Chapter 5
Parameter Estimation:

DataView's regression commands at their most basic are essentially intuitive. For example, consider:

IMP=F(GNP,CAPI)

which specifies that IMP is a function F() of the variables GNP and CAPI. This is the standard form: to execute a linear-in-parameters regression, simply create an expression that states that the dependent variable (left-hand-side) is a function of one or more independent (right-hand-side) variables. In the right context, provided that you have opened the demonstration data bank and set the date range and observation frequency, DataView will find the relevant series, perform the regression, and display the results shown in Figure 5-1.
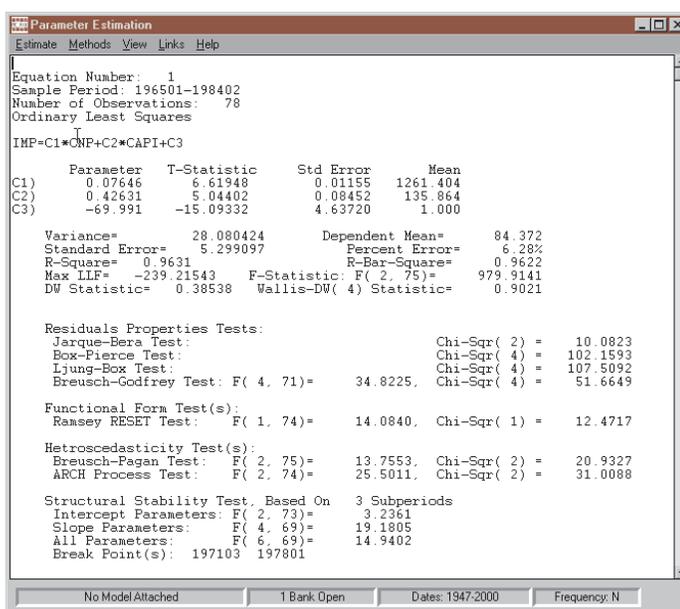

Figure 5-1. Regression Results Example

More generally, the extended   form of the Ordinary Least Squares regression command is:

$$Y=F(term_1,term_2,...,term_k;qualifiers)$$

where the terms themselves can be variables or expressions and the qualifiers–separated from the regressor terms by a semi-colon, as shown–further define the characteristics of the regression, as will be described.

However, before considering this command and its results in detail, it may be useful to identify the context in which it can be used. There are 2 ways to issue a regression command. The first is as an ordinary command in the context of the command space of the Central Control screen.   Alternatively, if you click on the Central Control screen menu category **Analysis**, and choose **Regression**, you will invoke the Parameter Estimation Screen and an initial display similar to that shown in Figure 5-2.
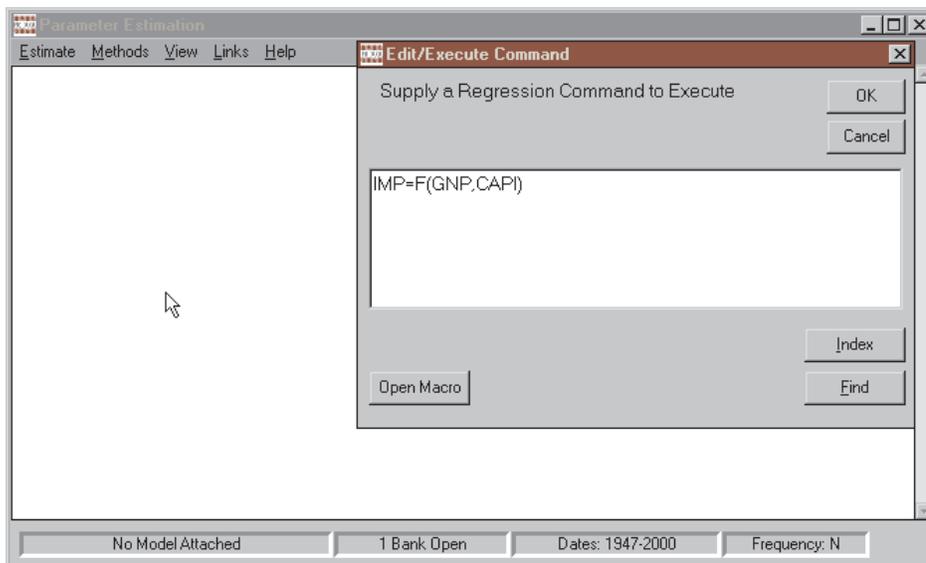

Figure 5-2. Parameter Estimation Screen Regression

Note, in particular, the **Index**, **Find**, and **Open Macro** buttons shown on the Edit/Execute form in Figure 5-2. Collectively, these provide information about available data series and allow you to capture mnemonics and possibly even entire regression commands from an existing macro file, for use in forming the current regression command.

