

Aggregating GTAP SAMs: SAMGator

Scott McDonald

Abstract

This paper is a user's guide to the SAMGator programme as applied to the GTAP database in SAM format. SAMGator is a Visual Basic for Applications programme that was developed as part of the PROVIDE project; the coding was done by Melt van Schoor. Variants of the SAMGator programme can be used to aggregate multiple different SAMs.

Address for correspondence:

Scott McDonald,
Department of Economics & International Business,
Oxford Brookes University,
Wheatley Campus,
Oxford, OX331HX, UK.
Email: smcdonald@brookes.ac.uk
Tel: +44 1865948

1. Introduction

The dimensions of the GTAP database require that for practical purposes the database must be aggregated before being used to calibrate a global CGE model. There are two main reasons for this:

1. If the database is not aggregated the resultant model would be too large to
 - a. solve within realistic limits of computer capacity, OR
 - b. yield results that would be susceptible to practical analyses and interpretation.
2. It is argued that models should be focused upon specific issues rather than being general and that part of the process of focus depends upon the identification of an appropriate aggregation for the purposes at hand.

An aggregation of the GTAP SAM requires that the database is (simultaneously) aggregated in three dimensions, i.e., the intra regional accounts for all regions are aggregated to reduce the number commodities/activities, factors, trade partners and associated transaction, and inter regional transactions are aggregated so as to accord with the aggregation of the trade accounts for each region. GAMS is an extremely efficient medium for implementing the calculations required to aggregate the GTAP SAM, but setting up the sets and (set) mappings to control the aggregation is potentially time consuming. One method for reducing the time costs involved in setting up a new aggregation, and simultaneously reducing the time required to configure the sets and other data is to use the SAMgator software (PROVIDE, 2004).

SAMgator is a Visual Basic for Applications (VBA) programme implemented using Microsoft (MS) Excel. The user declares, describes and defines the new (aggregated) sectors (commodities and activities), factors and regions in an Excel template, identifies the source and destination data files and then runs GAMS remotely from within the Excel programme¹. The SAMgator programme generates the set and mapping files, checks to ensure that all the mappings are 'legal' and then generates the GAMS programme file, which is then initiated from within Excel.

¹ The GAMS code at the heart of SAMgator is report in the Appendices.

Moreover with the same mapping sets can be used to aggregate other components of the GTAP database so that they are consistent with the transactions (SAM) data, e.g., the elasticities that are part of the GTAP database.

2 The Mechanics of SAMgator

At the heart of SAMgator is a single (GAMS) equation that aggregates the GTAP database in three dimensions (see below).

$$\begin{aligned} \text{NEWSAM}(sp, spp, k) &= \text{SUM}((ss, ssp, kk) \\ &\quad \$(\text{MAPSAMAG}(sp, ss) \\ &\quad \$\text{MAPSAMAG}(spp, ssp) \\ &\quad \$\text{MAPREGAG}(k, kk)), \\ &\quad \text{SAMG}(ss, ssp, kk)) ; \end{aligned}$$

The parameter $\text{SAMG}(ss, ssp, kk)$ contains the GTAP database where ss^2 is the set that defines the row and columns labels for the SAM for each of the kk regions. The parameter $\text{NEWSAM}(sp, spp, k)$ contains the GTAP database where s is the set that defines the row and columns labels for the aggregated (NEW)SAM for each of the k regions. The set $\text{MAPSAMAG}(sp, ss)$ defines the members of ss that aggregated into sp by the rows of $\text{SAMG}(ss, ssp, kk)$, while the set $\text{MAPSAMAG}(spp, ssp)$ defines the members of ssp that aggregated into ssp by the columns of $\text{SAMG}(ss, ssp, kk)$. (Note that there is only one mapping set MAPSAMAG with its implementation being solely driven by the sets ss and s and their aliases.) The set $\text{MAPREGAG}(k, kk)$ defines the aggregation of regions kk into k for the aggregated (NEW)SAM for each of the k regions.

2.1 Using SAMgator

The user interface to SAMgator consists of two worksheets ‘Control’ and ‘MappingConfig’. A user only needs to access these two worksheets; although more experienced users may wish to use other features of SAMgator the discussion here is limited to these two worksheets.

² The set ssp is an alias for ss .

