SUBSIM 2.0: A QUICK START

The objective of this note is to provide a quick introduction to the use of SUBSIM 2.0. SUBSIM is a SUBsidies SIMulation toolkit programmed in Stata and designed for:

- 1) Distributional analysis of subsidies;
- 2) Simulations of subsidy reforms;
- 3) Simulations of cash transfers as compensation for subsidy reforms.

The annex to this note provides the basic formulae behind SUBSIM.

For any problem concerning installation or use of SUBSIM, please contact Abdelkrim Araar (araar.abdelkrim@ecn.ulaval.ca) or Paolo Verme (pverme@worldbank.org).

INSTALLING SUBSIM

You can install SUBSIM directly from the internet or from a zipped folder if you already have a copy of all installation files.

Installation from the internet

set more off net from http://dasp.ecn.ulaval.ca/subsim2/Installer net install subsim2_part1, force net install subsim2_part2, force

Installation from a zipped file

Copy the zipped folder subsim2.zip in the C directory (or in another working directory used by Stata. You can download this folder from: http://dasp.ecn.ulaval.ca/subsim2.zip). Unzip the folder and check whether you have the following files properly installed:

```
c:/subsim2/installer/stata.tocc:/subsim2/ installer/subsim2.pkg;
```

In the Stata command window or in a do-file, execute the following commands:

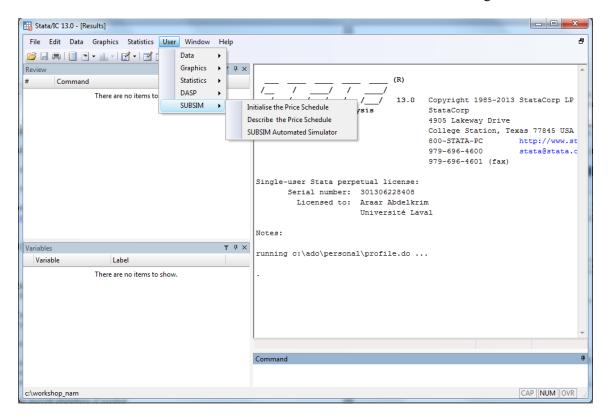
set more off net from c:/subsim2/Installer net install subsim2_part1, force net install subsim2_part2, force

Add the SUBSIM menu to Stata

Add the command "SUBSIM" to your Stata menu. You have three options to do that:

- 1) Add the command "_subsim_menu" to your Stata profile file "c:/ado/personal/profile.do". This will ensure that Stata loads SUBSIM each time that is launched.
- 2) If you don't have the "profile.do" file, you can create one by simply saving the command "_subsim_menu" in a do file with the name "profile.do" in the directory "c:/ado/personal/".
- 3) Alternatively, execute the command "_subsim_menu" when you open Stata and every time you wish to use the SUBSIM menu.

You will now be able to see SUBSIM under the User's menu as shown in the figure below.



EXAMPLES

For the examples, you will need to download first the data set and the examples files from the following website:

http://dasp.ecn.ulaval.ca/subsim2/examples.rar

EXAMPLE 1: LINEAR SUBSIDIES

The following examples are based on the dummy data "example.dta" provided with the toolkit. It is recommended to run the example with the data provided before testing SUBSIM with other data. This will ensure that SUBSIM has been correctly installed.

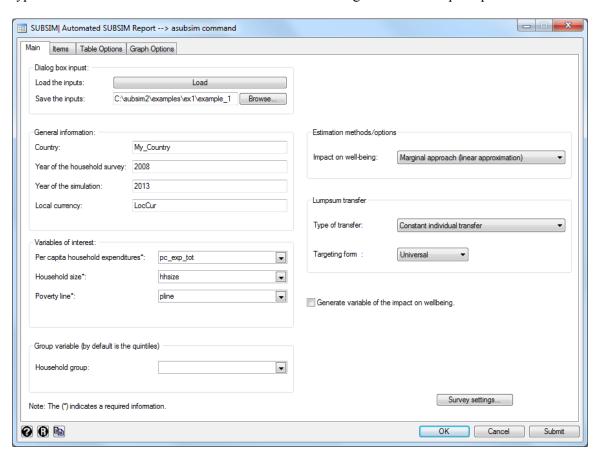
Suppose that our country of interest subsidizes two goods, flour and rice, and that we wish to simulate the impact of a subsidy reform (price increase) on wellbeing and government revenue.

