MyGTAP Data Program:

A Program for Customizing and
Extending the GTAP Database
Multiple Households
Split Factors
Remittances
Foreign Aid
Transfers

Peter Minor and Terrie Walmsley

For the most up to date MyGTAP documentation, model, or programing code, visit: **MyGTAP.ORG**.

Contents

Appendix 1: SAM Format of Shares required for MyGTAP

Forward

The GTAP Standard model has proved a useful analysis tool and data source for over 20 years. The GTAP model has been updated and refined overtime, but it largely maintains the same structure of a regional household, broken into three components: government, private and savings-investment expenditures. This accounting framework has been supported and constrained by available international data sources. The standard GTAP framework was\is acceptable and is sufficient for many applications. However, there have been numerous calls for a more detailed accounting system for the GTAP regional household, especially as it relates to estimating the potential impacts of policies and global shocks on poverty, sustainable and inclusive growth. A single household does not suffice for these purposes. The requirements for these types of analysis is firstly, that the household be representative of key demographic and income groups, such as rural and urban households, and secondly, that there are clear linkages to the sources of the household's income. A number of CGE models have been developed and/or adapted to incorporate multiple households. However, none of these models have provided a flexible, easily replicated approach to adjusting the GTAP database (Narayanan, Aguiar and McDougall, 2012) to include multiple households. The approach taken in our paper is to present not only a method for splitting the GTAP regional household and linking these households to factor incomes and taxes, but also to introduce a GEMPACK based application, MyGTAP, for splitting the GTAP regional household based on basic data and splitting shares which may be obtained from a Social Accounting Matrix (SAM); thereby greatly reducing the development requirements for introducing multiple households into the GTAP model. The introduction of a split regional household (which does not require splitting data for every region in a particular aggregation of the GTAP database) supports greater analysis of impacts based on detailed households, government, factor income, remittances, foreign aid and income transfers.

This paper is a guide to employing the data tools and programs for splitting the regional household and factors of production in the GTAP database. It is intended to be used in tandem with a complimentary paper detailing the theory and model code "MyGTAP Model: A Model for Employing Data from the MyGTAP Data Program: Multiple Households, Split Factors, Remittances, Foreign Aid, and Transfers", GTAP Working Paper by Walmsley and Minor 2012.

Introduction

MyGTAP is a set of GEMPACK programs (Harrison and Pearson, 1996) designed to re-organize and customize the standard GTAP database into multiple households and factors of production to facilitate analysis within a CGE framework[1]. The "My" in the program name reflects customization you define. The "GTAP" in the program name emphasizes that underlying the program is the GTAP database, and it is this that will be reorganized—all the original flows in the GTAP database are maintained after your customization. The standard GTAP model and database are organized around a single regional household, comprising both private households and government. One of the objectives of MyGTAP is to provide a means for the user to define more than one private household. This is especially helpful when one is concerned with the distribution of policy impacts across differing households. For example, will a shock to food import prices impact urban households and rural households equally? By re-organizing the GTAP regional household into multiple households, one can model, explicitly, the distributional impacts of shocks. Only available information and the user's time limit the number of households which may be specified.

Another common requirement from users is to divide the standard five endowment factors — land, skilled labor, unskilled labor, capital and natural resources — into more detailed categories. For example, it may be worthwhile to divide capital into agricultural and non-agricultural capital. Or, it may be meaningful to divide unskilled labor into rural and urban unskilled labor. MyGTAP provides tools to split the standard GTAP endowments into new categories. In addition, the user can combine analysis of multiple households with multiple endowment factors by providing factor ownership shares, thereby creating a solid link between factor earnings and household incomes. The same linkages between taxes and government income follow as a result of removing the regional household—tax revenues will be linked to government savings and deficits.

Sometimes analysis of household impacts requires an expanded definition of "income" to include foreign remittances, aid and household transfers. MyGTAP provides, in the cases of remittances and aid, country level estimates of these values, as well as tools for dividing the values between households and government[2]. Household transfers are entirely user defined.

The current version of MyGTAP is a DOS batch file driven program which executes a series of GEMPACK compiled files as well as several utility programs (e.g., Agghar.exe). The user must be familiar with working with a DOS command box and understand batch files on a basic level. GEMPACK is a suite of programs licensed by

---

[1] The MyGTAP set of programs are based on Mark Horridg's Splitcom program. The authors are grateful for the original open source programs provided by Mark, these programs greatly facilitated the development of MyGTAP and some of the underlying structure of Splitcom can be found in the source code.

[2] This is in contrast to bilateral data, which provides source and destination details, which MyGTAP does not provide. For user familiar with the GTAP database, the treatment of remittance and aid in the current MyGTAP database are similar to the global bank.

Monash University, and a user may be required to purchase a basic license
(http://www.monash.edu.au/policy/gpprice.htm) to fully execute data exercises of
unlimited size[3]. The programs are open source and may be altered by a user to further
extend their usefulness.

The following documentation provides step-by-step directions on the use of MyGTAP.
Since MyGTAP is constrained to return the GTAP data flows, users will find that their
inputs into the program for splitting the database stand as targets, with the actual data
returned often differing from the user's suggested divisions. This is a consequence of
maintaining the balance of the GTAP database. Most users will find several iterations of
their inputs will lead them closer to their desired output. This is expected. MyGTAP
provides several checks and comparison "headers" which provide concise summary
statistics on the information you provided, and how that information had to be modified
to achieve balance in the database. By referring frequently to these files, the user can
avoid surprises and begin a process of refining their input data.

The GTAP database is closely linked to the GTAP model. The Standard GTAP model is
based on a single regional household and it does not permit varying ownership of factor
returns, nor does it contain variables for foreign aid, remittances or household transfers.
An adaptation of the standard GTAP model (Hertel 1997) is now available which has
been modified and augmented to include these variables, along with an option to include
a linear expenditure function, which better represents household consumption in cases
where subsistence consumption is a significant portion of a household's expenditure.
The new model by Walmsley and Minor (2013) may be found at: mygtap.org.

