.
      and Reinert, K.A., (eds) *Applied Methods for Trade Policy Analysis: A Handbook*.
      Cambridge University Press: Cambridge

Rutherford, T.F., (1998). 'GTAPinGAMS: The Dataset and Static Model', *mimeo*.

Sadoulet, E. and de Janvry, A., (1995). *Quantitative Development Policy Analysis*. John
      Hopkins University Press: Baltimore. Chap 10.

UN, (1993). *System of National Accounts* 1993. UN: New York.

Walmsley, T. and Minor, P., (2013). 'MyGTAP: A Global Model with Single Country
      Flexibilities', (http://mygtap.org/wp-content/uploads/2013/12/Overview.pdf)

Walmsley, T., Winters, A. and Ahmed, S.A., (2007). 'Measuring the Impact of the Movement
      of Labor Using a Model of Bilateral Migration Flows' *GTAP Technical Paper* No.
      28 (www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=2529)

Appendix 1 Programme Instructions

<u>Overview</u>

The GAMS code for transforming the HAR file representation of the GTAP database into a SAM representation saved in GDX format are provided as a GAMS User Model library. There are currently versions of the programme for GTAP 5.3 to 9.1 in the User model library.[30]

All the files are in one WinZip archive – GTAPSAM.zip – together with the library programme gtaplib.glb.

<u>Software Requirements</u>

1. This programme is known to work with versions of GAMS 250 after and including distribution 24. It is believed to work with all versions of GAMS 250 that include the GDXXRW utilities, i.e., since distribution 21.0 in May 2003. Users do NOT need a GAMS license to use the programmes, i.e., it runs with the DEMO/Student versions of GAMS

<u>Data Requirements</u>

1. No GTAP data are provided. The user must have a licensed version of the GTAP database complete with the GTAPAgg programme.

2. All the sets and set mapping required to use the GAMS programmes for specific versions of the GTAP database are provided from a single Excel workbook.

3. An additional Excel workbook is also supplied that provides are simple semi-automated method for producing new set and mapping configurations.

<u>User Instructions</u>

(These instructions assume the user is working in a Windows 2000/XP environment using GAMSIDE and Microsoft Excel 2000/WP or later.)

1. The programme files are made available in a WinZip archive - gtapsam.zip.

2. The programme files are organised so that they can be used as a GAMS User Model Library (see the GAMS documentation for information on using User

---

[30] The original versions of the code used the SEEHAR utility to render the HAR data 'visible' to GAMS. The GDX2HAR and HAR2GDX utilities are more efficient.

Model Libraries). These instruction assume that files in gtapsim.zip have been extracted to a directory - gsamlib.

3. Create a directory for the programme. The path to this directory should follow standard/old DOS conventions - no directory with more than an 8 character name and no spaces in the names in the path.

   Create a project file (*.gpr) in this directory.

4. To copy the programme files to the project directory.

   In GAMSIDE with the project file identified do the following

   a) Locate the User Model Library - File/Model Library/Open User Model Library and then select the file gtablib.glb in the directory gsamlib

   b) Double click on the NAME of the version of the programme that corresponds to the version of the GTAP database with which you are working. ALL the files needed to run the programme other than the GTAP database file will be copied to the directory containing the project file.

   NB: A brief description of the programme is provided in the window at the bottom of the User Mode Library window.

5. Read and study the GAMS programme.

6. Create the database for the programme

   a) Use GTAPAgg to create an aggregation with 1:1 mappings for sectors, regions and factors.

   b) Extract the file Basedata.har from the WinZip archive created by GTAPAgg and copy to the project directory.

7. Check that programme compiles by using SHIFT+F9.

   If the steps above have been followed there should be no problems.

   a) Some users can have problems with the GDXXRW facility that generates the file data_in.gdx from the Excel file. These problems all seem to originate with a corruption of the 'LAYOUT' worksheet and result in an error message saying the file data_in.gdx does not exist. To overcome this type of problem

      i)   check that gdxxrw.exe exists in the GAMS directory

      ii)   copy the contents of the worksheet 'LAYOUT2' and replace the worksheet 'LAYOUT' using Paste Special..