Starting from the Stata main menu, open the menu:

User/SUBSIM/SUBSIM Automated Simulator

or

type "db asubsim" in the command window. The following window will open up.



The SUBSIM interface window includes four tabs:

- Main
- Items
- Tables options
- Graph options

The tabs "Main" and "Items" are designed for data inputs. The tabs "Tables options" and "Graphs options" are designed to select outputs options.

<u>Inputs that are compulsory for the simulations are indicated with an asterisk *.</u>

TAB: "Main"

The Tab Main contains six boxes for data input:

Dialog box input. Load and save input data. The box enables the user to load information already saved into the SUBSIM window or to save the information inserted in the dialogue box in a file to be stored for future simulations. This information is stored in text files with the extension *.prj. You can test this feature by uploading the file "example_1.prj" provided with the toolkit. Note that you can load the file from one directory ("Load the Inputs") and save it in a different directory with a different name ("Save the Inputs").

General information. The group box General information enables the user to insert some helpful information, like the name of the country or the local currency. This information will be saved in the file of results. This is important to remind the user of the basic background information about the simulation as this information is displayed and saved in the excel file of results.

Variables of interest. The box Variables of interest enables the user to insert key variables such as the per capita expenditures/income, the household size, and the poverty line.

Note that the key income variable should be in per capita terms.

Group variable. The box Group variable enables the user to insert a population group variable. This variable captures a socio-demographic group such as gender or urban-rural. By default, results are shown by quintile. When you select a different group variable the results will be shown by this variable. Note that only one variable can be chosen for each simulation. If results are needed by more than one variable the user will have to re-run SUBSIM each time.

Estimations methods/options. The group box Estimation method/Options enables the user to select different modelling estimation options. This concerns the selection of the approach to be adopted to assess the impact on wellbeing. In addition to the popular marginal approach, we propose to use a simple and popular form to model the consumer behavior which is the Cobb-Douglas function. Note that the impact on wellbeing is measured with the equivalent gain. For more information, see the Annex.

• The marginal approach

• The behavioral approach (Cobb-Douglas Utility Function)

Lumpsum transfer.

Type of transfer: The box "Lump-sum transfer" enables the user to indicate information on cash transfers (lump-sum transfer) and choose whether this transfer should be allocated to individuals or households. Further, the user can indicate a dummy variable to constraint the group that benefits from the transfer. Since the results are reported in graphs, the user can select a range of values of transfers (the min and the max). There is no need to indicate the amount of transfer because the analysis will cover all amounts included between the minimum and the maximum specified.

Targeting form: In this box, it is possible to indicate the group targeted by transfers. The user only needs to prepare in advance a dummy indicator variable for the group. For example, poor=1 and non-poor=0 if the poor only should be targeted. If the group is not specified, SUBSIM will produce the transfer graphs with universal transfers.

NB: Survey settings. Remember to set the survey settings before you launch SUBSIM including sampling weights and sampling design information. This can be done with the command "svyset" in Stata or you can use the button "Survey Settings..." located in the bottom right-hand corner of the SUBSIM "Main" tab. For more information on survey settings, see the Stata manual.

TAB "Items"

The Tab "Items" is conceived to insert information about the goods concerned by the simulation including initial prices, final prices and unit subsidies.

Initialize information with: The information on products can be initialized by manually inserting the information for each item (option "parameters values") or by selecting variables already created and available in the data set (option: "variables"). The "example.dta" already contains these special variables.

Number of items. The figure bellow shows an example concerning two goods: Flour and Rice. The user should start by selecting the number of items, which is two in our example.

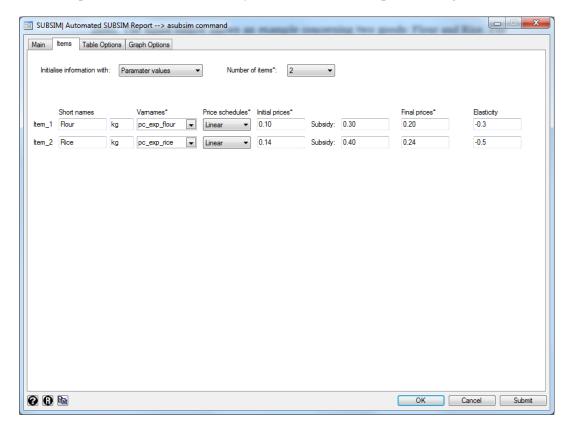
Option "Parameter values": With this option the user can input the information for each item manually including name, quantity, per capita expenditure variable, type of price schedule, initial price, unit subsidy, final price and elasticity.