There are plans to further expand the capabilities of MyGTAP. Currently, users are
limited to splitting the private household and factors in one region. This greatly reduces
data requirements by limiting required splitting information to one country. Often a
social accounting matrix (SAM) will contain the required information. The current
version of MyGTAP also limits the splitting of imports and domestic goods; users cannot
split them separately. Future releases will eliminate this restriction. Users may desire a
graphic user interface (GUI) which would make the processes of running MyGTAP
efficient and friendly. This too is planned for future releases. Finally, the current method
requires the user to iterate through several runs of the program to achieve a desired goal
in splitting shares. We are currently considering an option which will permit a user to
prioritize certain data in the splitting process or a program, perhaps in a spreadsheet,
which will permit the user to define ideal shares, which are feasible, before running
MyGTAP. Please provide any comments or questions to info@mygtap.org.

---

[3] A limited exe license, the most basic, should suffice to disable any size restrictions. An unlimited license
would be required to alter the code in the programs.

What Software Do I Need?

A user of MyGTAP will have to possess several software packages and data:

- GEMPACK (http://www.monash.edu.au/policy/gempack.htm);

- GEMPACK utility programs (http://www.monash.edu.au/policy/gpwingem.htm);

- GTAP Database (https://www.gtap.agecon.purdue.edu/databases/default.asp).

MyGTAP was developed on a Windows 7 operating system, but it is backwards compatible with most Windows operating systems. The programs are compiled in 32 bit Lahey FORTRAN and should function on 64 and 32 bit operating systems.

GEMPACK is a suite of programs; while not all components are required, we do require ViewHar, since it is the only program on earth which can read a GEMPACK har file. TABMATE or a good text editor is not required, but could be helpful. A Limited Executable License is required for larger aggregations.

Updated GEMPACK utility programs can be helpful for smooth running of the program, especially as future releases of GEMPACK and the GTAP Database evolve to more advanced specifications. However, we do include current versions of the most recent utility programs in the MyGTAP suite and install them into the MyGTAP directory: you may have to update them in the future.

Access to the GTAP database is also assumed. Some user may gain access to an aggregation of the GTAP database and will not require a license. MyGTAP assumes that you will provide an aggregation of the database in the first place. MyGTAP default is to reorganize the GTAP V8.1 database with 134 countries and 57 sectors. Users wishing to reorganize an aggregation of the GTAPV7 database (or any other GTAP version or year) will need to ensure that they have the relevant *extradata.har* file, as well as their GTAPAgg/FlexAgg output (*basedata.har, sets.har* etc), in the input directory. Several versions of the *extradata.har* file for v7 and the v8 2004 data are provided in the input directory. Should you need another version please contact us.

What Do I Need to Know to Run MyGTAP?

MyGTAP can be run by a novice user, with some practice and care. However it is best that a user is already familiar and comfortable with:

- GEMPACK programs and environment;
- DOS Batch file execution and modification to trouble shoot errors;
- The GTAP database, including the names of variables and their; relevance within an accounting structure;
- Familiarity with working Windows based file systems (Windows Explorer and command prompts).

If you feel you require a refresher on these topics, we recommend downloading and reading over "Exercises in Contributing I-O Tables to the GTAP Data Base" GTAP Resource 2850[4]. The reader should restrict their reading to the sections which introduce GEMPACK coding, DOS batch files and file manipulation within Windows.

Structure of MyGTAP Program

MyGTAP is comprised of a set of GEMPACK files, DOS batch files, and GEMPACK header array files. Tab files are the native format for GEMPACK programs, here you can read the code of all the MyGTAP processes. HAR files are data files in the native GEMPACK format and can only be read with GEMPACK programs like ViewHar. DOS batch files are scripts which help automate running a number of programs and processes all with one command (the technology is as old as the hills, but is the standard for most GTAP and GEMPACK applications). The files are organized into one root directory, ..*MyGTAPVXX* , where the XX will be the numbered version of the program you have downloaded. There are four subdirectories listed in table 1:

Table 1
MyGTAP Directories

Directory	Description
..\MyGTAP	Contains the CMF files for running tab files contained in SRC. MyGTAPf.bat file for running the GEMPACK .exe programs and MyGTAPg.bat for running GEMSIM programs (for users with an appropriate GEMACK license). Compallg.bat and Compallf.bat for compiling GEMSIM files and .exe files respectively
..\MyGTAP\SRC	Contains the original tab files listed in table 2
..\MyGTAP\Input	Directory with original GTAP data files (transferred by user from GTAPag) including: *basedata.har, BaseView.har, sets.har, default.prm*. Template for entering splitting sets: *HHsplit.har*
..\MyGTAP\Output	GTAP data files with split data. *Userwgt.har* file containing the weights applied in splitting the GTAP data, along with several checks of weights entered and weights actually applied
..\MyGTAP\Work	Numerous files created in the processes of running MyGTAP.

Initially, the main file of interest is the MyGTAPf.bat file in the root directory, which will run the MyGTAP programs[5]. After the first run of the program, the directories will be fully populated with the default files.

---

[4] https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=2850.
[5] An alternative file, MyGTAPg.bat is provided for users who have an appropriate GEMPACK license.

Table 2 lists the tab files employed by MyGTAP. The files are numbered in the order in which they are run. The programs must be run in this specific order, since subsequent programs rely on earlier ones. Each program is self documented and can be read and modified in GEMPACK. Most users will not have to be concerned with these details, unless there is an error message or in the case if you wish to modify the code.

