SCF Tabling Wizard Instructions

Introduction

The SCF tabling wizard contains a set of macros, written in Visual Basic code, that allow users to create a table (or tables) of weighted descriptive statistics using an Excel extract file of the SCF data (the Excel extract file is documented separately). This wizard allows users to output a table that contains estimates of one of four descriptive statistics (weighted median, weighted mean, weighted percent nonzero, or weighted percent in group) conditioned by as many as two conditions. These statistics are weighted by an analysis weight. While Excel maintains built-in worksheet functions that calculate medians, means, and frequencies (percents), the purpose of this wizard is to allow users to perform weighted mean, weighted median, weighted percent nonzero, and weighted percent in group calculations more efficiently than with the built-in Excel functions. This wizard should increase the accessibility of SCF data to users. Similar Visual Basic code was used to produce the tables included in the February 2009 *Federal Reserve Bulletin* article, "Changes in U.S. Family Finances from 2004 to 2007: Evidence from the Survey of Consumer Finances". Although this software has been tested extensively errors may still remain; the Federal Reserve does not take responsibility for the use analysts make of this software. **Special note: The tabling wizard does not currently work with Excel 2007.**

Opening the SCF extract that includes the SCF tabling wizard

The SCF website includes two data extracts. One extract includes only the aggregate data (e.g. scfp2007.xls); this version is made available to users who do not wish to use the tabling wizard. The other extract includes both the aggregate data and the SCF tabling wizard (e.g. Tabling.2007.pub.xls). To begin using the wizard download the second file and open it in Excel. This file contains macros.

When the file opens, a dialog box warns users that macros may potentially corrupt the workbook, and provides users the option to either enable or disable macros. In order to have access to the SCF tabling wizard users must click “Enable Macros”.

How to make the SCF tabling wizard appear

Opening the file that includes the tabling macro adds three menu options to Excel’s toolbars, which provide users access to the wizard. A button with the appearance of a table appears at the end of the command bar. When the cursor is placed over this button a message appears that reads, “SCF Tabling Wizard”. A menu item named “SCF Tabling Wizard” is appended to the bottom of the Tools menu. Finally, an “SCF Tabling Wizard” menu item is added to the bottom of the right-click shortcut menu. To locate this menu item right-click anywhere on the active worksheet and “SCF Tabling Wizard” will appear at the bottom of the shortcut menu that appears. When users click on any one of these menu items the SCF tabling wizard appears. As noted in more detail later, when Excel closes normally all of these menu items disappear until the next time this file is opened.

Creating User-defined variables and naming these variables

The user can add variables of interest to the worksheet entitled “Data”. These variables will then be used in user-interested calculations. First, make sure that these variables have a name and it appears in the first row of the column that contains the variable. Also, make sure that these variables have the same number of rows as the other aggregate variables. Second, make sure these variables are contained in columns directly next to the aggregate data. In other words, be sure the data are contiguous (no empty columns separate the aggregate data from user-defined data). If both of these conditions are satisfied user-defined variables should appear in the SCF tabling wizard’s lists. To ensure proper analysis of the added variables the user-defined name must begin with at least three

characters, after which any combination of numbers and characters may follow (e.g. “INC1”, “inc1234”). In addition, names that contain more than one word should be connected by an underscore, “_” (e.g. income_class, education_1234). Please contact the SCF Team at scf@frb.gov if any questions arise with respect to user-defined variables

Using the SCF tabling wizard

The SCF tabling wizard is organized in fourteen pages. The first page that appears always asks where you want to output your table. The second page asks which descriptive statistic you would like to include in your table. The subsequent pages that appear are determined by the selections you make. The SCF tabling wizard has many options. Users can select one of four descriptive statistics: weighted median, weighted mean, weighted percent nonzero, or weighted percent in group; all of which may be classified by as many as two conditions at a time. Users can write the requested table to an existing worksheet or to an entirely new worksheet. Once users make all of their selections from these pages and click “Finish” a set of macros calculates all of the requested statistics and places them in a table(s) in the specified worksheet. Please note that each calculation takes a few seconds and computing speed varies by machine. Large tables may take several minutes to produce. The following sections describe each page’s function.

Page 1: Output Worksheet

This page allows users to indicate where to store the requested table. Users can either select to output their table to an existing worksheet or to an entirely new worksheet. To output a table to a new worksheet click the button next to “New worksheet” and enter a valid Excel worksheet name in the box with the header “Type or select worksheet name”. Valid Excel worksheet names may not exceed 31 characters or include any of the following symbols: \ / ? : *. Users may also output their table to an existing worksheet. To output a table to an existing worksheet select the button next to “Existing

worksheet". All valid worksheet names (note: tables cannot be output to the worksheet that holds the data) are included in the drop down box with the heading "Type or select worksheet name". Click the arrow on the right side of this box and select the desired worksheet from the list. The requested table appears at the bottom of the selected worksheet. This page always appears first in the wizard.