Name and unit quantity: In the first two boxes the user can insert the name of the item and the unit quantity (kg, liters, etc). This information will be displayed in the results tables as headings.

Linear and non-linear prices. Users have an option to choose linear and non-linear prices. This refers to whether the price is equal for all quantities consumed by households or changes according to quantities. This is the case, for example, of a quota system where

households are entitled to subsidies prices only up to a certain quantity (quota). In this first example, we only consider linear prices. In Example 2 we will consider non-linear prices.

Prices. We are then ready to input the key prices (initial and final) and the unit subsidy. For example, for flour we assume that the actual initial price is 0.1 monetary unit and the unit subsidy is 0.3. In the absence of subsidies, the price of flour would be 0.4. We can simulate however any increase in price. For example, an increase in prices of 0.1 which leads to a final price of 0.2. In this case, our inputs will be 0.1 for the initial price, 0.3 for the unit subsidy and 0.2 for the final price. For rice, we can input, as an example, 0.14 for the initial price, 0.4 for the unit subsidy and 0.24 for the final price (see figure below).



Elasticity. This is the own-price (quantity/price) elasticity. The user can insert any value and this is used to estimate changes in quantities consumed and other impacts. See the Annex for a discussion on how to specify the value of elasticity.

Option "Variables": With this option, the user can select the data on products by selecting variables directly from a pre-prepared data set. This option is suitable when the user has a very large number of items so that it may be easier to prepare first a spreadsheet with all the key information including names of items, prices, units and elasticities. SUBSIM allows the user to upload this information and use it for the analysis. Note that the spreadsheet has to contain all the information needed for the analysis in the form of variables. This option will be treated more in detail in Example 3.

TAB "Tables options"

This tab allows the user to select the tables' options. The default option when you do not select the tables and override options is the production of all tables.

Tables: Select the tables to be produced. In case the user wishes to have only a selected number of tables the code of these tables can be indicated in the box. The list of codes with the titles of the tables can be seen by clicking on the question mark button ? For example, you can type "11 23" to produce tables 1.1 and 2.3 only (no commas, one space between numbers).

Override the default list of items. If the user wants to aggregate results for several products this is possible by indicating the codes of the products to aggregate and the name of the new aggregated item. For example, you may want to aggregate the results for various types of sugar (items 4, 5 and 6) and various types of flour (items 7 and 8). Or you may want to add results for rice and flour. This may be done by adding the option: "1 2: "Rice and Flour".

Excel file: Produce an Excel file of results. This box allows the user to define the Excel file where all tables should be stored. The user can select an existing file to override or create a new file. The user can either specify the name of the file or not. If the name is not specified, the file will be saved as "Tables" by default.

Note that, in the case of an existing file, the user should make sure that this file is closed when the program is launched, otherwise an error message will appear.

Language: Users can choose the language for all results. English and French are the two languages currently available.

TAB: "Graph options"

Graphs: Select the graphs to be produced. This option allows the user to save only selected graphs by indicating the code of each graph. The list of codes with the titles of the graphs can be seen by clicking on the question mark button ? For example, if the user wishes to produce only Graphs 1, 2 and 4, the user will simply type "1 2 4" (no commas, one space between numbers).

Select the folder of graphs results. This option allows the user to select the directory where the saved graphs should be stored. Note that all graph files are saved in three formats: .gph. .pdf and .wmf. SUBSIM will save a folder with the name "Graphs" in the directory selected.

Graph options. For each graph, the user can select options regarding the y-axis scale (min and max) and other two-way graphs options as indicated in the Stata graph help files.

EXAMPLE 2: NON-LINEAR SUBSIDIES

By non-linear subsidies we mean to describe subsidies that change according to different levels of quantities consumed by households. The case of non-linear subsidies is typically of two forms:

Quota system

Blocks system

The **quota** system refers to subsidies administered via quotas. For example, households may be entitled to a subsidized price for bread up to a certain quantity purchased - say - 10 kg./month. Beyond that quantity, consumers buy bread on the free market at unsubsidized prices. This system usually makes use of "cards" where households are given ratio (quantity) cards that allow them to purchase certain quantities at subsidised prices.