Table 2
My GTAP Program Files

File Name	Description
1_REMETC.TAB	Aggregates remittances, foreign capital income and foreign aid to match your aggregation
2_MyGTAP.TAB*	Creates default splitting shares
3_GTP2NORM.TAB*	Transforms GTAP database into "normalized" data format for splitting purposes***
4_NORMCHEK.TAB*	Checks that the normalized data matches the GTAP data
5_EXPWGTHH.TAB*	Weights consistent with the normalized format are created for splitting flows. Numerous re-scaling and checks take place in this processes
6_BALANCE1.TAB	Rebalances private household consumption by household to consumption recorded in the GTAP Database.
7_SPLITFLO.TAB*	The normalized GTAP data flows are split. Final checks are made
8_NORMCHECK.TAB*	The split data in normalized format are checked for consistency with the original GTAP data flows
9_NORM2GTAPHH.TAB*	The split data flows are re-organized into the standard GTAP header files
10_INCOME.TAB	Program which changes the GTAP income data to include aid and remittances
11_GTPVEWHH.TAB**	Adaptation of the popular GTAP view program to include the multiple households, factors, remittances, household transfers and aid flows
12_CHKDIFF.TAB*	Checks the new flows, in GTAP format, against the old flows
13_SAMVIEW.TAB	Foir those interested in seeing their output in the GTAP SAM (McDonald and Theirfelder (2004)

*Programs were based on code originally authored by Mark Horridge in SplitCom and adapted by the authors of MyGTAP.

**Programs originally authored by Hertel, Bacou, and Yu and adapted by the authors of MyGTAP.

***The normalized data set stands in contrast to a more conventional and comprehensive SAM approaches. The normalized format provides some advantages in its ease for splitting data. On the other hand, the normalized format is stylized and does not provide the comprehensive double entry accounting contained in a SAM. MyGTAP could be implemented in a SAM format with no impact on the results.

The following sections provide step-by step directions for entering the splitting data and running the programs outlined above.

LOADING YOUR BASE GTAP DATA

You will have to run either GTAPAgg or FlexAgg to create a version of the GTAP database suited to your purposes. Transfer files from FlexAgg or GTAPAgg after you have made your aggregation into the ..\input directory under MyGTAP. Several files are required and should have the file names specified below[6]:

- basedata.har;
- baseview.har;
- sets.har;
- default.prm.

RESTRICTION: THE FIVE GTAP FACTORS OF PRODUCTION, LAND, SKILLED AND UNSKILLED LABOR, AND CAPITAL CAN NOT BE RENAMED OR AGGREGATED[7].

Use ViewHar to open the file *HHsplit_Base.har* and use the menu options FILE | SAVE AS and rename *HHsplit_Base.har* to *HHsplit.har* by typing in the new name in the box when prompted and SAVE. After you have saved the file, Open it and review its contents. This file includes four headers listed in Table 3.

Table 3
Header Array File HHSPLIT.HAR

Row No.	Header	Name
1	REGS	Special region for which data will be split
2	ENDS	The new factor split--old and new factors
3	EMAP	Mapping from new factors to old (FAC1 to FAC0)
4	HHS	The household split--includes mainhh

In *HHSplit.har*, click on the header REGS. Enter the name of the country\region you wish to split under header REGS. The current version of MyGTAP supports one split region. There is one dummy region to start you off ("Reg_Split"). Overwrite this with your split region as follows in Table 4 (assuming you are following along with the example here, eventually you will provide your own data). Note that the country/region name must match the name given in your GTAP *sets.har* data file. Return to the top view of the headers in this header array file by clicking CONTENTS in the menu.

---

[6] If you use FlexAgg you may need to change the names of the output files to match these names.
[7] New household names and new factors should not use any unusual characters or spaces. Although GEMPACK allows for the use of flexible characters by entering a style statement into the CMF file, unpredictable errors can occur when using many of the GEMPACK "utility programs" such as agghar.exe and cannot be supported in a generalized application like MyGTAP without bypassing these utilities.

Table 4
Header REGS

Row No.	String
1	Tanzania

Click on the header ENDS. Enter your new factors into ENDS. There are six "dummy" names to start you off in this example ("fac1", "fac2", "fac3" fac4", "fac5" and "fac6"). You will likely have to resize the header array to fit the expected number of new factors, including any unsplit factors which you will want to use in your own application[8]. You may do this manually in Viewhar or cut and paste from an Excel file[9]. DO NOT UTILIZE THE ORIGNAL GTAP FACTOR NAMES FOR A FACTOR THAT IS SPLIT[10]. If a factor is unsplit, enter the standard GTAP factor name (e.g., "land" and "Natres" see Table 5). Return to the top overview again to view the headers by clicking Contents in the menu.

Table 5
Header ENDS

Row No.	String	Original factor name
1	Land	← Unsplit factor
2	unsk_rural	← Split factor (Unsklab)
3	unsk_urban	← Split factor (Unsklab)
4	skl_rural	← Split factor (Sklab)
5	skl_urban	← Split factor (Sklab)
6	Natres	← Unsplit factor
7	oCapital	← Split factor (Capital)
8	agCapital	← Split factor (Capital)

Note: The shaded are is added here to assist in understanding, it does not appear in the .har file.

After returning to the list of headers, click on the header EMAP. Edit EMAP to provide an appropriate mapping from the new factor names to the old factor names. We make some suggestion you may wish to follow in this example in Table 6.

Making a GEMPACK mapping file requires you to reference the positions \line numbers of the new factors as ordered in ENDS header and enter the "original" GTAP factor name next to the line number corresponding to the split factor name in ENDS[11]. The mapping

---

[8] Go to Edit | Change Size.
[9] Go to File |Advanced Editing Menu. Then Import | Paste to Screen from Clipboard.
[10] If you do, the program will abort prematurely
[11] For new users of GEMPACK, element order is important and meaningful. GEMPACK mapping files can take some time to master, despite their relatively straightforward structure and interpretation. New users would do well to take care and check their mappings after the program is run the first time, to verify what

includes any GTAP factors left unsplit (e.g., in Table 5 "Land" and "NatRes" are included).