Supplementary Regression Commands

There are also several supplementary commands and menu item equivalents that provide further information about a just-executed regression. For example, you can display the parameter estimator variance-covariance matrix by clicking on the menu category **View**, which offers the choice list:
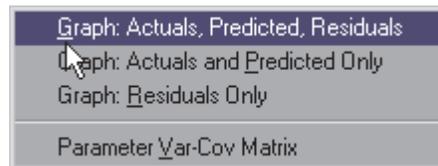
Parameter Variance-Covariance Matrix

If you choose the last of these, when there are two regressors and a constant term, a parameter variance-covariance matrix is displayed that takes the form:

```
PARAMETER VARIANCE-COVARIANCE MATRIX:


         1          2          3


1  .000133
2 -.000942     .007143
3 -.040256     .218121     21.503640
```

Predicted Values and Residuals

Alternatively, to display a comparison of actual and 'predicted' values of the dependent variable, a possible choice is the first **Graph: Actuals, Predicted, Residuals**, which produces the plot shown in Figure 5-3, displaying both the actual and predicted values of the dependent variable and the residuals.
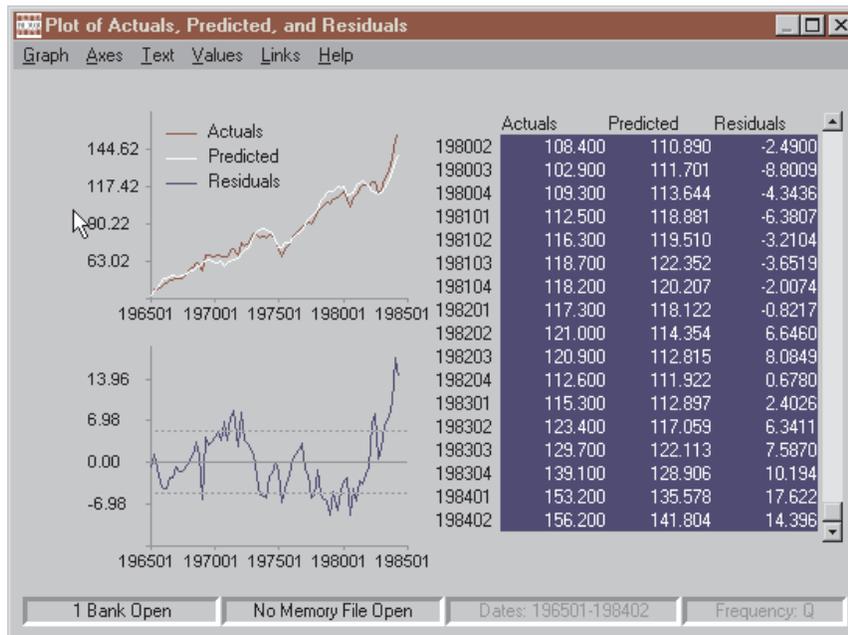
Figure 5-3. Sample Regression Plot

This regression plot has a number of the characteristics of an ordinary plot. For instance, you are able to edit it using the dropdown menu system shown in the figure and described in Chapter 2 or else capture it to the Clipboard, among other obvious options. In addition, you can generate a macro file that will replicate both the regression and graph; because Graph is a subsidiary regression option, this macro does not in fact directly reproduce the plot, but instead the regression, at which point the GRAPH option can again be chosen.

You can also create regression plots from scratch: to do this, first store the predicted and residual values in the Memory File either by clicking on the **Estimate** menu category of the Parameter Estimation Screen, which displays the choice list:
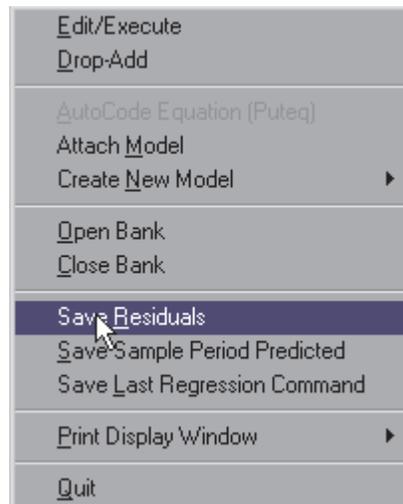
from which you can make the appropriate choices, or by using the commands:

SAVE PREDICTED AS seriesname
SAVE RESIDUALS AS seriesname

in a macro or from the Central Control screen. So long as no other operation is pending, you can return from the Parameter Estimation screen to the Central Control screen simply by pressing the <Esc> key, after which you can issue these commands.