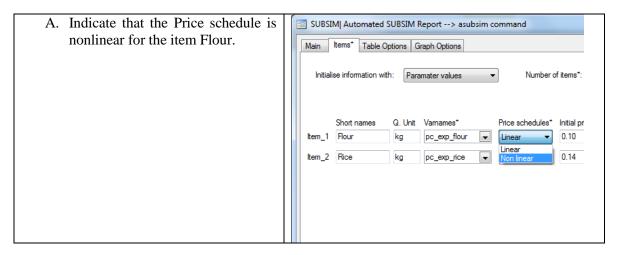
The **blocks** system refers to subsidies that change following a "block" structure with different prices applying to different blocks of quantities consumed. This is typically the case of electricity or water subsidies where the electricity or water prices are set by the regulator at different prices for each quantity block. For example, a price for a consumption of 0-150 kWh/month, a higher price for a consumption of 151-300 kWh/month and so on. In this case, the number of blocks can be small or large depending on the choice of the regulator.

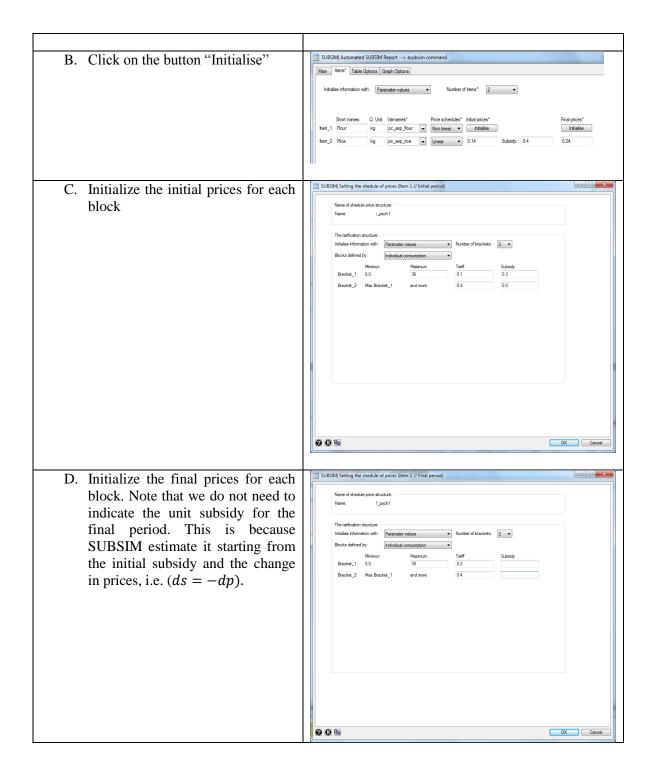
From an economic and modelling perspective the quota and blocks systems are equivalent. In fact, the quota system can be considered as a block system with only two blocks. Therefore, in what follows, we will limit our example to the quota system but the same explanations apply to the blocks system.

Suppose now that subsidies are administered through a quota system where all individuals are entitled to fixed quantities at subsidised prices. For example, imagine that the annual per capita quota for flour is 36 kg. Assume also that the non-subsidized market price is equal to 0.4. This implies that the price of flour is non-linear; it changes with different quantities consumed. Consumers pay a subsidized price up to 36 kg per person and the unsubsidized price for any additional quantity purchased. The following table summarizes the flour non-linear schedule price.

Block	by	Subsidy	Price
0-36 kg	individual	0.3	0.1
36 kg and more		0.0	0.4

This nonlinear schedule price must be first declared. To this end, the user has to perform the following steps:



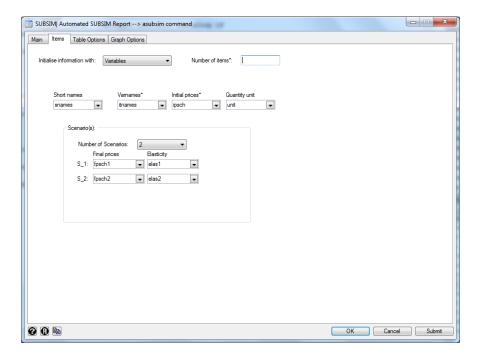


EXAMPLE 3: SIMULATION WITH LARGE NUMBER OF ITEMS

It may occur that the subsidy reform concerns a large number of items that may even exceed 10 items. In this case, the user can insert information on items using variables. This can be done by selecting the option "Variables" from the tab "Items". Note that the spreadsheet has to contain all the information needed for the analysis in the form of variables as shown in the example below.



Once the data are uploaded into STATA, the user can draw from the spreadsheet by using the items dialogue box as shown below (for this example, load the example_3.prj). When the information is uploaded through variables, it is possible to ask SUBSIM to perform the computation for up to three scenarios. For instance, in scenario 1 the reduction in subsidies is 30%, while it is 100% in scenario 2. In this case, the Excel output file will contain estimations for all scenarios.



