

Title: Towards consistent systematic sensitivity analysis in the GTAP model

Maros Ivanic, US Department for Agriculture
Hitesh Shah, Department for Business and Trade (UK)

Abstract

Systematic Sensitivity Analysis (SSA) is an important complement to GTAP CGE analysis. While tools in RunGtap/RunGEM allow for SSA in the GTAP model, they are limiting in more complicated scenarios: i) the GUI SSA tool only allows for varying either shocks or parameters but not both at the same time (enhancements allow simultaneous variation but require modifying TABLO model file, limiting the appeal to a wider user-base) ii) the choice of statistical distributions for capturing input uncertainty is restricted to the triangular or uniform distributions and iii) the number of multiple simulations is not in users' control rendering Monte Carlo type analysis infeasible.

In this paper we introduce an R package that overcomes some of the limitations of currently available off-the-shelf SSA tools. By allowing R to control the GTAP model--including its variables and parameters--it is possible to bypass the technical limits of the GEMPACK models and vary any variable or parameter simultaneously using any distribution (including empirical ones). Though not the first R package in this area -- the *rgtap* package (Villoria 2017) provides greater flexibility for SSA but does not appear to be up to date and requires refitting for addressing some of the above limitations -- *rssa* offers some practical advantages like minimal need for changing TABLO or R code for example.

To demonstrate the value of our software tool, we undertake a Monte Carlo-type analysis using the standard GTAP model, incorporating variation in shocks and parameters where uncertainty is captured via normal and uniform distributions.

1. Introduction

Many studies based on Applied General Equilibrium (AGE) models like the GTAP model (whether implemented in GAMS, GEMPACK or other frameworks) routinely report confidence intervals or ranges for model results reflecting uncertainty in estimates of parameters and/or inputs.

The ability to model sensitivity of results can help the researcher generate useful insights about key drivers and contribute to enhancing the credibility of the analysis by highlighting situations where model results are sensitive to certain inputs and instances where the model results are more robust.

Methods for translating input uncertainty into statistical confidence intervals or ranges, often referred to as Systematic Sensitivity Analysis (SSA), rely on computing tools for their implementation if more than simple SSA is desired.

While there are existing tools for performing SSA – for example, RunGtap/RunGEM and *rgtap*, an R package - some of their features are limiting in more complicated scenarios. In this paper we introduce another R package, named *rssa*, that overcomes some of the limitations.

The remainder of this paper is organised as follows: In section 2, we present a short literature review on existing SSA capabilities (sub-section 2.1), briefly discuss existing sensitivity analysis tools relevant to GTAP in GEMPACK (sub-section 2.2) and the comparative usefulness of *rssa* (sub-section 2.3). Section 3 discusses the functionality of the *rssa* package in greater detail: starting from its basic logic (sub-section 3.1) to system requirements (sub-section 3.2) and detailed application instructions (sub-section 3.3). Section 4 briefly discusses a number of specific applications and example outputs to illustrate *rssa* capabilities. Section 5 concludes.

2. Why another R-based SSA tool?

2.1 Literature Review

Pearson and Arndt (2000) document an early automated SSA procedure for use with GEMPACK – the SSA method uses numerical integration methods, especially Gaussian Quadratures, to significantly reduce the number of model-solves required for generating means and standard deviations for every single endogenous variable of the specified model.

Currently, SSA based on numerical integration methods can be implemented in GEMPACK via the RunGTAP and RunGEM programs. Details about using RunGem are documented in the GEMPACK Manual in chapter 45 (<https://www.copsmodels.com/gpmanual.htm>) while instruction for accessing the help files (distributed with the RunGtap program) for using RunGtap are here: <https://www.copsmodels.com/webhelp/rungtap/>.

Horridge and Pearson (2011) highlight the above limitations and offer a way for users to carry out SSA with respect to both parameters and shocks simultaneously using the existing RunGTAP or RunGEM setup for models implemented in GEMPACK. However, this approach requires the user to modify the TABLO model files and moreover, continues to be based on numerical integration methods thus not allowing the flexibility to specify the number of model solves to be carried out.

More recently, Villoria (2017) demonstrates the capability of the R package, *rgtap*, for conducting SSA using Monte Carlo methods by way of these examples¹:

- Change parameters (Armington Elasticity of Substitution) N times given an experiment where the exogenous variable TMS is shocked.
- Varying a global productivity shock, “aoall” based on a multivariate normal distribution for 3 modelled regions.

¹ Both serial and parallel processing versions are discussed for the first of these examples.

Chatzivasileiadis et. al. (2019) compare the sensitivity of CGE model results using both Monte Carlo and Gaussian Quadrature methods to conduct SSA. They model the impact of sea level rises using the GTAP model and provide Stata code for the specific purpose of replicating the results of their study.

2.2 Existing SSA tools for GTAP in GEMPACK

The most up-to-date, and readily available tool for conducting SSA using the GTAP model in GEMPACK is the one embedded in the package RunGtap/RunGem. This tool reduces computational burden by using discreet approximations to distributions for exogenous variables of interest and provides a Graphical User Interface for user-friendliness. While a GQ approach and a GUI is convenient in many cases, it can be limiting in more complex scenarios because:

- *A user is unable to vary Parameters or Shocks simultaneously;*
- *For capturing input uncertainty, a user can only choose from the triangular and uniform distributions.*
- *A user cannot select the number of model evaluations to construct a results range thus making Monte Carlo-type SSA infeasible.*

With the *rgtap* package, there does not appear to be a non-R based user-interface or a direct off-the shelf facility to simultaneously shock both parameters and variables². Furthermore, the last significant update to the *rgtap* package appears to have been in 2017.

2.3 Value added of *rssa*

The *rssa* tool resulted from a practical need for an easy-to-use SSA tool that addresses some of the above limitations in the context of AGE modelling conducted in a public institution for policy purposes.

As such, *rssa* offers what we consider are practical advantages over the available SSA tools including in respect of *user-friendliness*:

- No need for users to modify the underlying GTAP model code.
- An easier and efficient way of handling outputs in comparison to the RunGTAP SSA tool.
- Minimises the need to manipulate R code, as seems to be the case with *rgtap*.
- An excel based interface for specifying SSA requirements.