The ability to save these results makes it possible to treat the predicted and residuals as you can any other time series, including allowing you to build your own, specially tailored, plots of predicted or residual values, instead of or in addition to those pro-

duced via the **View** menu choices. For example, the commands:

> **SAVE PREDICTED AS PRED**
> **PLOTTITLE:Graph of Actual and Predicted Regression Values**
> **PLOT IMP,PRED**

together will produce a plot of actuals and predicted.

  Alternatively, consider:

> **SAVE RESIDUALS AS RESIDLS**
> **HIGH=SDEV(RESIDLS)**
> **LOW=-HIGH**
> **PLOT RESIDLS,HIGH,LOW,0**

Together these create a plot that shows the regression residuals and the points at which the values are within or greater than one standard deviation—bear in mind that the sum of the residuals is zero, at least so long as a constant term appears in the regression.

  Incidentally, there is an important difference between these last two plots and those produced by clicking on the **View** menu category of the Parameter Estimation Screen: the plots last produced have all the properties of ordinary plots; that is, they are generated from data series retrieved from the Memory File or a data bank. They can therefore be produced from a set of explicit plot commands.   In contrast, the regression Graph option plot is directly generated from the regression results as a special display.

Other Important Regression-related options

  Look back at the choice list shown on page 123, which you will

recall is displayed by clicking on the **Estimate** menu category. In addition to the commands just described that save residuals and predicted values, this list includes several others that manage the parameter estimation process. Among these are selections that permit the editing of regression commands, banks to be opened or closed, and results to be printed.

Edit/Execute Command

The **Edit/Execute** choice in particular recalls the command for the current regression and permits it to be directly edited in a variety of ways. Once a regression has been run there are two ways to invoke this option.   The first is to click on **Edit/Execute**. Alternatively, while the regression results are displayed, you can point your cursor at these and left-click the mouse.   You should next see a dialogue box that displays the regression command:
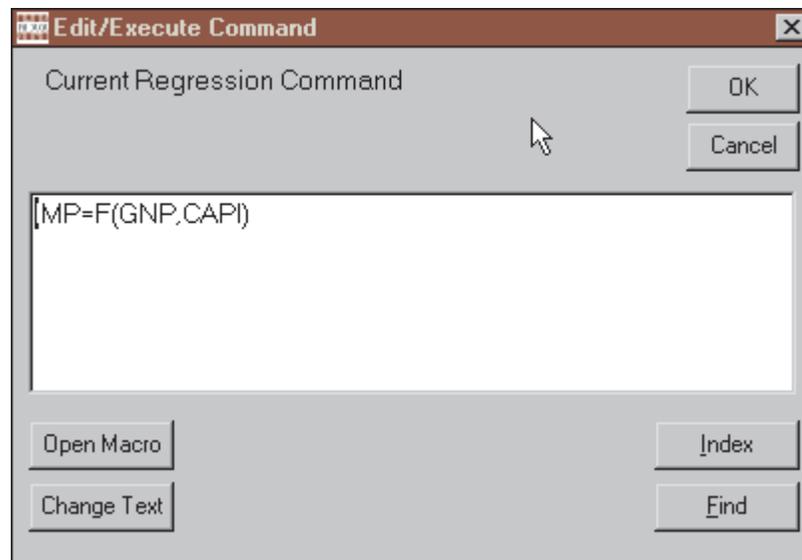

Figure 5-4. Edit-Execute Regression Command Dialog

Note that the white space that displays this command is meant to contain text; one of the consequences is that if you click on this space, using your *right* mouse button, and if you are using Windows 95 or a later operating system, a floating text edit menu will appear.   You will also find that if you click your mouse on the text of the command, the backspace, delete, and other editing keys work.   In short, you have the ability to edit the text in any way you see fit.   Note in addition the **Change Text** button on the lower left of the form, which can be used to make multiple text replacements.

As a bonus, the **Open Macro** button can be used to gain access to regression commands and other text that you may find relevant. And once you open a macro, if you highlight a segment of text there, it will be copied immediately to the Edit/Execute command space.

In turn, the **Index** and **Find** buttons can be used to discover variables in open data banks or the Memoryfile that you may wish to use in regression commands.

Drop-Add Command

  The Drop-Add facility is more specific in its function. Click on the **Drop-Add** entry of the **Estimate** menu category and you should see the dialog box shown on the next page.

This facility refers to regressor variables *only*; changes in the dependent variable must be made using the Edit/Execute command. Within this restriction, the terms to be dropped from the existing regression are identified by number, counting from the left-most regressor term.