8. Run the programme. The output to the working directory should be as follows:

   a) A series of GDX files for checking the intermediate stages

      gtptots1.gdx

      gtptots2.gdx

      gtptots3.gdx

      gtptots4.gdx

      gtptots5.gdx

      gtpmxmin.gdx

   b) A Global SAM in GDX format

      e.g., gtpsam60_p5.gdx

   c) A Global SAM in HAR format

      e.g., gsam60p5.har

9. Exports *fob* and Imports *fob* calculated directly do not always equate - indirect and direct methods of calculating imports are included search for TRADECAL to control the choice of method. All the checks labelled TRADECHK should equal ZERO.

10. GAMS code for aggregating the Global SAM programme is provided in the User Model Library.

11. These notes relate to situations where a user wishes to work with another version of the GTAP database. It is assumed this will be a later version.

    A new workbook - data_gtp_**.xls - will need to be compiled.

    a) Using the file gtp_sets_v60.xls or later as a template

       i)   Get the 'new' sector, region, factor and margin codes and descriptions from the GTAP database and copy them into the worksheet 'Set data from GTAPAgg'

       ii)   Check that the worksheets 'Set creation regions/sectors/factor/margins' have all updated correctly - NB the number of entries in each group of codes is the critical point.

       iii)   Compile the new sets in the worksheet 'Compiling Sets'

    b) Using the file data_gtp_60.xls or later as a template

       i)   Generate a revise worksheet 'Sets' - using copy and paste SPECIAL from the new file gtp_sets_V**.xls.

ii)   Generate a revise worksheet 'Maps' - using copy and paste from the new worksheet 'Sets'.

iii)   The worksheet 'Layout' should NOT need to be changed unless new sets or maps are created.

c) Update the programme file to read the new sets and maps.

d) Some general suggestions

i)   Spaces in the cells of the file data_gtp_**.xls are a problem - to GAMS spaces are characters.

ii)   If the orderings in the mapping sets are not consistent the mappings implemented will be wrong.

iii)   'Corruption' of the worksheet 'Layout' will cause the sets maps not to be read in.

iv)   The codes used for account names in the GTAP database may have changed.

v)   'Big' initial 'errors' (say entries in the first DIFFTOT that are greater than 1) are likely to be due to an error in the sets and maps files.

vi)   'Small' initial 'errors' (say less that 0.01) are more likely to rounding issues due to the degree of precision in the GTAP database.

vii)   When adjusting the programme to remove 'errors' due to the degree precision in the database, it is best by far to work through the accounts as they appear in the SAM, i.e., start with the import commodity accounts, then domestic commodities, etc.

Please note that the authors are NOT in a position to provide unconditional support for users of these programmes although they are interested in receiving well documented bug reports and will try to help within reasonable bounds.

Warning

Be warned that the files created when using these programmes can be large, although with development of GDX, for GAMS, and HAR, for GEMPACK, they are smaller now (2018) than there were in 2004.

Depending on the DISPLAY options chosen the *.lst file can exceed 100MB.

The programmes are not fast.