2.1.1 Control

The ‘Control’ worksheet is used to identify the location of GAMS.EXE, input and output file names, check for errors, write output and to run the aggregation programme. A screen shot of the worksheet is given in Fig2.1.1.1.

In order to run GAMS from within MS Excel it is necessary for Excel to know the location of the GAMS.EXE file; this information is provided in the ‘GAMS.EXE location’ box.³ In order to test that Excel can implement a GAMS programme first click on the ‘Pause after execution’ box – so it has a tick in the box as in Figure 2.1.1.1 – and then click on the ‘Test GAMS’ button. A DOS window will appear (Figure 2.1.1.2); if this indicates a normal completion this is confirmation that Excel can trigger GAMS correctly and the user can press any key to continue.

³ With some versions of GAMS spaces in the paths can lead to problems. Note how the screen shot indicates that GAMS.exe is stored in a GAMS directory on the C drive and not in the ‘Program Files’ that is the default destination when installing GAMS. This has been found to be time efficient since it avoids resetting the path every time a new version of GAMS is installed.

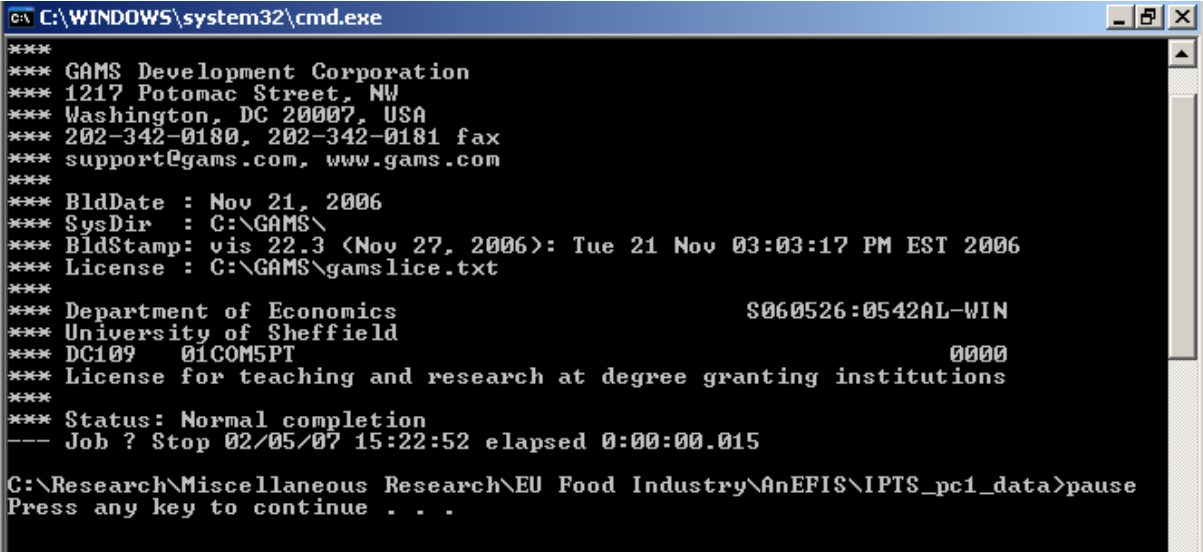
Figure 2.1.1.1 Control Worksheet for SAMgator⁴

	A	B	C	D	E	F	G	H	I	J	K	
	<div style="text-align: center;"> Aggregation program for global SAM from GTAP data <i>This version also generates weighted averages of the GTAP elasticities ESUBD, ESUBM and ESUBVA. A correctly formulated input database is required.</i> </div>											
Options:												
GAMS.EXE location:	C:\GAMS\GAMS.EXE							<div>Test GAMS</div>				
Pause after execution:	<input checked="" type="checkbox"/>											
	.GDX file name											
Input	worsam60_v3g.gdx											
Output	samg23_5_18_ipts.gdx											
Aggregation program (.gms):	globaggr_ipts1.gms											
	<div>Write output file</div>					<div>Write to Sets sheet</div>						
	<div>Check for errors</div>											
	<div>Run aggregation</div>											

The user should now define the input file, which is required to be a GAMS Data Exchange (GDX) file; in this case the file is called 'worsam60_v3g.gdx'. Then name the output file; in this case 'samg23_5_18_ipts.gdx', where the convention adopted is 'samgC_F_R*.gdx' where *C* is the number of commodities and activities, *F* the number of factors and *R* the number of regions. And finally the user should provide a name for the GAMS programme file; in this case 'globaggr_ipts1.gms'. The programme is not sensitive to the choice of naming conventions.