Up to the number of regressor terms, you can include as many numbers as you wish in the text space following the word DROP.
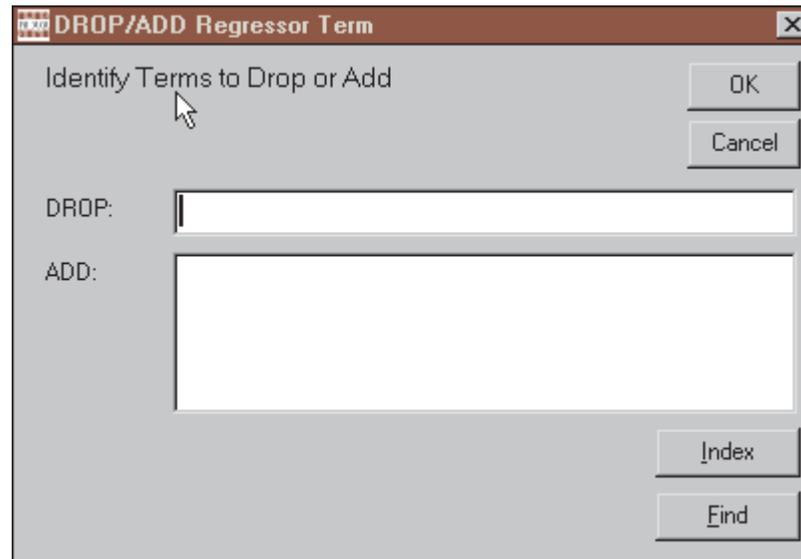
Figure 5-5. Drop-Add Contol

In the text space following the word ADD, you can include new regressor terms, separated by commas; these regressor terms can be variable mnemonics or expressions.   The **Index** and **Find** buttons in the bottom right quadrant of this form are provided in order to assist you in Adding additional terms.

Saving Estimated Equations: Single Equation Models

 DataView differs from an ordinary regression package in a number of ways, but a characteristic difference is that it allows you to save an estimated equation as a functioning single equation model.

 In order to create a model from an equation you have just estimated, simply click on **Create New Model: Single Equation**, which is found under the **Estimate** menu item. When you click, a form will appear, containing several text spaces to be filled in. Using this form, you will be asked to name the model, describe it

if you wish, and specify the frequency of the data to be used with it—which will usually be the same as the estimation frequency. In addition, if relevant, you may need to specify the normalization, the implications of which are further discussed below.

Once you have identified the model and its characteristics, the Parameter Estimation Screen menu will immediately change.   It will now include an additional category, **Projections**, as shown in Figure 5-6.
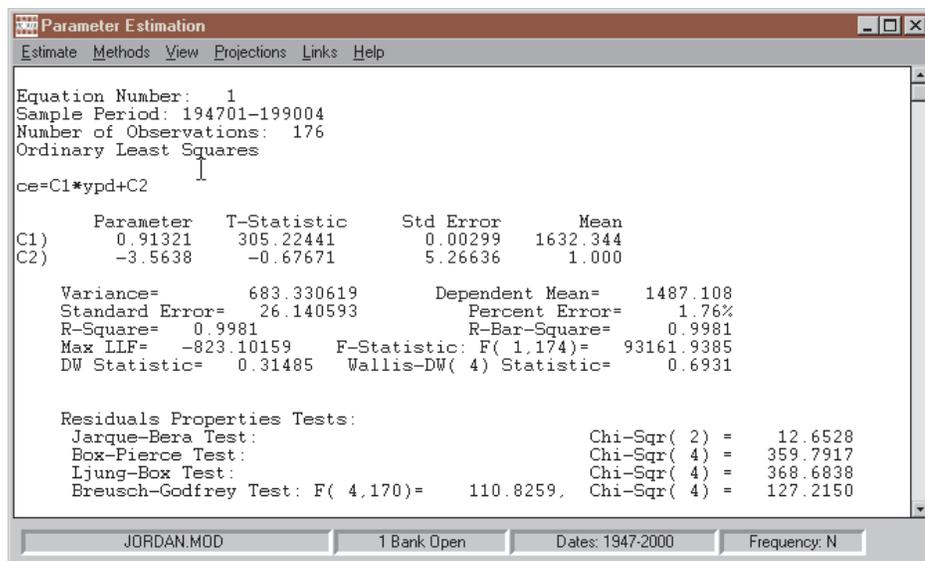
Figure 5-6. Estimation Screen with Projections

If you then click on **Projections**, you should see immediately a choice list in which only the first item will be enabled, **Project Equation**. The other items refer to the post-projection display of results, so are not relevant until an initial projection has been made.   Clicking on **Project Equation** will cause another form to be displayed, that shown in Figure 5-7.

