

Notes for Student Version of Soritec

Department of Economics

January 20, 2001

INSTRUCTIONS FOR USING SORITEC

This is a brief introduction to the use of the student version of the Soritec statistical/econometric program. The student version has most of the features of the full version but is limited to relatively small databases. On the other hand it is distributed for a nominal fee and is easy to use. Most of the Soritec commands described in these notes will work in the student version.

These instructions refer only to the interactive computing mode in which the user is in a conversational mode with the computer. Soritec is a Windows 95/98/NT program.

First let's begin with how to start Soritec and how to exit the program when you are finished.

1. To start Soritec, simply click on the Soritec icon on the Windows desktop.
2. To exit Soritec, you can click on the File menu at the top and then click on Exit or simply type STOP. Remember to hit the Enter key after each command that is typed in.

You are now ready to use the Soritec commands described below.

Printing your results

A simple way is to use the mouse to "paint" the screen containing the results you want to be printed out and then click on the print icon at the top of your screen. A dialog box appears and you just click on the OK button.

You probably will want to keep a hardcopy of your Soritec session in a file that can be printed out later. After you have started Soritec you can create an ASCII (text) file of your session by typing

```
ON LOG
```

The program will create a file labeled SORITEC.LOG. After your session you will want to hunt down this file and rename it as a txt file in order to avoid confusion. You can use the Find option on the Windows Start menu. Use Windows Explorer to go to where the SORITEC.LOG file is located. Click once on the file to highlight the file name. Right click on the mouse to access a menu that has the Rename option. Click on the Rename option and you will see SORITEC.LOG highlighted. For example to change the name to OHMY1.TXT, simply type this new file name over the old one

You can print out the file OHMY1.TXT by reading it in any word processing program as a text file and then using the print option. Microsoft Word would be a good choice as it preserves the formatting of the regression results.

Selected Soritec Commands and Procedures

A. Entering and Storing Data

Although Soritec has a spreadsheet type of data entry system, it would probably be best to enter your data in a spreadsheet program such as Excel.

The USE command

Suppose you are in Excel. Before entering data, you should inform Soritec of the type of data to be entered or used, i.e., cross-section, annual time-series, quarterly time series, etc. This is accomplished with the USE command. For example, suppose you want to use or enter cross-section data containing 48 observations. In the first cell of the spreadsheet you would type the command

```
USE 1 48
```

Or suppose that you wish to enter or analyze annual time-series data for the period 1970 to 1983. In the first cell simply type the command

```
USE 1970 1983
```

Assuming you have quarterly data to work with from the 3rd quarter of 1968 to the 2nd quarter of 1987, you would type the command

```
USE 1968Q3 1987Q2
```

Remember to type Q after the year, followed by the particular quarter (a number between 1 and 4). In the case of monthly data, the year would be followed by M and the particular month (a number between 1 and 12).

The Read statement

Let's suppose you have two variables named CONSUMPTION and INCOME. In the spreadsheet cell just below the USE command, type READ CONSUMPTION INCOME

Entering the data into the spread sheet

Suppose you have 6 years of annual data on the variables CONSUMPTION and INCOME. You would type in the data each year in the first and second columns just below the read statement. Your spreadsheet would look like this:

```
USE 1 6
```

```
READ CONSUMPTION INCOME
```

```
325                410
```

```
335                433
```

```
355                440
```

```
375                458
```

```
401                502
```

```
433                515
```

To save this as a text file on your diskette in Drive A: click on the File menu in Excel and select Save As. A dialog box appears. At the bottom of the box is a small window labelled Save as type. Scroll down the choices and click on Text (Tab delimited). Enter the name of the file you want to save in the small window above labelled File name. Let's suppose you want to name the file WHOA.TXT. You would just type in WHOA.TXT in the small window. Let's assume you want to save your data file on your diskette which you

have inserted in Drive A: of the computer. At the top of the dialog box is a small window labelled **Save in**. Scroll through the directory until you find something like $3\frac{1}{2}$ Floppy (A:). Just click on this choice. At the right of the dialog box is a button marked **Save**. Click on this. You will get some messages but all you need to do is click **OK**. Now you want to exit Excel by clicking on the **File** menu at the top. You will get another message but all you need to do at this point is click **NO** and you will be out of Excel.