Table 6
HeaderEMAP

No.	String	Maps to
1	Land	→ Land
2	UnSkLab	→ unsk_rural
3	UnSkLab	→ unsk_urban
4	SkLab	→ skl_rural
5	SkLab	→ skl_urban
6	NatRes	→ Natres
7	Capital	→ oCapital
8	Capital	→ agCapital

Note: The shaded area is added here to assist in understanding, it does not appear in the .har file.

Return to the header list by clicking on Contents and open header HHS and enter the element names of the split private household[12]. The base file includes three dummy elements for the split household ("hh1", "hh2", "hh3"). Again, you may need to edit the size of the array to correspond to the number of households you desire. In Table 7, the dummy headers have been edited to be meaningful descriptions of households.

Table 7
HEADER HHS

No.	String
1	Rural
2	Urban
3	Mixed

Note: No mapping file is required for the households, since they all map back to a single household in GTAP – the regional household. Government income and expenditure will be derived from existing tax rates and expenditures.

---

was done was what they intended to do. Refer to the GEMPACK documentation for more information on mapping files.
[12] The government component of the regional household will be split out by default. The user should only list private household splits.

RUNNING MYGTAP

These four arrays are all that are required for MyGTAP to begin its work and split the GTAP database. We recommend that you run the program at this point to:

- Ascertain any errors you might have made in data entry (prior section);

- Generate default splitting weights and annex data, which you can then edit to suit your specific situation (it is easier to start with these default splitting weights and modify them. The default weights will be contained in ..\output\userwgt.har)

Running MyGTAP involves executing numerous programs. This process is managed through a "DOS" batch program. The program is run from the command prompt, which can be accessed by selecting the MyGTAP directory in windows explorer and holding down the <SHIFT> key while right clicking on your mouse. A menu will appear and there will be an item "Open command window here." Select this option and the DOS box command window will appear (easily recognized by its distinct "retro" look of white letters (no graphics) on a black background[13].

From the command line in the DOS command box, in the ..\MyGTAP directory, type "MyGTAPf " without the quotes to execute the batch file MyGTAPf.bat.

The batch file will run and it will list the log files of the various programs. If you appropriately entered the correct data, the program will end with a prompt "BATCH JOB SUCCESFUL"

If your program did not complete successful, the program will list the log files. You will be interested in exploring the last log file, which will contain the offending error. You must correct this error to continue. The log file may be edited with any text editor, including TABmate. While in the DOS box, one can type Tabmate <filename.log> with the name of the relevant log file. TABmate should execute and open the log file[14]. Proceed to the end of the file to discover a clue to the error. Fix the error and re-run MyGTAP.

ENTERING SPLITTING WEIGHTS AND SHARES

After you have successfully executed the first run of MyGTAP, several files will appear in the output directory. These include new GTAP data files: basedata.har, BaseView.har, sets.har, default.prm and a new file called usewgt.har.

---

[13] Note, you could execute batch programs from inside windows, by doubling clicking them in Windows Explorer. However, you would not be able to monitor any of the screen output which would indicate errors in the execution of the program. We highly recommend running these from the "retro" DOS command window.

[14] If Tabmate does not respond to this command, you may have to insert the location of Tabmate, into your computer's PATH. See Windows documentation on the PATH statement.

Splitting weights or shares are required for at least the following three value flows for the split region defined in the prior section:

- factor use splits;

- household consumption splits;

- and factor ownership shares for redistributing income to the new, split, households.

Other user defined weights include savings rates, transfers to government and/or between households and remittances flows by households.

MyGTAP will generate default splitting weights if none are provided. These default weights are not very helpful in of themselves, but they do provide an excellent starting point for entering your own splitting weights. Simply run the program after specifying the sets above, and then refer to the file userwgt.har in the ..\output directory. The userwgt.har file has the following headers[15] (Table 8):

Table 8[16]
*Header Array File ..*output\userwgt.har

No.	Header	Description
1	WTF	Factor splitting weights
2	OWN	Factor ownership shares
3	WTC	Consumption weights by household
4	SRTE	Household Savings rate as a share of income
5	TRNH	Transfers between Households ($US Millions)
6	TRNG	Government transfers to households ($US Millions)
7	SHLD	Household Savings
8	SGOV	Government Savings
9	IHLD	Household income
10	IGOV	Government Income
11	PCON	Total Private consumption by commodity
12	SAVE	Regional Household Savings rate
13	KCON	Check consumption shares
14	KFAC	Check factor splits old to new
15	KOWN	Check ownership shares old to new
16	KSAV	Check household savings rates old to new

---

[15] If the headers are not arranged as presented here, you may have to change the header order option found in File | Options | Sort Headers Alphabetically. Make sure this option is not checked.
[16] Users interested in how these additional data relate to the SAM structure are referred to Appendix 1.

The first five headers are the subject of this section on factor and household splitting weights. The separate section covers transfers, remittances, and foreign income and aid to the government.

Factor Splitting Weights

The first header array WTF, includes the factor use splitting weights, which define how factor use by activity are split into the new factors. This header has three dimensions, ALLFAC and TRAD_COMM by REG. In order to see meaningful data, at least one dimension must be selected. In this case, it is meaningful to select the region (REG) concerned. We will pick "Tanzania", since it is the region for which we wish to split the household. Table 9 lists the default splitting weights for the new factors and the original factors.