In addition to user-friendliness, *rssa* also offers the following *analytical flexibility*:

² Theoretically, it may be possible to refit the *rgtap* package to enable simultaneous modelling of shocks and parameters and if this is the case, it is unclear whether this would involve a significant overhaul or a series of cosmetic changes.

- Vary both parameters and shocks simultaneously without having to modify underlying model files.
- Choice of Normal and Uniform distributions for capturing input uncertainty or with some manipulation of *rssa.R*, any other distribution including empirical distributions.
- Set, within limits, the number of simulations (sample runs) to be carried out to reflect input uncertainty.
- In models with higher dimensionality, a large sample size can significantly increase time and computational resource- *rssa* mitigates this somewhat by using parallel processing functionality available within R thus allowing a user to select a high sample number, taking the SSA towards a Monte Carlo approach.

Also, being based on an open-source language, our work could provide a route to community-driven enhancements - especially in relation to the user-interface, output handling and performance/solve times.

There are some drawbacks too:

- Some familiarity with R, though not essential would be helpful.
- The user-interface is not point-and-click.
- While the user may choose a high sample number due to parallel processing, available memory space on PCs might limit this though, for the examples presented in this paper, we did not encounter any issues for a sample size of 5000.

3. Details of *rssa*

3.1: Logic of *rssa*

The logic of *rssa* (kept simple to provide flexibility) for conducting SSA with GEMPACK based GTAP-type models is:

1. For N sample runs, we create a N new parameter files populated with N sample values of the chosen parameters drawn from a uniform/normal distribution.
2. For N sample runs, we create N CMF files populated with N sample values of the shocks of interest drawn from a normal/uniform distribution.
3. Run N simulations (using parallel processing) for each pair of parameter and CMF file associated with each sample
4. Gather and report sample results in Excel

3.2: System Requirements

3.2A: GEMPACK License

The key requirement and admittedly one that may entail some cost to some users is that they have an Executable Image and associated files of their model³.

For users who have a source-code license (<https://www.copsmodels.com/gpsc.htm>) and have a suitable Fortran Compiler, it is relatively easy to create an Executable image of their model (.tab file) using a batch file provided in the rssa package.

Our understanding of GEMPACK licensing requirements is that users without a source-code license *may* be able to obtain an Executable image and associated files from a Source-Code license holder *in certain situations*.

As a source-code license is usually a site wide license, it may be that users without a source-code version may be part of a site that has a source-code version and license – in this instance, users with the SC version can distribute the Executable image and related files of the relevant model to users with Executable image version which would then allow them to run rssa. See: <https://www.copsmodels.com/gpeisc.htm>.

In other instances, the most relevant information is in row 2 of the table 1 below (reproduced from the CoPS website - <https://www.copsmodels.com/gemfeat.htm>). This seems to suggest that, GEMPACK non source-code licensees are able to solve any-size model using a model-specific TABLO-generated EXE program and related files distributed by a source-code licensee subject to limitations as set out at <https://www.copsmodels.com/gpdist.htm>.

³ These are files created by TABLO and Fortran Compiler and include: .exe file, .axs file, .axt file, .tab model file associated with the .exe file, .inf file which has model set-up information, .sti file and .min file.

Table 1: summary of GEMAPCK versions

Summary of GEMPACK Versions					
The table below summarizes features of different versions of GEMPACK (more information here). For pricing see GEMPACK prices . New GTAP users should read this page .					
	No Licence	Introductory or Large Simulations Licence	Limited EXE-IMAGE Version and Licence	Unlimited EXE-IMAGE Version and Licence	Full Source-Code Version and Licence
Solve medium-size model using model-specific TABLO-generated EXE program (see notes 1 and 2)	yes	yes	yes	yes	yes
Solve any-size model using model-specific TABLO-generated EXE program (see note 1)	no	yes	yes	yes	yes
Create and solve medium-size model using GEMSIM (see note 2)	no	no	yes	yes	yes
Create and solve any-size model using GEMSIM	no	no	no	yes	yes
Create and solve any-size model using model-specific TABLO-generated EXE program	no	no	no	no	yes
Create and distribute model-specific TABLO-generated EXE program	no	no	no	no	yes
Requires Windows PC?	yes	yes	yes	yes	Windows or Linux
Requires FORTRAN compiler ?	no	no	no	no	yes
For fuller info, see:	-	Introductory licence	EXE-image licence	EXE-image licence	Source code licence

Note 1: Only source-code licensees can **create** TABLO-generated EXE programs. Such programs may be freely distributed to others, but they will only run large models if the user has some type of GEMPACK licence ([more details](#)).

Note 2: Here "medium" means limited by the sizes spelled out [here](#). Included amongst allowable "medium-sized" models are ORANI-G with up to 35 sectors and GTAP with up to 12 regions and 12 tradeable commodities.

Source: <https://www.copsmodels.com/gemfeat.htm> as at 12th April 2023.

In any case, for users without a Source-Code licence, we would recommend consulting CoPS about the feasibility of obtaining an Executable Image of the model of interest and related files based on their GEMPACK licence, if any.

3.2B: Software

- I. Operating system; Windows [64-bit]
- II. R
- III. R studio or any other text editor suitable for R
- IV. GIT – needed for downloading R packages on Github
- V. R packages – tidyverse, data.table, openxlsx, parallel, devtools, HARr and this.path
- VI. GtapAgg2
- VII. ViewHar

3.3: Conducting SSA with *rssa*

Assuming that the necessary software is downloaded, we set out step-by-step instructions for running *rssa* below in two parts:

Part A provides instruction for users with a Source-Code GEMPACK version & licence.

Part B provides instructions for users without SC GEMPACK version & license. It assumes that they have obtained the Executable image and associated files as per the conditions of the GEMPACK licensor.

Some steps (non-GEMPACK related) will be common to both sets of users.

Please note that *rssa* is set up such that a user only needs to interact with an excel front-end for:

- a) Determining the number of simulations;
- b) Selecting variables and parameters to vary;
- c) Choosing between a normal or uniform distribution relevant to b) above and
- d) Controlling the output of solution variables.