Page 2: Weighted Statistics

The SCF tabling wizard allows users to select one of four statistics: weighted median, weighted mean, weighted percent in group, or weighted percent nonzero. It should be noted that the public version of the SCF data are released without any of the data or weight adjustments that are applied in creating the tables in the *Bulletin* article. Consequently, the requested output—particularly weighted means—in some instances may be strongly affected by outliers. Further, many of the aggregate variables included in this data set have nonnormal distributions. For variables with nonnormal distributions the weighted median is a better measure of central tendency than the weighted mean, which is affected by outliers.

The default statistic is the weighted median.

Weighted median:

The weighted median corresponds to the data point that lies at the 50^{th} percentile of the ordered weighted distribution of the variable, range and classification specified. The effects of range and classification restrictions are further explained below.

Weighted mean:

The weighted mean is the summation of the product of the variable and the analysis weight divided by the summation of the analysis weight,

$$\sum_{i=1}^{n} X_i * WGT_i / \sum_{i=1}^{n} WGT_i$$

where X is the variable of interest, WGT is the analysis weight, and i denotes the observation number. As with the median, the mean is computed for the variable, range, and classification specified.

Weighted percent nonzero:

The weighted percent nonzero is the ratio of the summation of weights for those families with non-zero holdings of the variable divided by the summation of all weights,

$$\sum_{i=1}^{n} I(X_i \neq 0) * WGT_i / \sum_{i=1}^{n} WGT_i$$

where I is an indicator function that is set to 1 if X is not equal to zero and is set to zero otherwise. The percent is calculated for the variable, range, and classification specified. The primary use for this calculation is to determine the percentage of non-zero values for a variable of interest.

Weighted percent in group:

The weighted percent in group is the ratio of the summation of weights for those families that meet a user specified range divided by the summation of all weights of the population considered,

$$\sum_{i=1}^{n} I(L \leq X_i \leq U) * WGT_i / \sum_{j=1}^{m} WGT_j$$

where I is an indicator function that is set to 1 if the observation satisfies the upper (U) and lower (L) bound specified by the user. Note that if a lower bound is not specified, the user-specified range is unbounded from below; similarly, if an upper bound is not specified, the range is unbounded from above. If no conditions are specified, the summation of the weights in the denominator is taken over all

observations. Any conditions specified have the effect of limiting the population over which the weights in the numerator and denominator are summed,

$$\sum_{i=1}^{n} I(Cnd = True) * I(L \le X_i \le U) * WGT_i / \sum_{j=1}^{m} I(Cnd = True) * WGT_j$$

where I(Cnd=True) is an addition indicator function that is set to 1 if the observation satisfies the specified condition(s).

Page 3: Base-year dollars

The dollar values in each extract file are given in dollars of the survey year. Thus, all dollar-denominated variables in the 1998 extract file are stored in 1998 dollars. The default settings of the SCF tabling wizard produce dollar estimates in current dollars. The wizard provides users the option to calculate their tables in the real dollars of any one of the six survey years. The calculations included in *Bulletin* articles on the survey are always performed in dollars adjusted for inflation to the year of the most recent survey at that point. If users wish to make estimates comparable to those in the 2003 *Bulletin* article for 1998 SCF data, they would click the button next to “2001" in this page, adjusting the output from 1998 dollars to 2001 dollars.

Page 4: Variables

There are two versions of this page. The first version only appears if weighted median, weighted mean or weighted percent nonzero is selected. Users must select at least one variable from the box in this page to base their calculations on. This box lists all of the aggregate variables in the extract. Properly defined and stored user-defined variables also appear in this box and may be used in calculations. Users may select as many of these variables as they would like to include in their output table (hold in the Ctrl key to select multiple variables).

The second version of this page only appears if weighted percent in group is selected. This page contains a box that lists all of the aggregate variables and user defined variables, a check box that reads “Adjust bounds to base year dollars specified earlier”, a text box for a lower bound specification, and a text box for an upper bound specification. In this page only one variable can be selected from the list. If this variable is a dollar variable, the user may click the check box. Checking the box will adjust the upper and lower bound specifications to the selected base-year dollars; leaving the box unchecked indicates the upper and lower bounds specified are in dollars of the survey year. Finally, users may enter an upper and lower bound in the text boxes provided. Users may also leave either the upper or lower bound unspecified to specify a range that is unbounded from above or below, respectively. If neither an upper nor lower bound is entered the resulting calculations will always equal either 0 or 100 percent no matter what further conditioning is applied.