Notice that this form displays two sets of dates, identified respectively as the projection period and the sample period. The regression sample period is provided essentially as a reminder, for
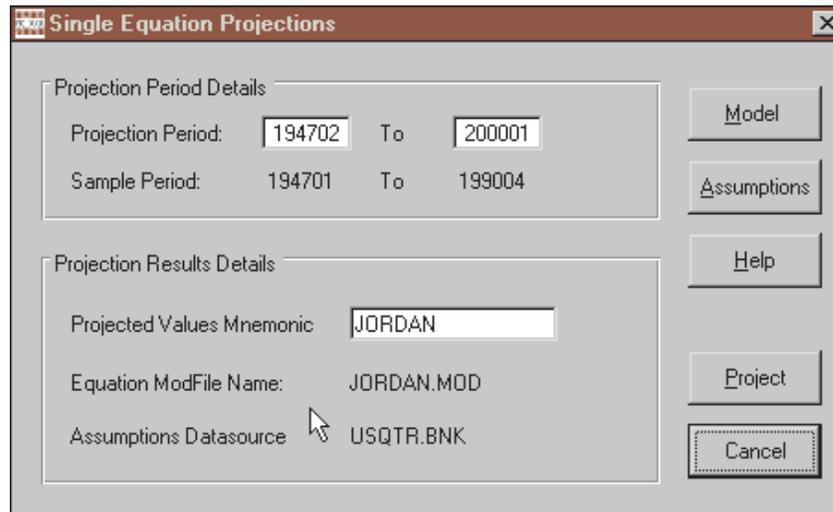
Figure 5-7. Projection Options

reference as you choose the projection period. The Projections Results Details in turn identify the file used for the model and the datasource from which the assumptions data are to be retrieved; unless you change it, the datasource will be that used for the regression. The Projected Values Mnemonic is the name under which the projections results will be stored in the Memoryfile; note that each projection will presumably result in a new set of projected values, so that in general you should change this mnemonic from one projection to the next—if you use the same name 2 or more times, previous values stored under the given name will be overwritten without warning.

 The **Project** button causes the projection to be made for the projection period specified. The **Model** button provides a description of the model's characteristics. **Assumptions** allows you to change the values of the regressor variables prior to a projection. Changes to regressor variable values are stored in the Memoryfile, not in the bank from which they are originally obtained; the effect of this convention is to preserve the original data. **Help** provides reasonably detailed context specific help, including how to dis-

play Single Equation model projections, which therefore will not be described here.

  An aspect of this process that you should keep firmly in mind is that any variables stored in the Memoryfile under the same name as that used in the original data bank automatically blocks retrieval of the original data bank values—because of data source precedence conventions, as described in Chapter 2. Thus, for example, if during a particular session you wish to create 2 or more single equation models in each case using the same orginal data values to perform the regressions, you must in between clear the Memoryfile, or at least delete from it the particular affected variables. Otherwise, if your projection period overlaps the regression sample period you will be in danger of using in the next regression assumed values from a previous projection. If you are making projections beyond the end of the regression sample period, there is in principle no problem, but you still should be aware that in the second case you will be using the Memoryfile as the immediate data source.

Sample Period Specification

  The regression sample period is specified most easily by clicking on the status bar entries for dates or frequency, as described in Chapter 2.   If you click on the status bar of the Central Control screen, the effect is to set, or reset, the dates and frequency globally for the session.   If you click on the status bar of the Parameter Estimation screen, the effect is local, generally affecting the regression sample period specifically.

  But in whatever way the sample period is set, as a standard convention DataView always attempts to perform a regression over the whole of the specified sample period. Thus if equations involve lags, DataView attempts to retrieve data for sufficient

periods prior to the initial sample period date to allow the regression to be performed as requested. The program truncates the sample period only when sufficient endpoint observations are not available.

 In some cases, within the range of available observations, DataView may discover particular missing observations. In this case, as a default, the program will issue an error message and refuse to execute the regression. At this point, if you wish, you can issue the command:

SET ALLOWNA=ON

which will have the effect of telling the program that missing observations should be ignored.   What occurs is that when a within-sample-period missing observation is found for a variable, DataView will delete the contemporaneous observations on all regression variables, thus in effect reducing the sample period by one observation selectively for each case. In addition, the regression results display produced subsequently will note specifically that missing non-available (NA) observations have been ignored.

There are circumstances in which ignoring missing observations makes sense. In particular, if there are no lagged variables in the regression, and certain other conditions prevail, it can be shown that *ceteris paribus* the effect is only to diminish the efficiency of the parameter estimates. However, in general, omitting NA observations can compromise the properties of a regression.   *It is your responsibility to make sure that you have taken into account all the possible effects*.

Trends , Dummy Variables, And Other Such Terms

 You can explicitly generate time trends, dummy variables, and

other such generated regressor terms if and when you want to include them. For example:

TREND=TIME(1,8401)

Alternatively, the TIME function can be directly embedded in a regression command, either using the ADD option or by typing it into the original command. In general, regression commands can contain both embedded expressions and functions. The most notable exceptions to this rule are the frequency conversion functions, which cannot be included in expressions or in regression commands. A list of DataView functions is shown earlier.