Apart from downloading the *rssa* package zip folder and ensuring that the required GEMPACK files are in place, for routine SSA, there is no need for the user to interact with either R or TABLO files.

3.3A: for users with Source-code GEMPACK version and License

Step 1. Create a top-level project folder (referred to as “top folder” henceforth for ease of reference) within windows file explorer – you are free to choose its location and name with one qualification – **the name should NOT contain any spaces.**

Step 2. Download the *rssa* package folder named GTAP_MC.zip from the gtap website in the same location as this paper and extract all its contents to your top folder created in step 1.

Step 3. Confirm that the MC top folder has,

These sub-folders: <<Inputs>> , <<Data>> , <<Log>>, <<Scenarios>>,
<<Reports>> and <<Solutions>> *plus*

These files: <rssa.R> , <runmc.bat> & <tablo_setup.bat> and

Optional files included for ease: <gtapv7.sti>, <gtapv7.tab>

Furthermore, confirm that the <<Inputs>> folder has an excel file named
<mc_input_template.xlsx>.

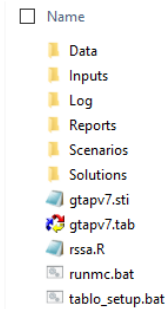
Step 4. Please copy these files to the top folder:

TABLO model file with the suffix **.tab** for your model.

The TABLO Stored Input File with the **.sti** suffix relevant to your model.

The standard GTAP model (v7) .tab and the associated .sti file are both available for download from: [GTAP Resources: Resource Display: GTAPv7.tab for version 7... \(purdue.edu\)](#) and are included in the zip folder for convenience. Replace these files with the tab and sti files associated with the model you are working with if different from gtapv7.

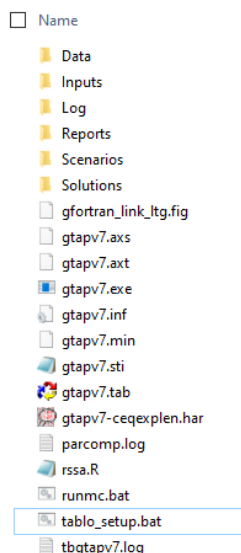
Your <<Top Folder>> should look something like below:



Step 5. Assuming you have a source-code version of GEMPACK, run the file <tablo_setup.bat> - if you saved the gtapv7.tab and gtapv7.sti as your model files and if the batch file has run successfully, your <<Top folder>> should now have the following *additional* files:

<i>gtapv7.axs</i> <i>gtapv7.axt</i> <i>gtapv7.exe</i> <i>gtapv7.inf</i> <i>gtapv7.min</i>	<i>gfortran_link_ltg.fig</i> <i>gtapv7-ceqexplen.har</i>	<i>tbgtapv7.log</i> <i>parcomp.log</i>
---	---	---

The “top folder” should now look like below:



Step 6. In the <<Inputs>> folder, take a copy of mc_inputs_template.xlsx and save it in the same folder as <mc_inputs.xlsx> - the filename has to be exactly as specified here and is case-sensitive⁴.

Then choose the inputs in this spreadsheet as per the instructions in the table below. Please also pay special attention to the notes below.

NB1:

The template file is set for the NCOR3x3 aggregation scheme with 3 sectors and 3 regions – clearly this will not apply to whatever aggregation you are using in all but the very exceptional situation. **So, please ensure that all data contained in the “mc_inputs.xlsx” file conforms to your chosen aggregation scheme.**

NB2:

The data input is set up in a way that parameter and variable names are Uppercase in line with the .HAR datafiles in the <<Data>> folder⁵.

NB3:

Currently the *rssa* package is set-up for handling only these parameters:

{SUBP, INCP, ESBT, ESBC, ESBV, ETRQ, ESD, ESBM, ESBQ}

Note that while the dimensions for these parameters can be either COMM*REG or ACTS*REG, we assume that a commodity is produced by only 1 activity and 1 activity produces only one commodity i.e. COMM*ACTS matrix has no non-diagonal entries. So, in the inputs spreadsheet, for the dimension names for these parameters, we select **AR** as the default.

Also, only endogenous variables whose domain is COMM*REG*REG or COMM*REG or ACTS*REG are allowable – again, in the case of the latter 2 items, “AR” is the default for the abbreviated symbol for the dimension names comm/acts*reg.

NB4:

A general point about how the inputs spreadsheet is set-up: all data is arranged in a long format so the early columns refer to the dimension names and total number of rows would equal to the product of dimensions elements. For example, the template file is currently set up for 3 sectors and 3 regions so, the tabs associated with each of the parameter (ESBM in the case of examples 1, 2, and 4 below) will have 9 rows. So, for your specified inputs, please make sure that each row corresponds to the correct dimension name for that input item and the total number of rows corresponds to the product of dimensions elements.

Further details about the spreadsheet are below:

There are 5 sets of tabs that are of interest in the mc_inputs_template.xlsx file:

Tab Set Colour	Tab Names	Purpose
Green	Sel_Vars, SAMP	Choose which parameters and or variables you wish incorporate in your SSA from within the allowable set
Yellow	CMF_START, CMF_INPUTS, SOL_METHODS	Set up the non-shock related relevant text that features in the model CMF file

⁴ You can use a different name but you will then need to edit the *rssa.R* file where the excel input file is referenced.

⁵ If you wish to use lower case letters, please edit the *rssa.R* package related to the HAR read and write statements –the current qualification “toLowercase = FALSE” simply needs to be deleted.

Violet	AOALL, AINTALL, TMS, AMS	Provide details for these variables to allow sampling
Red	ESBM, ESBV, SUBP	Provide details for these parameters to allow sampling

Each tab is discussed in detail in Table 2 below.