⁴ The precise version of SAMGator may differ slightly from the screenshots used in this document; this reflects on going developments and/or customisation.

Figure 2.1.1.2 Test GAMS



```

C:\WINDOWS\system32\cmd.exe
***
*** GAMS Development Corporation
*** 1217 Potomac Street, NW
*** Washington, DC 20007, USA
*** 202-342-0180, 202-342-0181 fax
*** support@gams.com, www.gams.com
***
*** BldDate : Nov 21, 2006
*** SysDir : C:\GAMS\
*** BldStamp: vis 22.3 (Nov 27, 2006): Tue 21 Nov 03:03:17 PM EST 2006
*** License : C:\GAMS\gamslice.txt
***
*** Department of Economics                      S060526:0542AL-WIN
*** University of Sheffield
*** DC109 01COM5PT                                0000
*** License for teaching and research at degree granting institutions
***
*** Status: Normal completion
--- Job ? Stop 02/05/07 15:22:52 elapsed 0:00:00.015

C:\Research\Miscellaneous Research\EU Food Industry\AnEFIS\IPTS_pc1_data>pause
Press any key to continue . . .

```

2.1.2 MappingConfig

The role of the 'MappingConfig' worksheet is to generate the sets for the aggregated SAM and set up the mappings, which will control the aggregation. The process has a number of safeguards built in to ensure there the mappings are legitimate and that there are no errors. Although users may choose to adapt a previous aggregation the description given here assumes that the user starts with a version of SAMgator that contains no aggregation specific set or mapping data.

Figure 2.1.2.1 illustrates how the 'MappingConfig' worksheet might look when there is no aggregation specific set or mapping data; the only difference between this image and the one seen when opening SAMgator is that the 'int_k – Regions' line in the 'Category' box (top left) has been highlighted. The user needs to create sets and mappings for the commodities/activities, factors and regions.

Figure 2.1.2.1 Mapping and Configuration Worksheet 1

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
													Check for Errors	
Category:			Elements in output set for this category:											
<div> int_c - Commodities-Activiti int_f - Factors int_k - Regions vstmarg - export margin acc </div>													Delete element Move Up Move Dn Edit Auto Sort	
													Name: <input type="text"/> Add Descr: <input type="text"/>	
Elements in input set mapped to selected element in output set:			Unmapped elements in input set:											
			<div> <div><<<</div> <div><</div> <div>></div> <div>>>></div> <div> aus - Australia nzl - New Zealand xoc - Rest of Oceania chn - China hkg - Hong Kong jpn - Japan kor - Korea twn - Taiwan xea - Rest of East Asia idn - Indonesia mys - Malaysia phl - Philippines sgp - Singapore tha - Thailand vnm - Vietnam xse - Rest of Southeast Asia </div> </div>											

The steps undertaken to declare and define aggregate sets and to generate the mapping file are as follows.

1. Select a category: choose commodities/activities, factors or regions in the 'Category' box (top left).
2. Declare (and describe) an aggregate set element: type the name in the 'Name' box and below it a description in the 'Descr' box and then click the 'Add' button. This element will be a member of an output set for which each element will appear with its description in the 'Elements in output set for this category' box.
 - Note that SAMgator will use the sets and mappings to generate automatically multiple sets by PREFIXING letters (between 1 and 5 characters) to each element of the output set.
3. Select elements in the input set that are to be mapped to the selected member of the output set: select an element in the 'Elements in output set for this category' box and then select those elements that are to be mapped to that output set element from the 'Unmapped elements of input set' box (bottom right) by clicking on each in turn (NB the SHIFT and Ctrl click options do not apply). Click on the <- button to define the mapping. Note how the selected element of

the input set are moved to the 'Elements in input set mapped to selected element in output set' box (bottom left).