Reading in an ASCII (text) data file

Suppose you wish to read in the data file WHOA.TXT in Soritec. After you are in Soritec click on the **File** menu and click on the **Open Data File** option. A dialog box appears. At the bottom of the box is a little window labelled **Files of type**. Scroll down and select **All Files (*.*)**. At the top of the box is a little window labelled **Look in**. Assuming your data file WHOA.TXT is located on your diskette in Drive A:, you would just scroll through and locate Drive A: and click twice on WHOA.TXT or click once and then click on the **Open** button. Presto, your data is read by Soritec and you are ready to roll. You can ignore the messages at the top. This would be a good time to type the ON LOG command if you have not already done so.

Saving your data in a data bank

If you have created some variables (see below) it is a good idea to save your newly created variables as well as the original ones for possible future use. For this purpose you need to create a databank only once. You can subsequently enter and save new variables of differing types (cross-section, quarterly time-series, etc.) in the original data bank.

To create a databank with the name OOHAAH (you can make up your own names), and save it on your diskette in Drive A: type

```
CREATE 'A:/OOHAAH'
```

Soritec is very picky about the ' single quote symbol. Use the same single quote ' symbol at the beginning and the end. Suppose you created variables LCON and LINC. You may now save these as well as your original variables in the OOHAAH databank by typing

```
KEEP CONSUMPTION INCOME LCON LINC
```

Accessing an already created databank

Let's suppose that you need to get back on the computer to run some sort of analysis on a variable that you previously entered and saved in a databank. Assume that the variable is named CONSUMPTION and that the databank is named OOHAAH. After getting into Soritec you would click on the **File** menu and select the **Open Data File** option. A dialog box appears. At the bottom of the box is a little window labelled **Files of type**. Scroll down and select **DataBanks (*.sdb)**. At the top of the dialog box is a little window labelled **Look in**. Assuming you named your databank file OOHAAH and saved it on your diskette in Drive A:, it will appear as OOHAAH.SDB. You would just scroll through and locate Drive A: and click on OOHAAH.SDB. Presto again, your data is read by Soritec.

Suppose that the data range you are interested in 1960 to 1965, just type

```
USE 1960 1965
```

The variable CONSUMPTION and any other variables defined for the same data range are now available for use with the appropriate Soritec commands. Sometimes after you have opened up a previously created databank, you may have forgotten exactly what variables are lurking in your data bank. You can obtain a directory of these variables by typing

CONTENTS

Of course there is no guarantee that you will remember what your variable names stand for! If at some future date you are sure that you will not be needing a particular variable anymore, you can permanently banish the variable by using the DISCARD command. Assume you want to get rid of the variable CONSUMPTION. Simply type

```
DISCARD CONSUMPTION
```

After you exit Soritec, all discarded variables are history.

B. Generating New Variables with Algebraic Transformations

You may create new variables from the data you have entered by using certain algebraic functions. Let's assume you are working with two variables whose values you have previously entered into your databank. To keep things brief we will assume these are named X and Y. You want to create a new variable called Z. Below are some of the most common examples of data transformations.

For Z defined to be X divided by Y, type $Z = X/Y$

For Z defined to be X multiplied by Y, type $Z = X*Y$

For Z defined to be X raised to the exponent 3.2, type $Z = X**3.2$

For Z defined to be the square root of Y, type $Z = \text{SQRT}(Y)$

For Z defined to be Y minus X, type $Z = Y - X$

For Z defined to be the quantity X plus 3 all divided by Y, type $Z = (X + 3)/Y$

For Z defined to be the natural logarithm of Y, type $Z = \text{LOG}(Y)$

For Z defined to be the antilog of X, type $Z = \text{EXP}(X)$

For Z defined to be the difference between X and its one period lagged value, type $Z = X - X(-1)$

For Z defined to be the first period lagged value of Y, type $Z = Y(-1)$

It is important to note that when lagged variables are used, the USE command must be altered to reflect the loss of data corresponding to the maximum lag length of any variable to be used. Thus if we had data on the variable CONSUMPTION over the period 1960 to 1969, and we wanted to use any statistical routine that relies upon CONSUMPTION(-1) we would alter the USE command to now refer to the period 1961 to 1969.