Table 2: mc_inputs.xls – instructions for use

Tab Name	Purpose & Action
{SEL_VARS} [columns B to F]	Column B: naming the parameters that you wish to vary - currently only these are allowed: {SUBP, INCP, ESBT, ESBC, ESBV, ETRQ, ESBD, ESBM, ESBQ} – see NB3 above. In column C enter the symbol for dimension names (AR for the allowable set), in column D – the number of dimensions and in column F, either Normal or Uniform. As there are some parameters that do not vary across region in the base data e.g. ESBM, after deciding what parameter you wish to sample from, in Column E, please select from {0,1} for parameters that you wish to NOT vary across regions [0] and those that you do wish to vary across regions [1]. !! Please do not change the titles in row 1.
{SEL_VARS} [columns G to J]	Enter uppercase name of the shock variables from the allowable set in Column G that you wish to sample. In Column H - enter either CSD or AR (abbreviations for COMM*REG*REG and ACTS*REG) – as only endogenous variables whose domain is CSD or AR are allowable. In Column I – enter the number of dimensions Column J - choose Normal or Uniform – case sensitive and first letter is capitalised.
{Sel_VARS}[Columns K to R]	This relate to variables from the .sol files the user wishes to collate and export to excel. In Column K – specify the variable name from the .sol file, lowercase In Column L – the number of dimensions of the chosen variable In Columns M: O – names of the dimensions respectively, uppercase In Column P:R – comma separated name of each element COMM, ACTS, REG, as they apply to the variable chosen, that you wish to report; for choosing all elements you can simply put ALL. Names are case sensitive and should match the name given in your aggregation for commodities, activities and regions. <i>Note: Currently no facility to pull out variable from .slc file but we plan to add this.</i>
{SAMP}[column B]	Set the number of times you want to sample for the chosen parameters and shocks and therefore the number of times the model is run.

	<i>Note: For initial set-up it is recommended that you run the model once with no changes (by setting sd=0 and min and max for uniform variates the same as central values – this should generate solutions that match your central run.</i>
{CMF_START}[Column A]	From row 2, use each row to copy the text that normally goes in your cmf file at the start. The template file as the text relevant to the NCOR3x3 model to give an idea of what is needed here. !!Please do not change row 1 title.
{CMF_INPUTS}[Column B:C]	Enter the list of exogenous variables and any swaps that you have in you have in your central model run – these data can be found in the CMF files associated with your central runs. Note also there are no swaps so column C is empty from row 2 onwards.
{Sol_METHOD}	Self-explanatory
{AMS} {TMS} {AOALL} {AINTALL}	There should be tab for each of the variables you have chosen in {SEL_VARS} [column G]- in the template file we have included 4 tabs for illustration purposes even though for NCOR3X3, we only need TMS. The structure of these tabs should follow these rules: 1 column for each dimension and these start at the left followed by the appropriate values – starting values in Column D, Target Values in Column E and in Columns F, G, the means and SDs if you have chosen Normal in [SEL_VARS]Col J or min and max boundaries if Uniform. For the NCORS3X3 experiment, the relevant .cmf is “TMSeufG.CMF”; note that in this file, the shock values come from the tms.shk file but essentially there is only one line for the simulation and this is shown in the mc_inputs_template.xlsx. !!Please do not change row 1 titles that refer to the dimensions. The names for mean, sds, min and max can be changed but the order has to remain the same.
{ESBM, ESBV, SUBP}	The same instruction apply as in the case of shock variables above except that there is no target value – all sampling is done on the value as specified in your dataset – the row 1 titles have a suffix “_bmk” to refer to these baseline values in column D for example. Columns E, F – mean/sds for Normal or Min, Max for Uniform. !!Please do not change row 1 titles that refer to the dimensions. The names for mean, SDs, min and max can be changed but the order has to remain the same.

Step 7. In the <<Data>> folder place the following files for your chosen aggregation scheme of the GTAP database appropriate to your model – these files are usually output in a zip folder when using GTAPAGG2 for aggregating the GTAP database:

basedata.har
 baserate.har
 default.prm
 sets.har

CMFSTART – for code lines to go into the CMF file usually at the start. See table above.

Use the package Viewhar to open the default.prm file and then save it as default.HAR file in the same <<Data>> folder. In summary the <<Data>> folder should have 6 files in total.

The data files associated with the NCOS3X3 model are pre-loaded In the Data folder of GTAP_MC.zip.

Step 8. Running the the MC program – here there are two options.

- i) Open the rssa.R via Rstudio or similar and run the it from that package; or
- ii) Run via the <runmc.bat> batch script – this looks for the path to where your R executable file is located and can sometime take a while so we'd recommend using option 1.

If running via the batch file, a CMD box will show up and though it may appear that nothing is happening, the processing is happening so wait until you get a flashing > at which point if you click the space-bar, the CMD box will close.

The table below gives and explanation for each of the sub-folders in the top folder and its purpose including the Solutions folder.

Folder Name	Purpose
Log	The log files associated with each sample run is saved here – names follow this pattern: log_mc <i>x</i> where <i>x</i> is sample number
Reports	This is only relevant to GTAP in GEMPACK and essentially is repository for the various Decomposition, Summary and Volume files associated with each sample run. The names of the files are self-explanatory. For the moment we do not save updated Data files to manage memory.
Scenarios	Saves three template files for cmf lines that are not shocks; The main cmf files associated with each sample run is saved in the subfolder: MC_SSA_CMFS and file names are self-explanatory
Solutions	Has two folder for saving the .SOL and .SLC output files associated with each sample run. Folder and file names are self-explanatory. There are also 3 spreadsheets that collects the user specified solution variables for each sample and is named: ssa_report_[x]d where <i>x</i> can be {1, 2, 3} referencing the output variable's dimension.
Data	A new <i>sub-folder</i> called MC_SSA_HARS is created and contains the parameter files for each sample.

3.3B: for users without Source-code GEMPACK version and License

Steps 1 to 3 are same as in section 1.

Step 4 is redundant.

Step 5: You will need to obtain the following files from a user with Source-Code GEMPACK version & license [the two log files are not needed]:

- gtapv7.axs*
- gtapv7.axt*
- gtapv7.exe*
- gtapv7.inf*
- gtapv7.min*
- gtapv7.sti*
- gtapv7.tab*
- gfortran_link_ltg.fig*
- gtapv7-ceqexplen.har*

Steps 6 and 7 are the same as in section 1.