Appendix 2 Extract from the Programme

```
*############## 8. CREATING THE GLOBAL SAM
#########################
* Submatrices are added by rows of the SAM
* Number in brackets (x:y) indicate row (x) and column (y) in the SAM

*#### Import commodity income components

* Import Intermediate Use matrices (1:3)

 SAMG(mcc,aa,kk)        = SUM((trad_comm,prod_comm)
                           $(MAPIMP(mcc,trad_comm)
                           $MAPACT(aa,prod_comm)),
                           VIFM(trad_comm,prod_comm,kk)) ;

* Household import demand vectors (1:6)

 SAMG(mcc,"hhous",kk)   = SUM(trad_comm
                           $MAPIMP(mcc,trad_comm),
                           VIPM(trad_comm,kk)) ;

* Government import demand vectors (1:13)

 SAMG(mcc,"ggovt",kk)   = SUM(trad_comm
                           $MAPIMP(mcc,trad_comm),
                           VIGM(trad_comm,kk)) ;

* Investment import demand vectors (1:14)

 SAMG(mcc,ii,kk)        = SUM((trad_comm,prod_comm)
                           $(MAPIMP(mcc,trad_comm)
                           $MAPINV(ii,prod_comm)),
                           VIFM(trad_comm,prod_comm,kk)) ;

*#### Domestic commodity income components

* Domestic Intermediate Use matrices (2:3)

 SAMG(dcc,aa,kk)        = SUM((trad_comm,prod_comm)
                           $(MAPDOM(dcc,trad_comm)
                           $MAPACT(aa,prod_comm)),
                           VDFM(trad_comm,prod_comm,kk)) ;

* Household domestic demand vectors (2:6)

 SAMG(dcc,"hhous",kk)   = SUM(trad_comm
                           $MAPDOM(dcc,trad_comm),
                           VDPM(trad_comm,kk)) ;
```

```
* Export taxes

 SAMG(teetax,dcc,kk)    = SUM((trad_comm,kkp)
                           $(MAPDOM(dcc,trad_comm)
                           $MAPTETAX(teetax,kkp)),
                           VXWD(trad_comm,kk,kkp))
                              - SUM((trad_comm,kkp)
                                 $(MAPDOM(dcc,trad_comm)
                                 $MAPTETAX(teetax,kkp)),
                                 VXMD(trad_comm,kk,kkp)) ;

* Government domestic demand vectors (2:13)

 SAMG(dcc,"ggovt",kk)   = SUM(trad_comm
                           $MAPDOM(dcc,trad_comm),
                           VDGM(trad_comm,kk)) ;

* Investment domestic demand vectors (2:14)

 SAMG(dcc,ii,kk)        = SUM((trad_comm,prod_comm)
                           $(MAPDOM(dcc,trad_comm)
                           $MAPINV(ii,prod_comm)),
                           VDFM(trad_comm,prod_comm,kk)) ;

* Export demand matrices (2:16)
* Export of transport services

 SAMG("ddotp","otppvst",kk)    = VST1("ddotp",kk) ;
 SAMG("ddwtp","wtppvst",kk)    = VST1("ddwtp",kk) ;
 SAMG("ddatp","atppvst",kk)    = VST1("ddatp",kk) ;

* Exports of Goods and Services

 SAMG(dcc,ww,kk)        = SUM((trad_comm,kkp)
                           $(MAPDOM(dcc,trad_comm)
                           $MAPTRAD(ww,kkp)),
                           VXWD(trad_comm,kk,kkp)) ;

*#### Activity Income

* Make Matrix (3:2)

 SAMG(aa,dcc,kk)    = (VDM(dcc,kk) + VST1(dcc,kk)
                           + SUM(ww,SUM((trad_comm,kkp)
                              $(MAPDOM(dcc,trad_comm)
                              $MAPTRAD(ww,kkp)),
                              VXMD(trad_comm,kk,kkp))))
                              * ID(aa,dcc) ;
```

```
*#### Factor Income                                          * Government Demand (9:13)

* Factor demands by activities (4:3)                          SAMG(smmtax,"ggovt",kk) = SUM(mcc
 SAMG(ff,aa,kk)           = SUM((endw_comm,prod_comm)                           $MAPSMTAX(smmtax,mcc),
                           $(MAPFACT(ff,endw_comm)                              SUM(trad_comm
                           $MAPACT(aa,prod_comm)),                             $MAPIMP(mcc,trad_comm),
                           VFM(endw_comm,prod_comm,kk)) ;                       VIGA(trad_comm,kk)))
*#### Indirect Tax Receipts for INDTAX(r) in GTAP model                          - SUM(mcc
                                                                                 $MAPSMTAX(smmtax,mcc),
* Import Duty Income (7:1)                                                         SUM(trad_comm