Suppose you wish to create a time trend variable named T that equals 1,2,3, etc. Simply type

```
TIME T
```

If you omit the T after typing the word TIME, your time trend variable will be named TIME.

C. Calculating Means, Variances, etc.

There are times when it is necessary to calculate basic statistics for a variable such as its mean, variance, standard deviation, etc. At the top of the screen you can click on **Stats** button. If you select the **SYNOPSIS** option the program will generate basic stats for all of the variables you have used. Or you can be selective if you are interested in the basic statistics for the variables **CONSUMPTION** and **INCOME**. Just type

```
SYNOPSIS CONSUMPTION INCOME
```

The statistics produced by the **SYNOPSIS** command for the variables **CONSUMPTION** and **INCOME** will be calculated for the observations corresponding to the current **USE** command.

If you wish to calculate the sample covariances between pairs of variables, it is easiest to use the **COVA** command. Let's say you are interested in estimating the covariances among the variables **CONSUMPTION**, **INCOME**, AND **GOVT**. You would just type

```
COVA CONSUMPTION INCOME GOVT
```

This command would result in the printing out of a 3x3 matrix. The off diagonal elements of the matrix are the estimated covariances among each pair of variables, and the diagonal elements are the estimated variances of each variable.

Or suppose you wish to estimated the simple correlations among the variables **CONSUMPTION**, **INCOME**, AND **GOVT**. You would simply use the **CORREL** command and type

```
CORREL CONSUMPTION INCOME GOVT
```

This command would result in the printing out of a 3x3 matrix. The off diagonal elements of the matrix are the estimated correlations among each pair of variables. Reassuringly, the diagonal elements are equal to 1.0 as they represent the correlation between a variable and itself.

D. Equation Estimation

The REGRESS command

Suppose you wish to use ordinary least squares to estimate an equation involving one dependent variable named CONSUMPTION and two independent variables named INVESTMENT and GOVT. Just type the command

```
REGRESS CONSUMPTION INVESTMENT GOVT
```

Note that the dependent variable must precede the list of independent variables. If for some reason you did not want a constant (intercept) term in your equation you would type

```
REGRESS (ORIGIN) CONSUMPTION INVESTMENT GOVT
```

If you wanted to run a regression of CONSUMPTION on INVESTMENT lagged one period and on GOVT, you would type

```
REGRESS CONSUMPTION INVESTMENT(-1) GOVT
```

Remember to shorten the data range by one period in the USE command before running the regression with a one period lagged variable.

The RECOVER command

There are occasions when one wishes to retrieve information generated from a regression. Suppose you wish to retrieve the predicted values of the dependent variable CONSUMPTION from the regression equation and to name this new variable CHAT. Immediately after running the REGRESS command type

```
RECOVER CHAT YFIT
```

You now have a new variable named CHAT. YFIT is the generic internal name Soritec gives to the regression predicted values of the dependent variable. You are merely assigning your own name CHAT to this variable. Suppose you wish to retrieve the residuals from your regression and to name these residuals EHAT. Immediately after running the REGRESS command type

```
RECOVER EHAT RESID
```

You now have a new variable named EHAT. RESID is the generic internal name Soritec gives to the regression residuals. You are merely assigning your own name EHAT to this variable. If you wanted to recover both the predicted values of the dependent variable and the residuals it would be easiest to type

```
RECOVER CHAT YFIT EHAT RESID
```

The TWOSLS command

Two stage least squares (2SLS) is a very popular econometric technique for estimating structural equations in a simultaneous equations framework. Before the 2SLS method can be used, you must specify the predetermined variables. Suppose in your model the predetermined variables are named X1,X2,X3, and X4. These variables would be listed in the following command:

```
EXOGENOUS X1 X2 X3 X4
```

Now you are ready to estimate your structural equations by 2SLS. Assume that one of the equations you want to estimate treats variable Y2 as a function of Y1, and X3. You would simply type

```
TWOSLS Y2 Y1 X3
```

In this example Y1 and Y2 are endogenous variables. Note that the number of variables specified in the EXOGENOUS command (4 in this example) must be greater than or equal to the number of right hand side variables in the equation you wish to estimate (2 in this example). If this condition is not met, you will receive an error message.

Or suppose you have entered all 48 observations previously, but you are interested in analyzing observations 11 through 48. In this case you would type the command

USE 11 48