Page 6: All families

This page appears only if the statistic requested is either a weighted median or a weighted mean. This page allows users to include “all families” in their calculations or condition their output to consider only those families with nonzero holdings of the variable of interest. The statistics included in the *Bulletin* article are based on all families with holdings with the exception of the calculations where the aggregate variables INCOME, NETWORTH, or KGTOTAL are the variables of interest. For these variables, zero is a legitimate value, so it is assumed that all families have holdings. Some variables in this file are held by only a small number of families; note that the weighted median is zero for any variable where less than fifty percent of families have holdings when “all families” are included in the calculation. To include “all families” in the requested calculations users should click the button next to “All families” in this page. To include only families with holdings users should select the button next to “All families with holdings”. The default setting is “All families with holdings”.

Page 7: Select first condition type

This page allows users to select from one of three types of conditions to place on their output. Conditions allow users to compute the statistics for the variable defined in the earlier screen for subsets of the population. Users can condition their output by *Bulletin* categories, the quantiles of a few selected variables, or a user-specific condition. Users can also select to not place another condition on their output. Clicking the button next to "No conditions" activates the Finish button allowing the user to output their table with no further conditions applied to their output.

Page 8: First condition *Bulletin* categories

If you select *Bulletin* categories in page 7, page 8 will appear with a list of all the *Bulletin* categories. There are seven *Bulletin* category groups to choose from: percentile of income, age of head, education of head, race or ethnicity of respondent, current work status of head, housing status, and percentile of net worth.

Percentile of income includes six categories:

INCCAT1: Less than 20

INCCAT2: 20-39.9

INCCAT3: 40-59.9

INCCAT4: 60-79.9

INCCAT5: 80-89.9

INCCAT6: 90-100

Age of head includes six categories:

AGECL1: Less than 35 years old

AGECL2: 35-44 years old

AGECL3: 45-54 years old

AGECL4: 55-64 years old

AGECL5: 65-74 years old

AGECL6: 75 or more years old

Education of head includes four categories:

EDCL1: No high school diploma

EDCL2: High school diploma

EDCL3: Some college

EDCL4: College degree

Race or ethnicity of respondent includes two categories:

RACECL1: White non-Hispanic

RACECL2: Nonwhite or Hispanic

Current work status of head includes four categories:

OCCAT11: Working for someone else

OCCAT12: Self-employed

OCCAT13: Retired

OCCAT14: Other not working

Current occupation of head includes four categories:

OCCAT21: Managerial or professional

OCCAT22: Technical, sales, or services

OCCAT23: Other occupation

OCCAT24: Retired or other not working

Family structure includes five categories:

FAMSTRUCT1: Single with child(ren)

FAMSTRUCT2: Single, no child, age less than 55

FAMSTRUCT3: Single, no child, age 55 or more

FAMSTRUCT4: Couple with child(ren)

FAMSTRUCT5: Couple, no child

Housing status includes two categories:

HOUSECL1: Owner

HOUSECL2: Renter or other

Percentile of net worth includes five categories:

NWCAT1: Less than 25

NWCAT2: 25-49.9

NWCAT3: 50-74.9

NWCAT4: 75-89.9

NWCAT5: 90-100

These are the categories included in the *Bulletin* article. For a further description of these categories please refer to this article. Users may select as many of these categories as they wish to condition their output by.

Page 9: Conditioning by the quantiles of a selected variable

Users may condition their output by the deciles, quintiles, or quartiles of a few of the aggregate variables. Note that the cut points for the deciles, quintiles, or quartiles are calculated over all households, regardless of any conditions specified. If users select "Condition by quantiles (calculated over all households) of selected variables" on page 7 this page appears. To condition by the quantiles of a selected variable three steps must be taken. In step 1, the user selects quartiles, quintiles, or deciles. Once one of these quantiles is selected the box in step 2 is populated with a list of corresponding quantiles. In step 2, the specific quantiles are selected to condition by. Finally, in the third step the user selects the variables to condition by. All of the variables in the list may be selected. Something in each step must be selected in order to proceed. User defined variables do not appear in the list of variables.