Table 9
Header WTF – Default Factor Splitting Weights (Tanzania)

WTF	1 GrainsCrops	2 MeatLstk	3 Extraction	4 ProcFood	5 TextWapp
1 Land	1	1	1	1	1
2 UnSkLab	0	0	0	0	0
3 SkLab	0	0	0	0	0
4 Capital	0	0	0	0	0
5 NatRes	1.0	1.0	1.0	1.0	1.0
6 unsk_rural	0.5	0.5	0.5	0.5	0.5
7 unsk_urban	0.5	0.5	0.5	0.5	0.5
8 skl_rural	0.5	0.5	0.5	0.5	0.5
9 skl_urban	0.5	0.5	0.5	0.5	0.5
10 OCapital	0.5	0.5	0.5	0.5	0.5
11 AgCapital	0.5	0.5	0.5	0.5	0.5
Total	5.0	5.0	5.0	5.0	5.0

Note, we have limited the number of commodities here to just five, several others may be present. The values in Table 9 indicate original factors in two ways: they have either a one or a zero value. A one indicates the factor is unsplit and hence the GTAP values should be taken, and a zero indicates the factor is split and hence the original aggregated factor is no longer relevant.

Where there is a split factor, the default splitting share is simply calculated by allocating equal parts to the new factor e.g., if the original GTAP factor was split in two parts, then each new split factor gets 0.5 (50%) of the factor income. If the factor was split into three factors, each would get 0.333 (33%) etc.

For maximum flexibility, the factors have been listed by PROD_COMM , so factor splits may be carried out by activity/commodity, allowing factors to be deferentially split/used by each produced commodity.

Editing a .har file directly can be somewhat cumbersome, so you may wish to export the .har file to Excel for editing, and import the results back into the .har file using export and import or CTL-C and CTL-V.

When you are done editing the *userwgt.har* file, be sure to SAVE THE FILE IN THE ..\INPUT directory. The *userwgt.har* file in the ..\output directory will be over written when the program is re-run.

After you have made and saved your new splitting weights, you will have to make one last change to the header to make MyGTAP recognize your unique splitting shares. The header must be renamed WTF1. Adding the last character, one (1), indicates that the header is the new, user defined, splitting share, not the default. The Viewhar feature for changing a header name can be found under EDIT CHANGE HEADER OR NAME. Be sure to change the header name, the coefficient name (WTF) is for information purposes.

SAVE YOUR NEW SPLITS: FILE| SAVE AS. Be sure to save your changes into the ..\INPUT directory.

Factor Ownership Shares

After defining factor splits, you will want to consider the factor ownership shares. Factor ownership shares define household incomes when combined with factor incomes. The second header, OWN, defines the ownership of factor incomes by household. Again, the header has three dimensions, requiring that we hold one dimension constant to view meaningful data. Again, we will look at Tanzania, the country with a split household in Table 10.

Table 10
Header OWN – Default Factor Ownership Shares (Tanzania)

OWNWGT	1 MainHHLD	2 Rural	3 Urban	4 Mixed	Total
1 Land	0	0.333333	0.333333	0.333333	1
2 UnSkLab	0	0	0	0	0
3 SkLab	0	0	0	0	0
4 Capital	0	0	0	0	0
5 NatRes	0	0.333333	0.333333	0.333333	1
6 unsk_rural	0	0.333333	0.333333	0.333333	1
7 unsk_urban	0	0.333333	0.333333	0.333333	1
8 skl_rural	0	0.333333	0.333333	0.333333	1
9 skl_urban	0	0.333333	0.333333	0.333333	1
10 OCapital	0	0.333333	0.333333	0.333333	1
11 AgCapital	0	0.333333	0.333333	0.333333	1
Total	0	2.666667	2.666667	2.666667	8

Table10 includes ownership shares for each of the factors in Tanzania. Original factors, split, have a zero value in the ownership shares. For original factors unsplit ("Land" and "NatRes"), a default splitting share of 1 / (# of households) or one-third is entered. New, split factors, have the same default splitting shares. The MainHHLD represents the non-split private household in the GTAP model and applies to the regions with only one household data. In this case, the MainHHLD has zeros in the case of Tanzania, since the household has been defined as split in the earlier section of our work. Each of the three new households ("Rural", "Urban", "Mixed") in Tanzania has an equal share of the factor incomes by default.

In the default case here, each household will be allocated exactly one-third of disposable factor income. If the shares of factor ownership are changed, as they should and likely will be, then disposable household income (EVOA) will vary by household. However, at all times, total disposable income by all households will be the same as in the GTAP database; only the shares of income of the split household will change.

Change the shares, using the method applied in the previous section on factor weights. You may change the ownership shares in Viewhar, or use a spreadsheet to provide rich and transparent calculations of factor ownership shares and the resulting household incomes implied by them. Then import the shares into the OWN header.

Factor Substitution and Factor Ownership
Many CGE models with multiple households will define a one-to-one correspondence between a factor and a household. This is common where you might have unskld_rural all going to the rural household. In this case, simply enter a one in the row and column for that factor and household. If there is not a one-to-one correspondence, you will define shares. The fact that MyGTAP allows you to map factors in something else than a one-to-one correspondence provides flexibility. When a factor is split, it then has its own price in the model, which will vary independently of other prices. Allowing a factor to have ownership, which crosses households, recognizes the fungibility of a factor, or substitution, resulting in the law of one price.

Finally, change the name of the header OWN to OWN1 so MyGTAP knows that these are your unique factor ownership shares.

SAVE YOUR NEW OWNERSHIP WEIGHTS by selecting FILE |SAVE.

Commodity Consumption Shares

The third header on the userwgt.har file is WTC, the weights of each commodity in the private household consumption basket. Once again, this header is defined over three dimensions, TRAD_COMM, ALLH, and REG, requiring us to hold one of the dimensions constant. Again, we will hold the region (REG) constant to "Tanzania" our region with a split household and factors.

The ten commodities are listed in the rows (Table 11), with each of the households listed in the columns. MyGTAP has populated each household with the share of each commodity in the GTAP regional household's consumption basket. In this case, the GTAP database (V8.1) asserts that 25.7% of private consumption in Tanzania is "GrainsCrops". The value is the same for each household, "Rural", "Urban" and "Mixed". For each household, the shares add to one, as illustrated in the bottom total row (the right hand totals are meaningless).