- The <- button moves selected elements in the unmapped element box to the mapped elements box; the -> button moves selected elements in the mapped element box to the unmapped elements box; the <<< button moves all elements in the unmapped element box to the mapped elements box; >>> button moves all elements in the mapped element box to the unmapped elements box.
 - Figure 2.1.2.2 illustrates that for the category 'int_c – Commodities-Activities' the elements 'pdr – paddy rice', 'wht – wheat' and 'gro – cereal grains nec' of the input set are mapped to the element 'gran – grains' of the output set.
4. Continue adding elements to the output set in each category until all the elements in the input set are mapped to elements in the output set. Note how each element in the input set can only be mapped to one element in the output set – thereby avoiding 'double counting'. The error checking also checks to avoid duplicates in the output set.
 5. Ordering of the output set: as each new element of an output set is declared it is added to the end of the list of members if the output set for that category and this defines the set ordering that will be used by GAMS. The order of elements within an output set can be altered by selecting elements in the 'Elements in output set for this category' box and using the 'Move Up' and/or 'Move Down' buttons.⁵
 6. Deleting an output set element: select the element to delete in 'Elements in output set for this category' box and click on the 'Delete element' button.
 7. Select another category and complete steps 2 to 7 as appropriate until output sets and mapping sets for all three categories have been generated.

⁵ Note that GAMS lists set members in the order that they were declared in the programme. If the same name, say 'abc', is used in 2 different sets this can create ordering problems for the user: assume that the order wanted for reports is the order of members in the second set declared, but because 'abc' is in the first set declared it will be the first member of the second set. Use names for members of sets that are unique to the sets to which they belong.

Figure 2.1.2.2 Mapping and Configuration Worksheet 2

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
													Check for Errors	

Category:

int_c - Commodities/Activities
int_f - Factors
int_k - Regions
vstmargin - export margin accounts

Elements in output set for this category:

gran - Grains
scb - Sugar cane and beet
ocrp - Other crops
pbf - Plant based fibres
lstk - Livestock
milk - Raw milk

Delete element Move Up Move Dn
Edit Auto Sort

Name: Add
Descr:

Elements in input set mapped to selected element in output set:

pdr - Paddy rice
wht - Wheat
gro - Cereal grains nec

<<<
<-
->
>>>

Unmapped elements in input set:

2.1.3 Implementing an Aggregation

Once the aggregation is defined in the *MappingConfig* worksheet return to the *Control* worksheet. There are then five steps to follow:

1. Check that the input and output GDX files are correctly specified and that the aggregation program name is as wished.
2. Click on the *Write Output* button.
3. Click on the *Write Set Sheet* button.
4. Click on the *Check for Errors* button – if there is an error in the mappings associated solely with the processes implemented by the VBA programme a box will appear that indicates the nature of the error.
5. Click on the *Run Aggregation* button.

If everything has been done correctly a DOS box appears indicating the progress of the GAMS execution of the aggregation programme. If there are any errors these will cause the programme to fail to return a normal completion, i.e., there will have been some syntax or execution error in GAMS OR the programme will have aborted due to an imbalance in the

SAM. The easiest way to work out the source of these errors is to re run the `globaggr**.gms` programme and examine the output.

Typical errors are

1. The programme aborts because the row and column totals of the aggregated SAM are not identical. There is a conditional abort statement that is set to require a maximum difference between each and every row and column total of 0.00001. There are two usual reasons for this ‘error’: (a) the row and column totals of the disaggregated SAM are not identical, and (b) there are errors in the sets and hence mappings.
2. Syntax errors usually arise because the set names expected by SAMGator are not those used in the disaggregated SAM.
 - a. If the user has used the set names specified by the programme that generates a SAM from the GTAP database (McDonald and Thierfelder, 2004) this should not happen, if not then the ‘InputSet’ worksheet needs changing to accord with the sets in the source data.
 - b. There are sometimes slight differences in the 3 letter codes used by GTAP between one version and another; if this happens the the ‘InputSet’ worksheet needs updating.
 - c. Different versions of the GTAP database contain different numbers of accounts – between versions 5 and 7 these were limited to changes in the number of regions; if this happens the ‘InputSet’ worksheet needs updating.
3. Execution errors to date have indicated that the user has changed something in the one or other of the worksheets. The simplest solution is to restart with the version of SAMGator that was downloaded – you did of course remember to make a backup before you started!!

References

McDonald, S. and Thierfelder, K., (2004). ‘Deriving a Global Social Accounting Matrix from GTAP version 5 Data’, *Global Trade Analysis Project Technical Paper 23*, 2004
Center for Global Trade Analysis: Purdue University.