                                                                                  $MAPIMP(mcc,trad_comm),
 SAMG(tmmtax,mcc,kk)      = SUM((trad_comm,kkp)                                    VIGM(trad_comm,kk))) ;
                           $(MAPIMP(mcc,trad_comm)
                           $MAPTMTAX(tmmtax,kkp)),            * Investment Demand (9:14)
                           VIMS(trad_comm,kkp,kk))
                               - SUM((trad_comm,kkp)          SAMG(smmtax,ii,kk)      = SUM(mcc
                               $(MAPIMP(mcc,trad_comm)                           $MAPSMTAX(smmtax,mcc),
                               $MAPTMTAX(tmmtax,kkp)),                          SUM((trad_comm,prod_comm)
                               VIWS(trad_comm,kkp,kk)) ;                        $(MAPIMP(mcc,trad_comm)
                                                                               $MAPINV(ii,prod_comm)),
*# Import Sales Taxes                                                           VIFA(trad_comm,prod_comm,kk)))
                                                                                   - SUM(mcc
* Intermediate Inputs (8:3)                                                        $MAPSMTAX(smmtax,mcc),
 SAMG(smmtax,aa,kk)      = SUM(mcc                                                  SUM((trad_comm,prod_comm)
                           $MAPSMTAX(smmtax,mcc),                                  $(MAPIMP(mcc,trad_comm)
                           SUM((trad_comm,prod_comm)          *# Domestic Sales Taxes                $MAPINV(ii,prod_comm)),
                           $(MAPIMP(mcc,trad_comm)                                  VIFM(trad_comm,prod_comm,kk))) ;
                           $MAPACT(aa,prod_comm)),            * Intermediate Inputs (8:3)
                           VIFA(trad_comm,prod_comm,kk)))
                               - SUM(mcc                       SAMG(sddtax,aa,kk)      = SUM(dcc
                               $MAPSMTAX(smmtax,mcc),                            $MAPSDTAX(sddtax,dcc),
                               SUM((trad_comm,prod_comm)                        SUM((trad_comm,prod_comm)
                               $(MAPIMP(mcc,trad_comm)                          $(MAPDOM(dcc,trad_comm)
                               $MAPACT(aa,prod_comm)),                          $MAPACT(aa,prod_comm)),
                               VIFM(trad_comm,prod_comm,kk))) ;                 VDFA(trad_comm,prod_comm,kk)))
                                                                                   - SUM(dcc
* Household Demand (9:6)                                                            $MAPSDTAX(sddtax,dcc),
                                                                                   SUM((trad_comm,prod_comm)
 SAMG(smmtax,"hhous",kk) = SUM(mcc                                                 $(MAPDOM(dcc,trad_comm)
                           $MAPSMTAX(smmtax,mcc),                                  $MAPACT(aa,prod_comm)),
                           SUM(trad_comm                                           VDFM(trad_comm,prod_comm,kk))) ;
                           $MAPIMP(mcc,trad_comm),
                           VIPA(trad_comm,kk)))              * Household Demand (8:6)
                               - SUM(mcc
                               $MAPSMTAX(smmtax,mcc),          SAMG(sddtax,"hhous",kk) = SUM(dcc
                               SUM(trad_comm                                     $MAPSDTAX(sddtax,dcc),
                               $MAPIMP(mcc,trad_comm),                           SUM(trad_comm
                               VIPM(trad_comm,kk))) ;                            $MAPDOM(dcc,trad_comm),
```

```
                    VDPA(trad_comm,kk)))                                                    $(MAPIMP(mcc,trad_comm)
                      - SUM(dcc                                                              $MAPACT(aa,prod_comm)),
                        $MAPSDTAX(sddtax,dcc),                                               VIFA(trad_comm,prod_comm,kk)))
                          SUM(trad_comm                                        + SUM(dcc,SUM((trad_comm,prod_comm)