Table 11

Header WTC – Default Consumption Shares(Tanzania)

WTC	1 MainHHLD	2 Rural	3 Urban	4 Mixed	Total
1 GrainsCrops	0.0000	0.2257	0.2257	0.2257	0.6771
2 MeatLstk	0.0000	0.1118	0.1118	0.1118	0.3355
3 Extraction	0.0000	0.0323	0.0323	0.0323	0.0968
4 ProcFood	0.0000	0.1974	0.1974	0.1974	0.5923
5 TextWapp	0.0000	0.0547	0.0547	0.0547	0.1640
6 LightMnfc	0.0000	0.0525	0.0525	0.0525	0.1574
7 HeavyMnfc	0.0000	0.1437	0.1437	0.1437	0.4311
8 Util_Cons	0.0000	0.0270	0.0270	0.0270	0.0810
9 TransComm	0.0000	0.0840	0.0840	0.0840	0.2520
10 OthServices	0.0000	0.0709	0.0709	0.0709	0.2127
Total	0.0000	1.0000	1.0000	1.0000	3.0000

You can edit the shares of each commodity in the consumption bundle of each household. However, bear in mind, your new shares will be recalibrated so that the total regional consumption of a given commodity will be equal to the original private household consumption by commodity in the GTAP database. In practice, the shares you input are guidance, or targets, which often must be adjusted to fit the existing GTAP data.

An additional header, PCON, for private household consumption, by commodity, at market prices, is provided in header 11 and may be useful for analysis when defining WTC and savings in the following section. The data are extracted from the GTAP database and must hold before and after the splitting of households.

 The user should be careful NOT TO ASSIGN A CONSUMPTION SPLITTING SHARE OF ZERO weight for any commodity in any household, or the model may not solve. This is an artifact of the model's elasticity calibration and the data balancing. Instead, code zero shares as very small numbers, for example 0.00001 — which is one-one-thousandth of a percent-so close to zero as it does not matter to anything other than software running the model. Remember to rename the header WTC1 and save in the input directory.

Household Savings Rates

The fourth header in the userwgt.har file is the household savings rate SRTE. In the earlier headers, we defined factor splits, factor ownership shares and consumption by commodity and household. In order to derive private expenditure shares by household, we have to subtract household savings (header SHLD) from disposable household income (header IHLD and equal to factor incomes minus depreciation, plus net foreign labor and capital remittances). In the case of SRTE, the header is defined across two dimensions – HHLD and REGS.

Note that comparing the savings or the savings rate for Tanzania in the GTAP database with the post-split database is not helpful, since the regional household savings rate in the standard GTAP database includes both private household (SHLD) and government (SGOV) savings. Moreover the inclusion of foreign remittances and aid alters the total savings in the economy, although not global savings.[17] Our split region therefore has two savings rates: one for government and one for the split household. The government's savings depends on Government income (from taxes and foreign aid) less government spending and transfers. Household savings, on the other hand, will depend on their income (factor income plus foreign remittances and transfers) less expenditures.

Table 12 illustrates the household savings rate for each of the split households in the split region ("Tanzania"). The default values are calculated from the GTAP database and are implied from the data on total disposable household income (IHLD) and private household expenditure (PCON). The values in Table 12 suggest a household (non-government) savings rate of around 0.87 percent —a fairly small value. Overall, the private household savings rate of 0.87 percent must hold. New savings rates can be entered, perhaps lowering the savings rate on rural household and raising it on urban households. You can enter the savings rates in absolute terms and MyGTAP will rescale them to comply with the overall rates implied in the data. Enter the new savings rates in an appropriately dimensioned header in *input\userwgt.har* called SRT1 (a copy of header SRTE will work).

> ### Savings Rates – Without the Regional Household
> *The regional household in GTAP was designed to combine information on factor incomes and tax flows in one account. This design was driven, in part, by the fact that domestic tax data for many countries, especially for developing countries, could be difficult to reconcile. Therefore, detailed flows of income and expenditure were not required and the savings rate for the country as a whole could be determined with confidence.*
>
> *MyGTAP eliminates the regional household and separates government income from household income and expenditures. This results in two savings rates: one for the household and another for government. Both are derived as revenue (taxes or factor incomes plus net capital and labor remittances plus net foreign aid) minus expenditures. In some cases, a country may have a positive savings rate, but the household may have a small or even negative savings rate, since expenditures exceed factor incomes. The user will want to consider this situation carefully and consider changes to the database which may be required.*

---

[17] Header SAVE in Table 8 contains the overall savings rate for the regional household for the split region. It therefore includes the average of both government and household savings, after taking remittances into account. If you find your household savings rates are lower (SRTE) than the average savings rate (SAVE), then government is the primary saver n the economy.

Table 12

Header SRTE Household (non-government) Savings Rates as a Share of Total Income (Tanzania)

SRTE	ASAVERATE
1 MainHHLD	0.0000
2 Rural	0.0087
3 Urban	0.0087
4 Mixed	0.0087
Total	0.0261

If you find the rates contained in header SRTE to be totally unreasonable, and you have better data on tax rates than GTAP offers, you can use Altertax (Malcom, 1998) to change disposable (after tax) household income to support your analysis. Alternatively, you can use the aid and transfer data provided (reviewed latter in this documentation) to move revenue from government to households etc..

Depreciation Splitting Weights

Depreciation splitting weights are used to allocate depreciation across households. Depreciation is split based on the proportion of capital rents in household income. It is highly recommended that you use these splits since ill-conceived shares can result in a household receiving negative returns to capital. For this reason we have not allowed users to change these shares independently.

Government Transfers

Transfers of income from government to households, for example pension payments, may be added in the header "TRNG". The amount to be transferred from the government, to the household should be entered in millions of dollars, rather than as shares. Unlike WTF, WTC or WTO, there are no default values, since the default is considered to be zero in all cases. It is therefore not required that you rename the variables with a "1" at the end of the header name: simply use "TRNG".