4. Illustrative Applications

We use the *rssa* tool following the instruction above to model four example experiments.

4.1: Example 1

NCOR3x3 model – remove tariffs on Food exports from Sub-Saharan Africa to the EU.

We assume that ESBM is variable as captured by a Normal distribution with mean=baseline value of ESBM and sd=2.

We run 5000 simulations and pull out solution variables: qo, po and qxs.

Firstly note that we need two additional tabs over and above the Green and yellow set of tabs - these are TMS and ESBM – this is in line with what we have chosen in the SEL_VARS tab as shown below.

Note that even though TMS is not being sampled, we assume that it is but then make it invariant by setting the lower and upper limits the same as the target value.

Also, note that we are only interested in pulling out qo, po and qxs and so the columns K to R have to reflect this. Finally, to run 5000 simulations, the Tab SAMP needs to be adjusted too.

Excerpts from the relevant tabs are shown below:

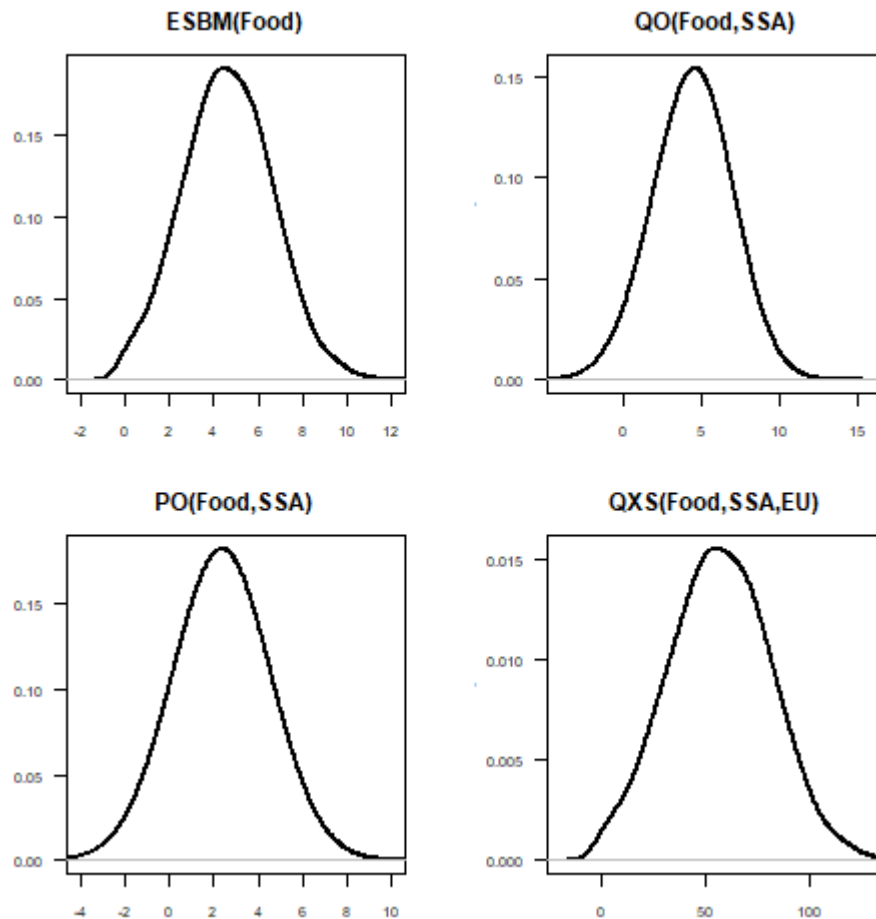
Tab: ESBM						Tab: TMS							
A	B	C	D	E	F	A	B	C	D	E	F	G	H
1	sector	region	item	esbm_bmk	esbm_mean	esbm_sd	1	sector	from_r	to_r	item	tms_s	tms_tgt
2	Food	SSA	ESBM	4.63334	4.63334	2	2	Food	SSA	EU	TMS	13.042218208	0
3	Mnfcs	SSA	ESBM	5.98184	5.98184	0	3					0	0
4	Svces	SSA	ESBM	3.81532	3.81532	0							
5	Food	EU	ESBM	4.63334	4.63334	2							
6	Mnfcs	EU	ESBM	5.98184	5.98184	0							
7	Svces	EU	ESBM	3.81532	3.81532	0							
8	Food	ROW	ESBM	4.63334	4.63334	2							
9	Mnfcs	ROW	ESBM	5.98184	5.98184	0							
10	Svces	ROW	ESBM	3.81532	3.81532	0							

Tab: SEL_VARS									
A	B	C	D	E	F	G	H	I	J
1	id	par_name	par_dimt	par_dimn	par_secvar_flag	par_dist	cmf_var_name	cmf_var_dimt	cmf_var_dimn
2	1	ESBM	CR	2	0	Normal	TMS	CSD	3
3	2								Uniform
K	L	M	N	O	P	Q	R		
	solvar_names	dimnum	dim1	dim2	dim3	dls1	dls2	dls3	
qo		2	ACTS	REG		ALL	SSA,EU		
qxs		3	COMM	REG		ALL	SSA,EU		
po		2	ACTS	REG		ALL	ALL	SSA,EU	

Tab: SAMP	
A	B
1	ITEM VALUE
2	nsamp 5000

The graphic below shows the values of ESBM in Food and the variability of qo, po and qxs as density plots. Note the purpose of this example is to illustrate the workings of *rssa* and so we do not discuss economic meaning or intuition.

Chart 1: Zero tariffs on SSA exports of Food to the EU with variable ESBM



4.2: Example 2

NCOR3x3 model – Lower iceberg costs on Food Exports from Sub-Saharan Africa to the EU-shock AMS by 10.

ESBM varies as a Normal distribution with mean=baseline value of ESBM and $Sd=2$.

We need two tabs: AMS and ESBM and we need to select both ESBM and AMS in the SEI_VARS tab.

Again, note that even though we are not sampling AMS, in order to have the shock in place, we work as if we are shocking AMS but then impose $sd=0$. This is the same as example 1 where TMS was deemed to be sampled but invariant.