Page 10: User-specific condition

This page allows users to condition their output by a user-specific condition. This page includes a list of all of the aggregate variables, two text boxes, and a check box. The user may select one aggregate variable from the list of variables to condition their output table by. The first box to the right is for a lower bound specification. The second box is for an upper bound specification. If no lower bound is specified, the range of the conditioning variable is unbounded from below; similarly, not entering an upper bound indicates that the range is unbounded from above. If the variable selected is a dollar variable then the check box ("Adjust bounds to base year dollars specified earlier") may be checked. Selecting this check box indicates that the user-specified upper and lower bounds correspond to amounts in real dollars of the base-year selected on page 4. Leaving the box unchecked indicates that the bounds are in specified in survey-year dollars. Do not select the check box if the variable used to specify the condition is not a dollar variable (e.g., conditioning on age, education, or income percentile group). To return descriptive statistics for families with income greater than or equal to $80,000, INCOME would be selected from the list of variables, 80000 would be entered into the lower bound text box, and the check box would be clicked. This conditions the output to consider only households that meet this restriction in the resulting table.

Page 11-14: A second set of conditions

Pages 11-14 perform exactly the same functions as pages seven thru ten, allowing users to place a second set of conditions on their output. Please refer to the text above for a discussion of these pages and their functionality.

A few examples on how to use the SCF Tabling Wizard are provided below.

Example 1:

To write a table with weighted medians in current dollars to a new worksheet “Table1" for “families with holdings” of the variables DEBT, CCBAL, and INSTALL, conditioned by *Bulletin* categories INCCAT1 through INCCAT6 and EDCL1 through EDCL4 perform these steps:

First page that appears: Type Table 1 in the text box. Click forward.
Second page that appears: Click forward. Weighted median is the default.
Third page that appears: Click forward.
Fourth page that appears: Select DEBT, CCBAL, and INSTALL from the list of variables. To select multiple variables hold down the ctrl key and click the variable names. Click forward.
Fifth page that appears: Select the button next to “Bulletin categories”. Click forward.
Sixth page that appears: Select INCCAT1-6 and EDCL1-4 from the list of Bulletin categories. Click finish.

Example 2:

To write a table of weighted means to an existing worksheet “deciles_of_asset” for all families for the variables INCOME and NETWORTH, conditioned by the deciles of ASSET perform the following steps:

First page that appears: Click the button for existing worksheet and then select “deciles_of_asset” from the list of existing worksheets. Click forward.
Second page that appears: Click the button next to weighted mean. Click forward.
Third page that appears: Click forward.
Fourth page that appears: Select INCOME and NETWORTH from the list of variables.
Click forward.
Fifth page that appears: Click the button next to all households. Click forward.

Sixth page that appears: Click the button next to “Condition by quantiles (calculated over all households) of selected variables”. Click forward.

Seventh page that appears: In step 1, click the button next to “Deciles”; in step 2, select all ten deciles; in step 3, select Asset from the list of variables. Click Finish.

Example 3:

To write a table showing the percent of households that have non-zero holdings of the variable RETQLIQ (retirement accounts) by AGECL to the worksheet “ret_age” perform the following steps:

First page that appears: Type ret_age in the text box. Click forward.

Second page that appears: Selected weighted percent nonzero. Click forward.

Third page that appears: Click forward.

Fourth page that appears: Select RETQLIQ from the list of variables. Click forward.

Fifth page that appears: Select the button next to “Bulletin categories”. Click forward.

Sixth page that appears: Select AGECL1-6 from the list of Bulletin categories. Click finish.

Example 4:

Here is an example of the steps necessary to output a table of weighted percent in group that calculates the percent of households with greater than or equal to one million dollars in Net Worth holdings that is stored in an existing worksheet “WGF_NW”.

First page that appears: Click the button next to “Existing worksheet”, select WGF_NW from the list of worksheet names, and click forward.

Second page that appears: Select the button for weighted percent in group. Click forward.

Third page that appears: Click forward.

Fourth page that appears: Select NETWORTH from the list of aggregate variables, enter 1000000 in the lower bound text box, and then click finish. Please do not enter one million with comma separators. The Visual Basic programs do not recognize comma separators.

To determine the percent of households with less then or equal to one million dollars in Net Worth holdings repeat the steps above but enter 1000000 in the upper bound text box instead. Likewise, to determine the percent of households with 1000 to 20000 dollars of new worth holdings follow the steps above except enter 1000 in the lower bound text box and 20000 in the upper bound text box. All values entered as an upper or lower bound are inclusive.

Example 5:

To calculate two tables of weighted medians for the variables BUS and CCBAL that are conditioned by AGECL 1-6 and the quantiles of the variable FIN (financial assets) in a new worksheet named BUS_CCBAL_2c perform the following steps:

First page that appears: Enter BUS_CCBAL_2c into the text box in step 2. Click forward.

