EXERCISE: Estimating the Relationship Between Aggregate Income and Consumption

The purpose of this exercise is to show how one might model the relationship between two trended variables. The variables of the example are the quarterly data on aggregate consumption and disposable income in the U.K. for the period 1955Q1–1994Q4, which are recorded in the prices of 1990. There are altogether 160 observations.

Data Preparation

It is appropriate to take the logarithms of the data and then to remove the seasonal fluctuations. These operations can be performed by the *IDEOLOG* program. Either the Wiener–Kolmogorov time-domain method or the frequency-domain method can be used.

In either case, a linear function should be fitted to the logarithms of the data and the regression residuals should be saved in place of the trended data. The seasonal adjustment procedure is applied to the detrended residual sequence and, thereafter, the linear trend is added back to the resulting sequence to create the seasonally-adjusted version of the trended data.

Reading the Data into gretl

The *gretl* program can be used in estimating the relationship between the seasonally-adjusted logarithmic data on consumption and income. The two data sequences are the first and the second columns, respectively, of the `incANDcons` data file, from which an ASCII(.txt) file should be formed, which can be read into *gretl*.

To read the file, one must access the *File* menu of *gretl*. The appropriate sequence of commands, which are chosen from the *File* menu and it sub menus, is ⟨*File*⟩ ⟨*Open Data*⟩ ⟨*text/CSV*⟩. In the resulting *data delimiter* window, you may specify *auto detect*. On typing *OK*, you will be confronted by a large data window, which enables one to specify the location of the the files to be read.

At the bottom right-hand corner of the window is an option box in which you must specifiy that the file to be read has the ASCII(.txt) format. A small window titled *data structure wizard* will appear in which it must be recorded that the data are *time series.* Thereafter, one can record that the frequencies are *quarterly* and that the start date is *1955.1*

The label $v1$ is given automatically to the the consumption sequence, which is to be the dependent variable, and the label $v2$ is given to the income sequence, which is to be the independent variable

The Regression in Levels

To begin, the logarithmic consumption data should be regressed against the logarithmic income data and a constant. The regression is performed most readily by clicking on the $\hat{\beta}$ symbol at the base of main window of the program.

The slope coefficient of the regression should be noted as well as the coefficient of determination (R-Squared) which will strike you as remarkably close to unity. The effects of the regression can be further investigated via the *Graphs* menu.

The Regression in Differences

The regression in levels produces results that are commonly regarded as spurious. Therefore, we should consider a regression in which the variables are in differenced form. To take the first differences of the variables $v1$ and $v2$, one must access the *Add* menu of the data window. Thereafter, one can create some lagged versions of the d_v2 income sequence.

The lagged sequences are revealed by clicking on the plus sign that appears beside d_v2. The regression of d_v1, which is the differenced consumption data, on d_v2 and its lags will deliver will a remarkably low coefficient of determination, which will increase by very little as further lags of d_v2 are included amongst the explanatory variables.

The Cointegrating Regression

A cointegrating regression can be obtained by supplementing the explanatory variables of the regression in differences by a so-called error-correction term, which comprises a linear combination of the variables in levels. The appropriate linear combinations is the variables $v3 = v1 - \gamma v2$, where γ is the slope coefficient from the regression of v_1 on $v2$ and a constant. This variable has been included in the `incANDcons` data file, where it is ready for use.

The inclusion of the error-correction term will have the effect of inceasing the coefficient of determination, which will continue to be low.