Excerpts from the relevant tabs are shown below:

Tab: ESBM

	A	B	C	D	E	F
1	sector	region	item	esbm_bmk	esbm_mean	esbm_sd
2	Food	SSA	ESBM	4.63334	4.63334	2
3	Mnfcs	SSA	ESBM	5.98184	5.98184	0
4	Svces	SSA	ESBM	3.81532	3.81532	0
5	Food	EU	ESBM	4.63334	4.63334	2
6	Mnfcs	EU	ESBM	5.98184	5.98184	0
7	Svces	EU	ESBM	3.81532	3.81532	0
8	Food	ROW	ESBM	4.63334	4.63334	2
9	Mnfcs	ROW	ESBM	5.98184	5.98184	0
10	Svces	ROW	ESBM	3.81532	3.81532	0

Tab: AMS

	A	B	C	D	E	F	G	H
1	sector	from_r	to_r	item	ams_s	ams_tgt	ams_mean	ams_sd
2	Food	SSA	EU	AMS	0	10	10	0
3	Mnfcs	SSA	EU	AMS	0	0	0	0
4	Svces	SSA	EU	AMS	0	0	0	0
5	Food	EU	SSA	AMS	0	0	0	0
6	Mnfcs	EU	SSA	AMS	0	0	0	0
7	Svces	EU	SSA	AMS	0	0	0	0

Tab: SEL_VARS

A	B	C	D	E	F	G	H	I	J
id	par_name	par_dimt	par_dimn	par_secvar_flag	par_dist	cmf_var_name	cmf_var_dimt	cmf_var_dimn	cmf_var_dis
1	ESBM	CR	2	0	Normal	AMS	CSD	3	Normal
2									

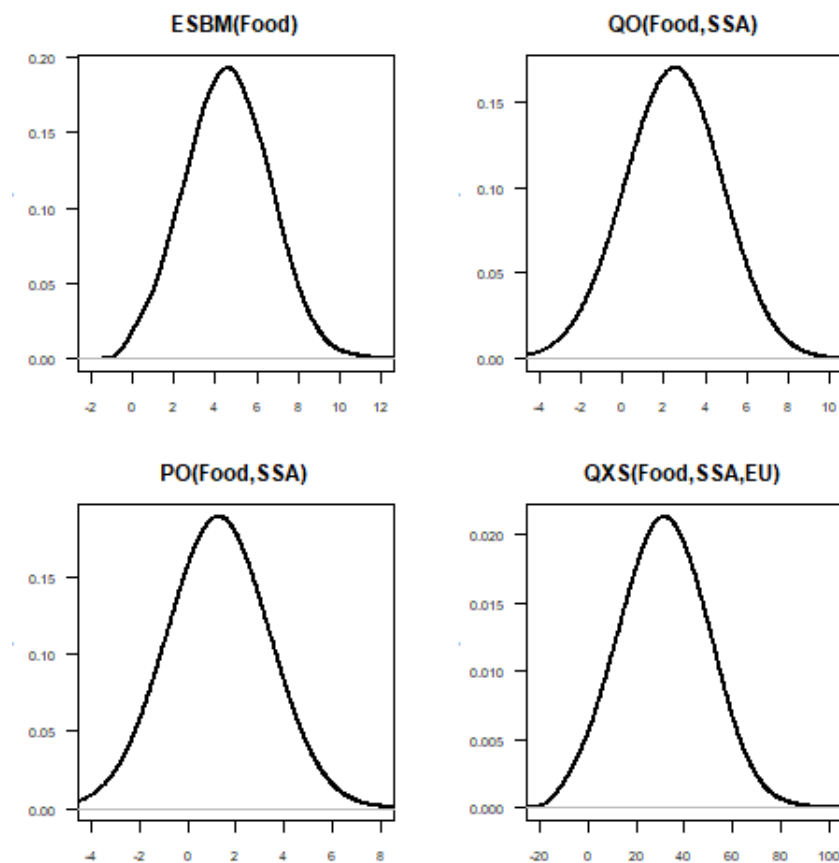
K	L	M	N	O	P	Q	R
solvvar_names	dimnum	dim1	dim2	dim3	dls1	dls2	dls3
qo	2	ACTS	REG		ALL	SSA,EU	
qxs	3	COMM	REG	REG	ALL	SSA,EU	SSA,EU
po	2	ACTS	REG		ALL	ALL	

Tab: SAMP

	A	B
1	ITEM	VALUE
2	nsamp	5000

The density plots for AMS, qo, po and qxs are as below:

Chart 2: Lower Iceberg costs on SSA exports of Food to the EU with variable ESBM



4.3: Example 3

NCOR3x3 model – Remove iceberg costs on Food Exports from Sub-Saharan Africa to the EU
– shock AMS by values from a normal distribution with mean 10 and sd 2.

We need only one additional Tab - AMS & the SEL_VARS tab now has AMS in columns G to J.

As parameters are not being varied there is no need to specify these unlike shocks where even if invariant they have to be specified as in examples 1 and 2 above.

Relevant extracts are shown below:

Tab: AMS

	A	B	C	D	E	F	G	H
1	sector	from_r	to_r	item	ams_s	ams_tgt	ams_mean	ams_sd
2	Food	SSA	EU	AMS	0	10	10	2
3	Mnfcs	SSA	EU	AMS	0	0	0	0
4	Svces	SSA	EU	AMS	0	0	0	0
5	Food	EU	SSA	AMS	0	0	0	0
6	Mnfcs	EU	SSA	AMS	0	0	0	0
7	Svces	EU	SSA	AMS	0	0	0	0