Household Transfers

Transfers of income between households, for example remittances between urban and rural populations within the same country, may be added in the header "TRNH". The amount to be transferred from one household, to another household should be entered in millions of dollars. Like government transfers, no initial splitting shares are provided and the values are assumed to be zero. It is not required to rename this header with a "1" at the end of the header name: simply use "TRNH".

Augmenting GTAP Income Flows

The standard GTAP database does not distinguish important income flows such as foreign labor remittances, foreign capital remittances or foreign aid flows. These flows can be

important components of household and government income and are included in the MyGTAP data application. For more information on augmenting income flows in the MyGTAP model and application, see McDonald and Sonmez 2012 and Walmsley and Minor 2012. In the cases of foreign remittances and capital payments and foreign aid, the MYGTAP application augments household income for all regions for these values automatically based on international estimates of these values. While the estimates on a national scale are fixed, the user has the capability to distribute the flows by altering selected shares.

Foreign Remittances (labor)

Labor remittances out of and into a given country are divided up by each households share in labor income. It is strongly advised not to change these shares since ill-conceived shares may result in negative income. Should a user wish to change these shares they can do so by adding two headers RMI1 and RMO1, of dimensions HHLD and REGS, for remittances in and out of the split country respectively. The total values of remittances in and out are set out in input/*extradata.har*.

Foreign Income (capital)

Capital remittances (payments on foreign owned capital) out of and into a given country are divided up by each households share of capital in household income. It is strongly advised not to change these shares since ill-conceived shares may result in negative income. Should a user wish to change these shares they can do so by adding two headers FYI1 and FYO1, of dimensions HHLD and REGS, for remittances in and out of the split country respectively. The total values of remittances in and out are set out in input/*extradata.har*.

Foreign Aid

Foreign aid is included in government income. The user is not able to change this flow directly. However, a user can transfer income from the government to the households employing "TRNG" (see earlier section).

Checking Your Results

MyGTAP is a program which splits the GTAP database, maintaining all of GTAP's value flows. Often a user will employ a SAM for a local economy. That SAM or other data sources are not likely to correspond exactly to the value flows within the GTAP database. In many cases, the differences will be modest; in other cases the differences could be very large. The presence of large structural differences in the two data sources must be reconciled. MyGTAP will employ the splitting weights and shares you provide as ideal targets; however, the data must balance and be consistent with the original GTAP database. This means MyGTAP will have to modify your targets (shares and weights) to best fit the structure of the GTAP data. Checking tables are provided to help you ascertain how your scaling weights have been modified from their original values.

Consumption Shares

The example applied in the following sections diverges from the one used in previous sections of this document, if only because it provides a richer view of the data. The following data can be generated by renaming *hhsplit_Tanzania.har* to *hhsplit.har*, overwriting the earlier hhsplit file based on the example file. The *userwgt_Tanzania.har* file is also renamed to *userwgt.har*. Re-run MyGTAP by typing MyGTAPf.

MyGTAP allows for a number of splits in factors, households, and household ownership of factor incomes. In most cases, these values will remain unchanged, unless you have submitted values which were not reasonable (i.e., factor ownership shares which add to more than one). In the case inconsistent data were entered, MyGTAP will rescale weights and shares to make them compatible with economically reasonable splitting shares.

In the case of consumption shares, WTC, it is likely that no matter what shares you put into this value, MyGTAP will likely rescale them to satisfy the requirement that all private consumption in the GTAP database are consumed by the split households. You will need to review these changes any time you run MyGTAP.

Header kCON, includes several tables for comparing your original household consumption shares, WTC1, with the new household consumption shares calculated by MyGTAP. The table has four dimensions, TRAD_COMM, ALLHH, CHK1, and REG. The region you will be interested in is the split region, in this case Tanzania. That will leave three dimensions for your analysis. CHK1 contains four elements:

   a. UserIN: user input (WTC1 or basic splits where no user input),

   b. ProgOUT: output from MyGTAP,

   c. Dif_Level: the difference between the user specified shares (UserIN) and the programs calculations in levels (ProgOUT), and

   d. Scale: (ProgOUT) / (UserIN). If no changes are made, the value will be 1. If the proportion was increased, the value will be 1+the percentage change in the proportion.

It may be helpful to select a particular commodity, the split region, and display the households and CHK1 variable at once. This will give you a full display of changes by commodity. Table 13 includes an analysis of Tanzania using actual weights derived from Tanzania's SAM and multiple households.

Table 13

Header kCON – Check household consumption values ("Tanzania", "MeatLstk")

kCON	1 UserIN	2 ProgOUT	3 Dif_Level	4 Scale	Total
1 MainHHLD	0.0000	0.0000	0.0000	1.0000	1.0000
2 hapr	0.0774	0.1121	0.0347	1.4482	1.6724
3 hanp	0.0798	0.1128	0.0329	1.4123	1.6378
4 hrpr	0.0606	0.0864	0.0258	1.4255	1.5982
5 hrnp	0.1006	0.1381	0.0375	1.3725	1.6487
6 hupr	0.0622	0.0879	0.0258	1.4145	1.5903
7 hunp	0.0858	0.1068	0.0211	1.2456	1.4593
Total	0.4663	0.6440	0.1777	9.3186	10.6067

Note: this table has different data then the default values used earlier. Using the default values would result in a trivial example.

In Table 13, we have selected "Tanzania", our split region, and one commodity "MeatLstk". The four CHK1 values are listed across the top of the table. In the case of "hapr" (poor agricultural households), there was a decrease in the user defined consumption share in household consumption for "MeatLstk". The share was increased from 7.7 percent of consumption, to 11.2 percent of consumption, a 44.8 percent increase (Scale-1) or 3.5 percentage points. In the case of, "hunp" (poor urban households) had an increase in the consumption share spent on "MeatLstk" of -24.45 percent or 2.1 percentage points from 8.6 percent to 10.7 percent.