  However, if you are running regressions in order to build and use single equation models, then you should also take into account that functions defined on a particular time domain (such as TIME, MEAN, SDEV, etc) cannot be included in a model, even if they can be freely used in regression equations.   The use of these functions is not allowed precisely because they can be unintentionally misused in the context of a model; however, as in the case of the variable TREND above, it is always possible to solve the function and then include the variable (or a particular constant) in a model.

Correcting Mistakes

  It is easy to make mistakes when specifying functional relationships of any complexity. Even when you build a regression step-by-step using ADD and DROP commands, you are almost certain at some stage to make a typing mistake or to get a series name wrong. When this happens DataView displays an error

message and the offending command, generally offering an opportunity to make the correction.

As a design characteristic, any errors you make while using DataView should be immediately correctable. Occasionally, however, errors can occur that have no obvious source, in which case you may see the message:

Possible Program Malfunction, Please Document and Report

This message indicates that the program is "aware" of some sort of problem, but cannot specifically determine its source or, possibly, the appropriate corrective action. Please document and report such problems. Such problems are often context dependent and are therefore difficult to discover independently—or to correct. The more details that you can provide as to how to re-create the error, for the purpose of determining its cause, the better it is—that is, the more likely it is that the error can be corrected.

Functional Forms

The terms in a regresion can be names of series or, more generally, functions and expressions. As an example, you can directly estimate a logarithimic relationship by typing

LN(IMP)=f(LN(GNP),LN(CAPI),TREND)

yielding the result shown in Figure 5-8.

```
Parameter Estimation                                          _ □ ×
Estimate  Methods  View  Links  Help

Equation Number:    1
Sample Period: 196601-198402
Number of Observations:    74
Ordinary Least Squares

LN(IMP)=C1*LN(GNP)+C2*LN(CAPI)+C3*TREND+C4

          Parameter   T-Statistic    Std Error      Mean
C1)        0.86218      2.52878       0.34095       7.143
C2)        0.38851      2.95961       0.13127       4.909
C3)        0.00455      2.45595       0.00185      41.500
C4)       -3.8368      -1.92092       1.99738       1.000

   Variance=       0.003130    Dependent Mean=        4.418
   Standard Error=    0.055949     Percent Error=       1.27%
   R-Square=   0.9670             R-Bar-Square=       0.9656
   Max LLF=    110.42042    F-Statistic: F( 3, 70)=     683.8300
   DW Statistic=   0.54931   Wallis-DW( 4) Statistic=     1.0973


   Residuals Properties Tests:
    Jarque-Bera Test:                          Chi-Sqr( 2) =     0.6257
    Box-Pierce Test:                           Chi-Sqr( 4) =    59.8082
    Ljung-Box Test:                            Chi-Sqr( 4) =    63.1874
    Breusch-Godfrey Test: F( 4, 66)=    16.2954,  Chi-Sqr( 4) =    36.7691

   Functional Form Test(s):
    Ramsey RESET Test:     F( 1, 69)=     0.6382,  Chi-Sqr( 1) =     0.6782

   Hetroscedasticity Test(s):
    Breusch-Pagan Test:   F( 3, 70)=     4.1273,  Chi-Sqr( 3) =    11.1220
    ARCH Process Test:    F( 2, 70)=     5.3664,  Chi-Sqr( 2) =     9.5718

   Structural Stability Test, Based On    3 Subperiods
    Intercept Parameters: F( 2, 68)=    15.4587
    Slope Parameters:     F( 6, 62)=    10.1579
    All Parameters:       F( 8, 62)=    14.6060
    Break Point(s):  197203  197803

       No Model Attached        1 Bank Open      Dates: 1947-2000    Frequency: N
```

Figure 5-8. Logarithmic Regression

Note that the left-hand-side of a regression may also contain expressions. This flexibility makes it possible to estimate relationships with imposed coefficients, for example:

LN(IMP)-0.5*LN(GNP)=F(LN(CAPI),TREND)

Inclusion of a trend in a regression sometimes helps to identify short-term inter-relationships between variables. Another technique for doing this is to express variables as deviations from moving-average values. The following example uses 16-quarter moving averages:

IMP-MAVG(16,IMP)=F(GNP-MAVG(16,GNP),
CAPI-MAVG(16,CAPI))

Incidentally, observe that this command extends to a second line. If a regression command requires more than one line, you should end each incomplete line with a comma to indicate that further

terms will be supplied on the next line. But if you use the comma as a line continuation character, make sure that you break the command after each complete term. Extraneous comma *must not be used*: break the expression at existing commas.

Dynamic Specifications

  Lag terms may be specified freely in order to estimate dynamic relationships between variables; you may include as many as 20 terms altogether.