Tab: SEL_VARS

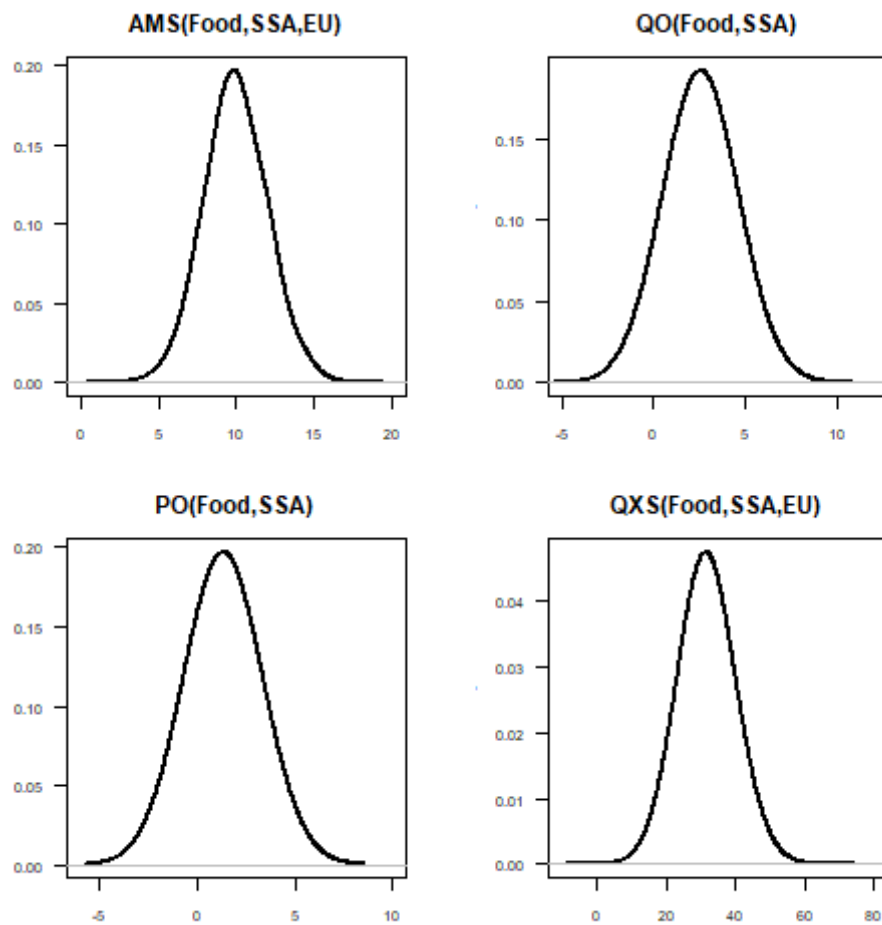
	A	B	C	D	E	F	G	H	I	J
1	id	par_name	par_dimt	par_dimn	par_secvar_flag	par_dist	cmf_var_name	cmf_var_dimt	cmf_var_dimn	cmf_var_dis
2	1						AMS	CSD	3	Normal
3	2									
	K	L	M	N	O	P	Q	R		
	solvar_names	dimnum	dim1	dim2	dim3	dls1	dls2	dls3		
qo		2	ACTS	REG		ALL	SSA,EU			
qxs		3	COMM	REG	REG	ALL	SSA,EU			
po		2	ACTS	REG		ALL	ALL			

Tab: SAMP

	A	B
1	ITEM	VALUE
2	nsamp	5000

The density plots for AMS, qo, po and qxs are as below:

Chart 3: Lower Iceberg costs on SSA exports of Food to the EU with variable AMS



4.4: Example 4

NCOR3x3 model – we combine examples 2 and 3.

Shock AMS by values sampled from a normal distribution with mean= 10 and sd=2 &
Sample ESBM from a Normal distribution with mean=baseline value and sd=2.

We need two additional Tabs -AMS and ESBM and the SEI_VARS tab now has ESBM in columns B to F and AMS in columns G to J.

Relevant extracts are shown below:

Tab: ESBM

	A	B	C	D	E	F
1	sector	region	item	esbm_bmk	esbm_mean	esbm_sd
2	Food	SSA	ESBM	4.63334	4.63334	2
3	Mnfcs	SSA	ESBM	5.98184	5.98184	0
4	Svces	SSA	ESBM	3.81532	3.81532	0
5	Food	EU	ESBM	4.63334	4.63334	2
6	Mnfcs	EU	ESBM	5.98184	5.98184	0
7	Svces	EU	ESBM	3.81532	3.81532	0
8	Food	ROW	ESBM	4.63334	4.63334	2
9	Mnfcs	ROW	ESBM	5.98184	5.98184	0
10	Svces	ROW	ESBM	3.81532	3.81532	0

Tab: AMS

	A	B	C	D	E	F	G	H
1	sector	from_r	to_r	item	ams_s	ams_tgt	ams_mean	ams_sd
2	Food	SSA	EU	AMS	0	10	10	2
3	Mnfcs	SSA	EU	AMS	0	0	0	0
4	Svces	SSA	EU	AMS	0	0	0	0
5	Food	EU	SSA	AMS	0	0	0	0
6	Mnfcs	EU	SSA	AMS	0	0	0	0
7	Svces	EU	SSA	AMS	0	0	0	0