                          $MAPDOM(dcc,trad_comm),                                            $(MAPDOM(dcc,trad_comm)
                          VDPM(trad_comm,kk))) ;                                              $MAPACT(aa,prod_comm)),
                                                                                             VDFA(trad_comm,prod_comm,kk)))
* Government Demand (8:13)                                                     + SUM(ff,SUM((endw_comm,prod_comm)
                                                                                             $(MAPFACT(ff,endw_comm)
 SAMG(sddtax,"ggovt",kk) = SUM(dcc                                                            $MAPACT(aa,prod_comm)),
                          $MAPSDTAX(sddtax,dcc),                                             EVFA(endw_comm,prod_comm,kk))) )
                          SUM(trad_comm                                          ;
                          $MAPDOM(dcc,trad_comm),
                          VDGA(trad_comm,kk)))
                      - SUM(dcc                                   * Factor Income Taxes (12:4)
                        $MAPSDTAX(sddtax,dcc),
                          SUM(trad_comm                            SAMG("ddirtax",ff,kk)   = SUM(aa,SUM((endw_comm,prod_comm)
                          $MAPDOM(dcc,trad_comm),                                            $(MAPFACT(ff,endw_comm)
                          VDGM(trad_comm,kk))) ;                                             $MAPACT(aa,prod_comm)),
                                                                                            VFM(endw_comm,prod_comm,kk)))
* Investment Demand (8:14)                                                    - SUM(endw_comm
                                                                                 $MAPFACT(ff,endw_comm),
 SAMG(sddtax,ii,kk)      = SUM(dcc                                                EVOA(endw_comm,kk)) ;
                          $MAPSDTAX(sddtax,dcc),
                          SUM((trad_comm,prod_comm)                 *#### Components of Regional Household Income
                          $(MAPDOM(dcc,trad_comm)
                          $MAPINV(ii,prod_comm)),                   * Factor Incomes after depreciation (5:4)
                          VDFA(trad_comm,prod_comm,kk)))
                      - SUM(dcc                                     SAMG("rreghous",ff,kk)        = SUM(endw_comm
                        $MAPSDTAX(sddtax,dcc),                                                     $MAPFACT(ff,endw_comm),
                          SUM((trad_comm,prod_comm)                                                EVOA(endw_comm,kk))
                          $(MAPDOM(dcc,trad_comm)                                                      - DEPREC(ff,kk) ;
                          $MAPINV(ii,prod_comm)),
                          VDFM(trad_comm,prod_comm,kk))) ;         * Import duties and export taxes (5:7)

* Factor taxes paid by activities                                  SAMG("rreghous",tmmtax,kk)  = SUM(mcc,SAMG(tmmtax,mcc,kk)) ;

* Factor taxes paid by activities                                  SAMG("rreghous",teetax,kk)  = SUM(dcc,SAMG(teetax,dcc,kk)) ;

 SAMG(fftax,aa,kk)          = SUM((endw_comm,prod_comm)             * Domestic Sales taxes (5:8)
                          $(MAPFTAX(fftax,endw_comm)
                          $MAPACT(aa,prod_comm)),                   SAMG("rreghous",sddtax,kk)  = SUM(aa, SAMG(sddtax,aa,kk))
                          EVFA(endw_comm,prod_comm,kk))                                          + SAMG(sddtax,"hhous",kk)
                          - SUM((endw_comm,prod_comm)                                            + SAMG(sddtax,"ggovt",kk)
                          $(MAPFTAX(fftax,endw_comm)                                             + SAMG(sddtax,"kkap",kk) ;
                          $MAPACT(aa,prod_comm)),
                          VFM(endw_comm,prod_comm,kk)) ;           * Import Sales taxes (5:9)

* Production taxes paid by activity                                SAMG("rreghous",smmtax,kk)  = SUM(aa,SAMG(smmtax,aa,kk))
                                                                                                + SAMG(smmtax,"hhous",kk)