Launch SUBSIM

You are now ready to launch SUBSIM. When SUBSIM is launched, it will display all results in the Stata output window. The user can stop the command at any stage of execution by using the Stata "Break" button. If the user has selected to save the table results in an Excel file, this file is automatically opened once the computation ends. The excel file produced contains one table per sheet and all tables produced by the program. All graphs produced by the program are instead saved in a folder with the name "Graphs", and in three formats (pdf, gph and wmf).

The complete set of tables and graphs can then be used to prepare a report on the distribution of subsidies, on the impact of subsidies reforms on household welfare and the government revenues and on the impact of compensatory cash transfers on poverty and the government budget. If the user is familiar with SUBSIM and all input information is available, SUBSIM will produce results in a few minutes and a full report can be prepared in a few days. Moreover, all the data input are saved by users in a file with the .prj extension. This allows for an easy update or reproduction of results at any time.

ANNEX - SUBSIM Basic Formulae

This annex provides a brief introduction to the basic formulae used by SUBSIM. The first version of SUBSIM (SUBSIM 1.0) was accompanied by a full paper: Araar, A. and Verme, P. (2012) Reforming Subsidies: A Toolkit for Policy Simulations, World Bank Policy Research Working Paper No. 6148. The paper includes a general section on subsidies simulations, a section on the economic theory behind SUBSIM and the SUBSIM 1.0 users' guide. The sections below integrate and update the theoretical part of the paper for SUBSIM 2.0.

Changes in welfare

Let w=welfare measure; e=expenditure; p=price, the subscript 1 representing the subsidized product and the subscript 2 representing the bundle of all other consumed products. The change in welfare due to an increase in the price of product 1 (reduction in subsidies) can be estimated as:

$$\Delta w = -e_1 dp_1$$

where dp represents the relative price change $(\Delta p_1/p_1)$. This is the most popular method to estimate changes in welfare subject to changes in prices and is the same approach proposed by Coady et al. (2006) among others.

Note that this formula applies with any behavioral response on the part of households including changes in quantities consumed of the subsidized products or substitution of the subsidized product with consumption of other products. This means that the use of elasticities in SUBSIM does not affect the estimation of the impact of subsidies reforms on household welfare. Households can reorganize consumption as they wish but the impact on total household welfare will not change.

In the case of multiple pricing of the product considered (for example electricity with different tariffs for different quantities consumed) the formula for the changes in household welfare is as follows:

$$\Delta w_h = -\sum_{h=1}^B e_{1,h,b} dp_{1,b}$$

where b represents the blocks and h households. The sum across households represents the total change in welfare.

SUBSIM 2 also allows to model household behavior using a Cobb-Douglas function. In the case of multiple pricing of the product considered the formula is as follows:

$$\Delta w = e_{1,h} \left(\frac{1}{\prod_{m=1}^{M} \varphi_{m,h}^{\alpha_{m,h}}} - 1 \right)$$

Where $\varphi_{m,h}$ is the average weighted post reform price (the post reform price in the linear case) of household h for the good m and $\alpha_{m,h}$ is the expenditure share of household h for the good m.

The marginal approach is the most common method and it is usually accurate for small or moderate price increases. For very large price increases, the marginal approach tends to overestimate the welfare impact and it is better to use the Cobb-Douglas approach.

Changes in quantities

Estimates of changes in quantities in the consumption of the subsidized product are useful to have an idea on the impact of the subsidy reforms on quantities consumed and, by consequence, on production of subsidized goods. They are also essential to estimate the impact of reforms on government revenues given that the government reduces expenditure on subsidies when households reduce consumption of subsidized products. Estimates on changes in quantities, in turn, require knowledge of the demand function and the price-quantity elasticity of the subsidized product.

SUBSIM assumes a linear demand function and allows for imputing elasticities. The basic formula for the estimation of changes in quantities of the subsidized product is

$$\Delta q_1 = q_1 dp_1 \varepsilon_1$$

where the own price elasticity ε_1 is typically negative and between 0 and -1. Note that we are assuming that all households behave equally so that the total impact on quantities is just the sum of the changes in quantities consumed across all households.

Elasticity

The formula for the estimation of changes in quantities consumed uses the own-price uncompensated elasticity. One of the main difficulties in subsidies simulations is to specify the value of this elasticity correctly. There are at least three major difficulties.