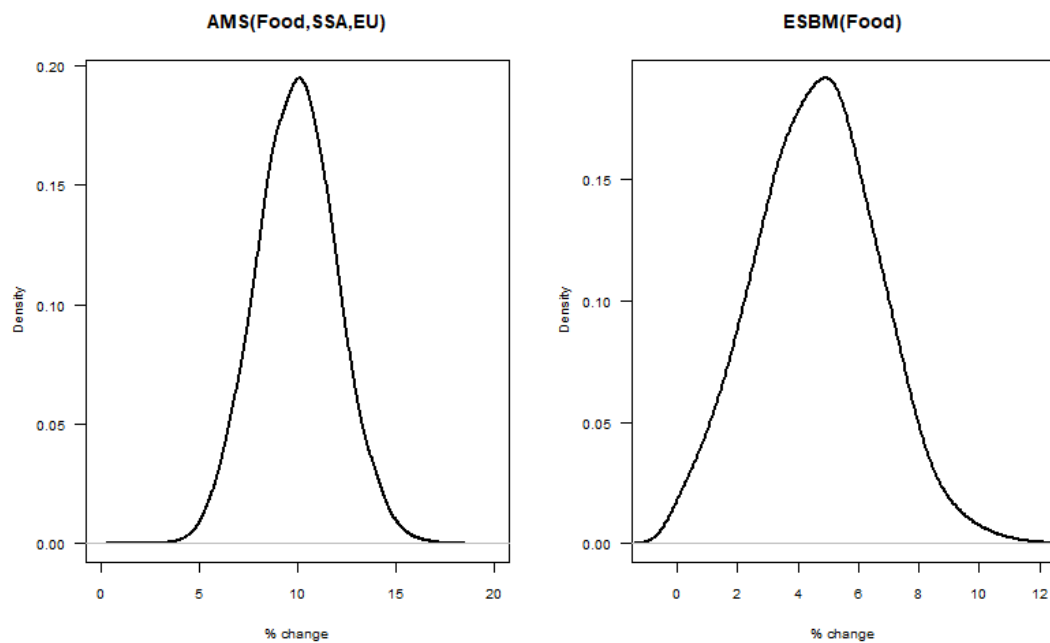
Tab: SEL_VARS

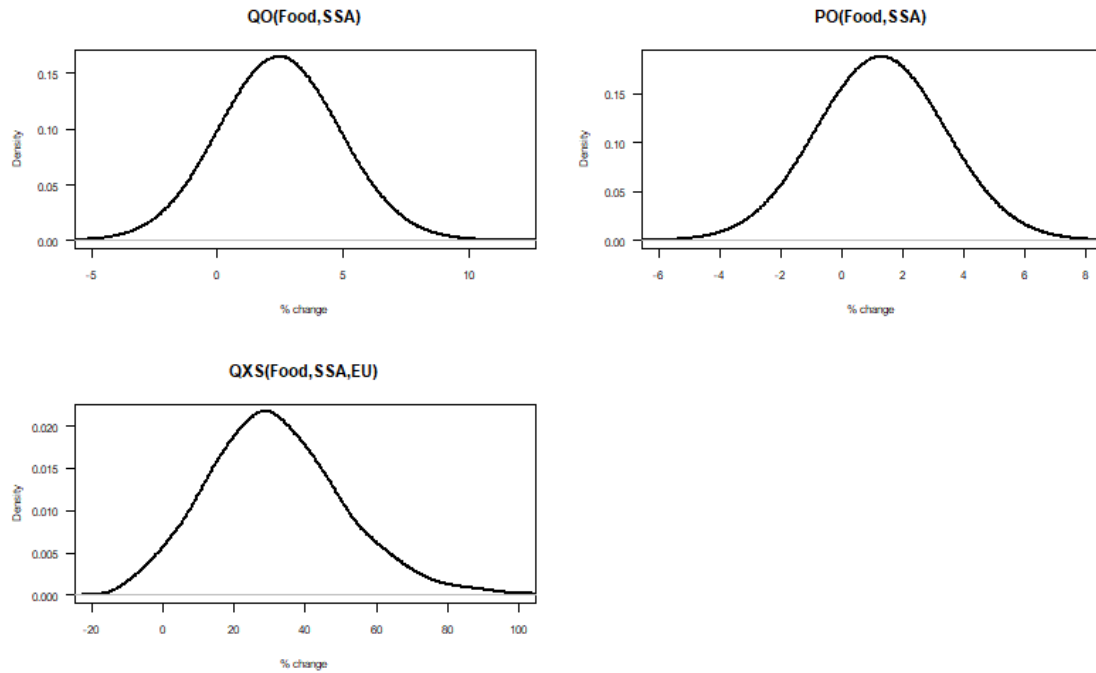
A	B	C	D	E	F	G	H	I	J
id	par_name	par_dimt	par_dimn	par_secvar_flag	par_dist	cmf_var_name	cmf_var_dimt	cmf_var_dimn	cmf_var_dis
1	ESBM	CR	2	0	Normal	AMS	CSD	3	Normal
2									
	solvar_names	dimnum	dim1	dim2	dim3	dls1	dls2	dls3	
qo		2	ACTS	REG		ALL	SSA,EU		
qxs		3	COMM	REG	REG	ALL	SSA,EU		
po		2	ACTS	REG		ALL	ALL	SSA,EU	

Tab: SAMP

	A	B
1	ITEM	VALUE
2	nsamp	5000

The density plots for AMS, ESBM, qo, po and qxs are as below:





5. Closing remarks

In this paper we have presented a new R based tool, called *rssa*, that lets GEMPACK users perform systematic sensitivity analysis for CGE modelling using GTAP-like models. The tool permits SSA with an excel interface and without having to change underlying model files or editing R files. Thus, we hope it will appeal to a wider user-base including GEMPACK users familiar with using RunGtap but with minimal knowledge of R and TABLO.

Also, users will be able to perform SSA with greater analytical flexibility than readily available SSA tools - vary shocks and parameters simultaneously, choose between a normal and Uniform distributions and specify the number of simulations to perform. This benefit does however depend on having access to the exe image and associated files of their model and sacrificing a point-and-click interface like rungtap. Furthermore, true Monte Carlo type analysis with thousands of simulations may face memory constraints but we hope that this can be addressed as part of the ongoing evolution of *rssa*.

Due to R's open source nature, we believe this tool can be adapted and developed further to enhance user-friendliness, improved control of reported outputs and potentially widening its application to non-GEMPACK models.

References

Chatzivasileiadis, T, Estrada, F, Hofkes, M W and Tol, R S J (2019) Systematic sensitivity analysis of the full economic impacts of sea level rise. *Computational Economics*, 53 (3). pp. 1183-1217. ISSN 0927-7099

Horridge, M., & Pearson, K. (2011). "Systematic Sensitivity Analysis (SSA) with Respect to Correlated Variations in Parameters and Shocks" (Presented at the 14th Annual Conference on Global Economic Analysis, Venice, Italy). Purdue University, West Lafayette, IN, Global Trade Analysis Project (GTAP).

Pearson, K., & Arndt, C. (2000). "Implementing Systematic Sensitivity Analysis Using GEMPACK." Purdue University, West Lafayette, IN, Global Trade Analysis Project (GTAP) Technical Paper No. 03.

Villoria, Nelson (2017). "R Meets GEMPACK for a Monte Carlo Walk." *Journal of Global Economic Analysis*, Volume 2 (2017), No. 2, pp 128-154.