Second page that appears: Select weighted median. Click forward.

Third page that appears: Click forward.

Fourth page that appears: Select BUS and CCBAL from the list of aggregate variables. Click forward.

Fifth page that appears: Click forward.

Sixth page that appears: Select the button next to *Bulletin* categories. Click forward.

Seventh page that appears: Select AGECL 1-6 from the list of Bulletin categories. Click forward.

Eighth page that appears: Click the button next to “Condition by quantiles (calculated over all households) of selected variables”. Click forward.

Ninth page that appears: In step 1, select the button next to “Quintiles”; in step 2 select all five quintiles in the list; in step 3, select FIN from the list of variables provided. Click finish.

Potential questions and a few warnings:

Is it ok to have other Excel workbooks open while working with the SCF tabling wizard?

While it is possible to have other workbooks open it is best to close all other Excel workbooks before opening the extract and using the SCF tabling wizard. There are several reasons to close all other workbooks. First, the extract is a very large data set for Excel to manage. The less you have open the faster the wizard will run. Further, the SCF tabling wizard has gone through extensive testing before its release. However, it is a new product that may still contains bugs. The more workbooks users have open the greater the chance these bugs will appear.

What should be done if multiple SCF Tabling menu items appear?

The SCF Tabling menu items that appear in the Tools menu, on your command bar, and the shortcut menu are placed there when this workbook's "open event" runs. These menu items are removed when the workbook's "close event" runs. If the workbook's "close event" does not run properly, for any reason, these menu items are not removed. The next time this workbook is opened two sets of menu items appear. When this happens users will want to remove the extra set of menu items. In order to do this go to the tools menu, select macro, and then macros. When the list of macros appears select the macro DelSCFTablingBtn from the list and click run. Repeat these steps for the macro DelSCFTablingFromShortCut and the macro DelSCFTablingMI. This process removes the extra set of menu items.

Why are some of the letters in the wizard underlined?

These are accelerator keys. Accelerator keys allow the user to use keystroke combinations to navigate through the wizard. While the wizard can not be used independently of a mouse these accelerator keys come very close to allowing users to use only keystrokes to output tables with the wizard. To perform the function of the object that has an accelerator key hit Alt and the underlined letter at the same time.

What should be done if the SCF Tabling menu items do not appear or seem to suddenly vanish?

The SCF Tabling menu items are removed when this workbook's "close event" runs. The close event runs when users click close from the file menu or click the "x" in the upper right corner of the workbook. The close event prompts users to save their work. If users select cancel from this prompt the SCF menu items have been removed while the extract remains open. If this happens, close the wizard and reopen it.

How can I determine if macros have been disabled and how can I enable macros?

When macros are disabled the extract data will appear but the SCF tabling wizard will not be available. In general, Excel's macros will be disabled because the security setting is set to 'High'. To enable macros in the SCF Tabling Wizard file, follow these steps: Open Excel, select 'Tools' from the Excel Menu, then 'Macro', and then 'Security'. In the dialog box, select 'Medium Security' and Press 'OK'. Close the file, and then open the SCF Tabling Wizard file. This time you should see the 'Enable Macros' dialog box.

Why does some of the output appear in bold?

Some cells in a table may appear in bold. This happens when calculations are based on fifty or fewer implicates (ten or fewer actual interviews or households). The bold font warns users that these calculations may be influenced by outliers.

What if user-defined variables do not appear in the variable lists in the SCF tabling wizard?

This could happen for a couple of reasons. First, make sure that these variables have a name and it appears in the first cell of the column that contains the variable. Also, make sure that these variables have the same number of rows as the other aggregate variables. Second, make sure these variables are contained in columns directly next to the aggregate data. In other words, be sure the data are contiguous (no empty columns separate the aggregate data from user-defined data). If both of these conditions are satisfied user-defined variables should appear in the SCF tabling wizard's lists. If they do not, contact the SCF team.

What if a Visual Basic run-time error occurs?

We hope this never happens but there is a chance it will. If this happens please write down the error number and the steps you took to produce the error then press end. After pressing end please send us an email at scf@frb.gov.

What if an error log appears?

The tabling wizard has been developed for the public so that more SCF users can utilize this complex data set. Consequently, we need your help when the wizard falls short as a useful tool. Whenever an error log appears please email or fax it to the SCF team. This will allow the SCF team to

fix the error that you found making it a more useful tool for you and other users like yourself in the future.

Please do not change either the workbook name or the name of the worksheet that stores the data. The macros that run the wizard are dependent on these names. The SCF tabling wizard does not work on Macintosh operating systems.