 SAMG("pprodtax",aa,kk)  = SUM(dcc,SAMG(aa,dcc,kk)) -                                            + SAMG(smmtax,"ggovt",kk)
                          (SUM(mcc,SUM((trad_comm,prod_comm)                                     + SAMG(smmtax,"kkap",kk) ;
```

```
* Production Taxes (5:11)

 SAMG("rreghous","pprodtax",kk)
                  = SUM(aa,SAMG("pprodtax",aa,kk)) ;

* Factor Use Taxes

 SAMG("rreghous",fftax,kk)  =  SUM(aa,SAMG(fftax,aa,kk)) ;

* Direct Taxes (5:12)
 SAMG("rreghous","ddirtax",kk)  =  SUM(ff,SAMG("ddirtax",ff,kk)) ;

*#### Components of Regional Household Expenditure

* Income of Private Households (6:5)

 SAMG("hhous","RREGHOUS",kk) = SUM(mcc,SAMG(mcc,"hhous",kk))
                     + SUM(smmtax,SAMG(smmtax,"hhous",kk))
                     + SUM(dcc,SAMG(dcc,"hhous",kk))
                     + SUM(sddtax,SAMG(sddtax,"hhous",kk))
;

* Income of Government (13:5)
 SAMG("ggovt","RREGHOUS",kk) = SUM(mcc,SAMG(mcc,"ggovt",kk))
                     + SUM(smmtax,SAMG(smmtax,"ggovt",kk))
                     + SUM(dcc,SAMG(dcc,"ggovt",kk))
                     + SUM(sddtax,SAMG(sddtax,"ggovt",kk))
;

* Domestic Savings (14:5)

 SAMG("kkap","RREGHOUS",kk)  = SAVE(kk) ;

*#### Capital Account Income

* Depreciation charges (14:4)

 SAMG("kkap","ffcapital",kk) = VDEP(kk) ;

*#### Rest of World

* Import Margin Matrices (15:1)

 SAMG(OTPPMARG,mcc,kk)    = SUM((trad_comm,kkp)
                   $(MAPIMP(mcc,trad_comm)
                   $MAPOTPMARG(otppmarg,kkp)),
                   VTWR("otp",trad_comm,kkp,kk)) ;

 SAMG(WTPPMARG,mcc,kk)    = SUM((trad_comm,kkp)
                   $(MAPIMP(mcc,trad_comm)
                   $MAPWTPMARG(wtppmarg,kkp)),
```

```
                   VTWR("wtp",trad_comm,kkp,kk)) ;

 SAMG(ATPPMARG,mcc,kk)    = SUM((trad_comm,kkp)
                   $(MAPIMP(mcc,trad_comm)
                   $MAPATPMARG(atppmarg,kkp)),
                   VTWR("atp",trad_comm,kkp,kk)) ;

* Import Supply Matrices (16:1)

* SMcD TRADE

* Direct method
$ontext
 SAMG(ww,mcc,kk)          = SUM((trad_comm,kkp)
                     $(MAPIMP(mcc,trad_comm)
                     $MAPTRAD(ww,kkp)),
                   VIWS(trad_comm,kkp,kk))
                     - SUM(marg_comm,SUM((trad_comm,kkp)
                       $(MAPIMP(mcc,trad_comm)
                       $MAPTRAD(ww,kkp)),
                       VTWR(marg_comm,trad_comm,kkp,kk)))
;
$offtext

* Indirect Method using Exports fob

Parameter
 SAMExports1(mcc,ww,rr)    Exports of commod mcc to region ww by
region rr
 SAMExports2(mcc,kk,rr)    Exports of commod mcc to region kk by
region rr
 SAMExports3(mcc,kk,ww)    Exports of commod mcc to region kk by
region ww

 ;

 SAMExports1(mcc,ww,rr) = SUM((dcc,kk)
                     $(MAPCOM(mcc,dcc)
                     $MAPREG(rr,kk)),
                     SAMG(dcc,ww,kk)) ;

 SAMExports2(mcc,kk,rr)  = SUM(ww
                     $MAPREG4(kk,ww),
                     SAMExports1(mcc,ww,rr)) ;

 SAMExports3(mcc,kk,ww)  = SUM(rr
                     $MAPREG5(ww,rr),
                     SAMExports2(mcc,kk,rr)) ;

 SAMG(ww,mcc,kk)          = SAMExports3(mcc,kk,ww) ;

* Expenditures on transport services
```

```
 SAMG("otppvst",otppmarg,kk)  = SUM(mcc,SAMG(otppmarg,mcc,kk)) ;
 SAMG("wtppvst",wtppmarg,kk)  = SUM(mcc,SAMG(wtppmarg,mcc,kk)) ;
 SAMG("atppvst",atppmarg,kk)  = SUM(mcc,SAMG(atppmarg,mcc,kk)) ;


* Capital Account Balances
* Capital Account balance for trade in transport services

 SAMG("kkap","otppvst",kk)  =
SUM((mcc,otppmarg),SAMG(otppmarg,mcc,kk))
                              - SUM(dcc,SAMG(dcc,"otppvst",kk)) ;

 SAMG("kkap","wtppvst",kk)  =
SUM((mcc,wtppmarg),SAMG(wtppmarg,mcc,kk))
                              - SUM(dcc,SAMG(dcc,"wtppvst",kk)) ;

 SAMG("kkap","atppvst",kk)  =
SUM((mcc,atppmarg),SAMG(atppmarg,mcc,kk))
                              - SUM(dcc,SAMG(dcc,"atppvst",kk)) ;

* Capital Account balances by trade partner on commodity trade

 SAMG("kkap",ww,kk)          = SUM(mcc,SAMG(ww,mcc,kk))
                              - SUM(dcc,SUM((trad_comm,kkp)
                                    $(MAPDOM(dcc,trad_comm)
                                    $MAPTRAD(ww,kkp)),
                                    VXWD(trad_comm,kk,kkp))) ;
```