Savings Rates

Like the consumption shares, the household savings rates are likely to differ from the values you submitted in SRTE. The new savings rate by household, SRTE, is obtained by referencing the private savings rate implied by the GTAP data base, then scaling the GTAP savings rates by household according to the relative values submitted in SRTE. The pre and post savings rates are included in the header kSAV. Again, this rate is different than the savings rate for the regional household, which includes government savings as well.

Table 14

Header kSAV – Check Household Savings Rates ("Tanzania")

kSAV	1 UserIN	2 ProgOUT	3 Dif_Level	4 Dif_PCT	Total
1 MainHHLD	0.0000	0.0000	0.0000	0.0000	0.0000
2 hapr	0.0800	0.0071	-0.0729	91.1075	91.1217
3 hanp	0.1000	0.0089	-0.0911	91.1075	91.1253
4 hrpr	0.0500	0.0044	-0.0456	91.1075	91.1164
5 hrnp	0.0800	0.0071	-0.0729	91.1075	91.1217
6 hupr	0.0400	0.0036	-0.0364	91.1075	91.1146
7 hunp	0.1200	0.0107	-0.1093	91.1075	91.1288
Total	0.4700	0.0418	-0.4282	546.6448	546.7284

Other Shares and Weights

kCON and kSAV are likely to be the headers you are most interested in and will differ from the shares and weights you originally submitted on WTC1 and SAV1. There are three other check headers associated with the factor splitting shares (kFAC) and ownership shares (kOWN). These headers take on the same format as kSAV and kCON and they include an analysis of any re-weighting of the original values you submitted in FAC1 and OWN1. The difference here is that any changes made by MyGTAP are likely due to data entry errors and you may wish to revisit your data tables to ascertain the source of the inconsistency which MyGTAP has had to reconcile. If the data in these headers is entered accurately, and consistently, there is no reason they should be rescaled.

Differences in GTAP Value Flows

Up to this point, all the checks were between the weights you suggested and the weights applied by MyGTAP. As a final check, we want to make sure all the value flows in the new GTAP database equal the old, unsplit, database. Header array file ..\work*alldiffs.har* includes fourteen headers corresponding to the differences in key value flows, from the original data, to the new database. The values in each of these headers should be zero, or contain very small values[18].

Homogeneity Check

After any modification of the GTAP database, the user is encouraged to run a simulation with a numeraire shock to test for homogeneity. All changes in quantity should be zero, or very close to zero. This final check cannot be taken for granted. MyGTAP is a complex set of programs and it is impossible for the developers to predict every possible set of inputs a user might supply, correct or not. Running this check is your insurance all values have been modified to maintain the original balance in the GTAP database.

---

[18] Viewhar can be setup to see the total of all values in a particular header. This option can be helpful to look at the values in all fourteen headers at once, without opening each header. See File | Options in Viewhar.

Bibliography

Harrison, W. J. and K. R. Pearson, 1996. "Computing Solutions for Large General Equilibrium Models Using GEMPACK", *Computational Economics*, Vol. 9, pp.83-127.

Horridge, M., R. McDougall, B. Narayanan, and T. Walmsley, 2008. "Exercises in Contributing I-O Tables to the GTAP Data Base" GTAP Resource #2850. Purdue University, West Lafayette, Indianan, USA.

Malcom, G., 1998. "Adjusting Tax Rates in the GTAP Data" GTAP Technical paper Series, 12, Purdue University, Global Trade Analysis Program.

McDonald, S. and Y. Sonmez, 2004. "Augmenting the GTAP Database with Data on Inter-Regional Transactions" mimeo, Purdue University, Global Trade Analysis Program.

McDonald, S. and K. Theirfelder, 2004 "Deriving a Global Social Accounting Matrix from GTAP Versions 5 and 6 Data" GTAP Technical paper Series, 22, Purdue University, Global Trade Analysis Program.

Narayanan, G., B., A. Aguiar and R. McDougall, Eds. 2012. *Global Trade, Assistance, and Production: The GTAP 8 Data Base*, Center for Global Trade Analysis, Purdue University

Walmsley, T. and P. Minor, 2013 "MyGTAP Model: A Model for Employing Data from the MyGTAP Data Application including Multiple Households, Split Factors Remittances, Foreign Aid, and Transfers " GTAP Working Paper, Center for Global Trade Analysis, Purdue University, West Lafayette, Indiana.

Appendix 1: SAM Format of Shares required for MyGTAP

	Commodities	Activities	Factors	Households	Government	Capital	Margins	Rest of World	Totals
Commodities	0	Combined Intermediate Use Matrix	0	WTC: Private Consumption	Government Consumption	Investment Consumption	Exports of Margins (fob)	Exports of Commodities (fob)	Total Demand for Commodities
Activities	Domestic Supply Matrix	0	0	0	0	0	0	0	Total Domestic Supply by Activity
Factors	0	WTF: Expenditure on Primary Inputs	0	0	0	0	0	0	Total Factor Income
Households	0	0	OWN: Distribution of Factor Incomes	TRNH: Transfers	TRNG: Transfers	0	0	RMI, FYI: Remittances	Total Household Income
Government	Taxes on Commodities	Taxes on Production / Taxes on Factor Use	0 / Direct Taxes	Direct/Income Taxes	0	0	0	Foreign aid	Total Government Income
Capital	0	0	Depreciation Allowances	SHLD: Household Savings	Government Savings	0	Balance on Margins Trade	Foreign Savings	Total Savings
Margins	Imports of Trade and Transport Margins	0	0	0	0	0	0	0	Total Income from Margin Imports
Rest of World	Imports of Commodities (fob)	0	0	RMO, FYO: Remittances	Foreign aid	0	0	0	Total Income from Imports