The first difficulty is that it is very hard to estimate elasticities when products are subsidized. When prices are subsidized and especially when only one price is applied nationally and on all quantities, it is not possible to estimate the own-price elasticity cross-section with a model based on household data (there is no price variation). Sometime, the subsidized price changes over time and one may have available several household consumption surveys that cover the period when price changes occurred. However, this is rare and it is very difficult to isolate the impact of the price change in the subsidized product from other effects on expenditure over time. Therefore, subsidies analysts can rarely estimate elasticities for the country of interest.

The second difficulty relates to the use of known elasticities from the literature and other countries. Sometimes, it is possible to derive elasticity parameters from the specific literature on products. For example, the own-price elasticity for gasoline is quite well known and has been estimated widely worldwide and the user could simply use estimations made for similar countries to the country of interest. However, known elasticities are typically estimated at free market prices and they are point elasticities that apply to prices that are not subsidized. The point elasticities at subsidized prices may be very different and cannot be assumed to be the same. Therefore, it is very difficult for subsidies analysts to simply "borrow" elasticities from elsewhere.

The third difficulty is that the formula presented in the previous section is designed for small changes in prices (marginal changes) and does not function well for large price changes. When the product between changes in prices and elasticity $(dp_1\varepsilon_1)$ is greater than 1, the post-reform quantity can become negative using this formula. Unlike other simulations of price changes, changes in subsidized prices can be very large, especially when governments want to remove subsidies altogether. In these cases, it is not unusual to have price increases of several folds so that dp_1 can be very large. Therefore, subsidies analysts cannot simply use standard parameters for elasticities like -0.3 or -0.5 but have to consider more specifically the relation between subsidized and unsubsidized prices before specifying elasticities.

To overcome these problems, SUBSIM has three main solutions. The first solution is that, by design, SUBSIM does not allow quantities to become negative $(-Q_0)$ because the post-reform quantity has a lower bound of zero. However, one should be aware that when results on quantities in the Excel output file show zero values, it is most likely that the specified elasticities are too large. Subsidized products are usually essential consumption items and it is unlikely that households stop consuming these products altogether if the price increases. It is more likely that our specification of elasticity is incorrect.

The second solution is to use the value of elasticity at unsubsidized prices from another country and derive from this elasticity the correct elasticity to use for the subsidized price. When the subsidized price is several folds lower than the unsubsidized price, this means that the subsidized price is extremely low. But if this price is extremely low and quantity is initially high, we should expect the own-price elasticity to be very low. If prices increase a little around the subsidized price, consumers will tend to reduce quantities by very small amounts. On the contrary, if the subsidized price is very close to the unsubsidized price then it is more likely that increases in prices will lead to large decreases in quantities and that the elasticity will be large. Hence, either the elasticity ε_1 is large or the relative change in price dp_1 is large but they should not be both large at the same time. As a rule of thumb, if the new price is three times the current price and the known elasticity at unsubsidized prices is (say) -0.3, then the elasticity to use in the formula may be around a third of that value, say 0.1.

With the assumption of a straight linear demand function, it is also possible to calculate precisely the initial elasticity (the elasticity at the subsidized price) using the final elasticity (the elasticity at the unsubsidized price). The formula is as follows:

$$\varepsilon_{1} = \frac{\left(\frac{\frac{1}{1-\frac{\varepsilon'_{1}(p'_{1}-p_{1})}{p'_{1}}}-1}\right)}{(p'_{1}-p_{1})}p_{1}$$

The third (and perhaps the most sensible) solution is to run SUBSIM with different assumptions about the elasticity and compare results. In this case, it is useful to use zero as a lower bound and the expected value of elasticity at the unsubsidized price as an upper value. This is what we would recommend especially when price increases are very large.

Changes in government revenues

Having discussed elasticities and changes in quantities, we can now estimate changes in government revenues. We may face two cases, one where we know the unit subsidy and one

where we don't know the unit subsidy in advance. If we know the unit subsidy, the formula is as follows:

$$\Delta r = \sum_{h=1}^{H} e_{k,h} dp_k (1 - \varepsilon_k (s_k - dp_k))$$

where s_k is the unit subsidy for product k. If we don't know the unit subsidy in advance, we can then approximate the change in government revenues with the change in producers' profits as follows:

$$\Delta r = \sum_{h=1}^{H} -e_{k,h} dp_k \left(1 + \varepsilon_k \left(1 + dp_k \right) \right)$$

SUBSIM will use one or the other formula depending on whether users specify unit subsidies or not in the Tab "Items".