   **IMP=F(GNP,GNP(-1),GNP(-2),CAPI,CAPI(-1),CAPI(-2),
   IMP(-1),IMP(-2))**

But note that it is sometimes better to write a dynamic formulation with difference terms. For example the relationship above is nearly equivalent to:

   dif(IMP)=f(dif(GNP),lag(1,dif(CAPI))

  This ability to embed expressions directly into regression commmands is one of the reasons that DataView provides considerable flexibility in the way in which regression commands may be specified.

Optional Tests: Serial Correlation of the Error Term

  One of the assumptions of the general linear model is that equation errors are independently distributed, when in fact it is often true that they are serially correlated. There are a number of specification tests automatically produced by DataView that bear on the properties of the error term, as you can see from Figure 5-5.

But a specific test for autocorrelation is the Durbin-Watson. However, if the regression includes a lagged dependent variable you can also generate the more appropriate Durbin h-statistic, by adding **TEST:DBH** to the regression command:

IMP=F(GNP,CAPI,IMP(-1);TEST:DBH)

However, note that, when computing this statistic, the lagged dependent **must** be the last regressor term, as shown here. The Durbin h-statistic is just one of a number of optional tests available; others are described in the online help that can be referenced by pressing **Help**.

  A characteristic of tests for autocorrelation, in particular, is that programs such as DataView also incorporate specific corrections, although you should be aware that it can be argued that serial correlation can be indicative of more general misspecification errors. You should therefore consider carefully whether you should apply a particular automatic correction or try an alternative specification. In any particular case, DataView is designed not to force you to take one action rather than another, but instead to support whichever choice you make.
  .
   To estimate an equation and include autoregressive correction, requires the addition of **AUTO** to the end of the regression command. For example:

IMP=F(GNP,CAPI;AUTO,ORD=2)

not only specifies autoregressive corrections (Hildreth-Lu, by default), but also defines (ORD=2) that second order autoregressive corrections should be performed; the results are shown at the top of the next page. Incidentally, since the grid search associated with Hildreth-Lu may result in a thousand or more regressions
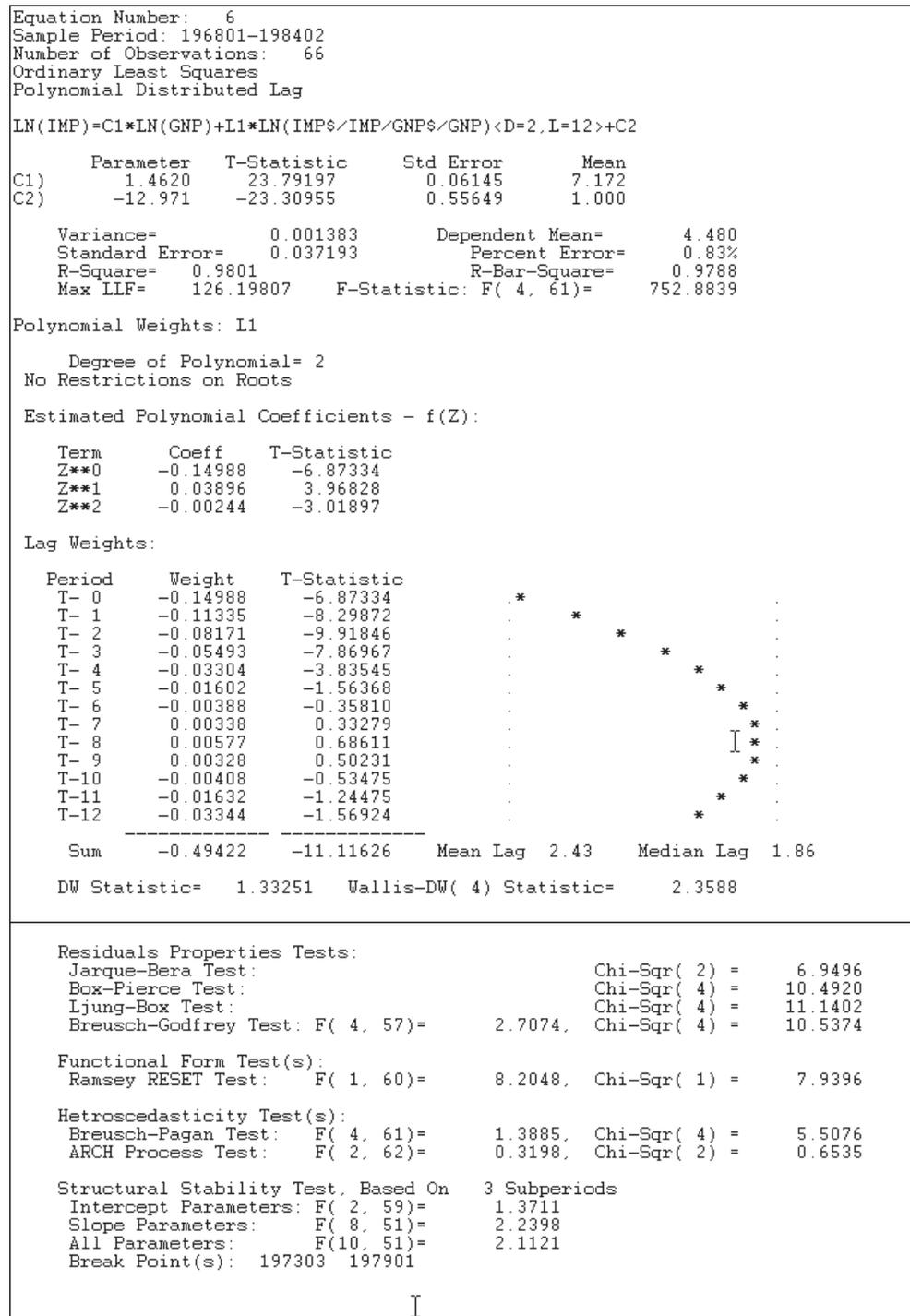
```
Equation Number:    6
Sample Period: 196801-198402
Number of Observations:    66
Ordinary Least Squares
Polynomial Distributed Lag

LN(IMP)=C1*LN(GNP)+L1*LN(IMP$/IMP/GNP$/GNP)<D=2,L=12>+C2

        Parameter    T-Statistic     Std Error       Mean
C1)        1.4620      23.79197       0.06145        7.172
C2)      -12.971      -23.30955       0.55649        1.000

    Variance=          0.001383      Dependent Mean=       4.480
    Standard Error=    0.037193         Percent Error=      0.83%
    R-Square=   0.9801                  R-Bar-Square=       0.9788
    Max LLF=    126.19807    F-Statistic: F( 4, 61)=    752.8839

Polynomial Weights: L1

     Degree of Polynomial= 2
 No Restrictions on Roots

 Estimated Polynomial Coefficients - f(Z):

    Term     Coeff    T-Statistic
    Z**0    -0.14988   -6.87334
    Z**1     0.03896    3.96828
    Z**2    -0.00244   -3.01897

 Lag Weights:

  Period    Weight    T-Statistic
   T- 0    -0.14988    -6.87334         .*
   T- 1    -0.11335    -8.29872         .       *
   T- 2    -0.08171    -9.91846         .           *
   T- 3    -0.05493    -7.86967         .              *
   T- 4    -0.03304    -3.83545         .                *
   T- 5    -0.01602    -1.56368         .                  *
   T- 6    -0.00388    -0.35810         .                    *
   T- 7     0.00338     0.33279         .                    ] *
   T- 8     0.00577     0.68611         .                    ] *
   T- 9     0.00328     0.50231         .                     *
   T-10    -0.00408    -0.53475         .                    *
   T-11    -0.01632    -1.24475         .                   *
   T-12    -0.03344    -1.56924         .               *
          ------------- -------------
    Sum    -0.49422    -11.11626    Mean Lag  2.43    Median Lag  1.86

    DW Statistic=   1.33251    Wallis-DW( 4) Statistic=    2.3588
```

```
    Residuals Properties Tests:
     Jarque-Bera Test:                            Chi-Sqr( 2) =    6.9496
     Box-Pierce Test:                             Chi-Sqr( 4) =   10.4920
     Ljung-Box Test:                              Chi-Sqr( 4) =   11.1402
     Breusch-Godfrey Test: F( 4, 57)=    2.7074,  Chi-Sqr( 4) =   10.5374

    Functional Form Test(s):
     Ramsey RESET Test:    F( 1, 60)=    8.2048,  Chi-Sqr( 1) =    7.9396

    Hetroscedasticity Test(s):
     Breusch-Pagan Test:   F( 4, 61)=    1.3885,  Chi-Sqr( 4) =    5.5076
     ARCH Process Test:    F( 2, 62)=    0.3198,  Chi-Sqr( 2) =    0.6535

    Structural Stability Test, Based On   3 Subperiods
     Intercept Parameters: F( 2, 59)=    1.3711
     Slope Parameters:     F( 8, 51)=    2.2398
     All Parameters:       F(10, 51)=    2.1121
     Break Point(s):  197303  197901
```

Figure 5-9. Distributed Lags Example

being performed, second order autoregressive equations may take a moment, although in the past few years such "moments" have shrunk: you are likely to notice a pause only if you are using an older machine with a 486 or earlier CPU.

  Alternatively, if you specify:

IMP=F(GNP,CAPI;AUTO:CO,ORD=2)

the regression will be performed using the Cochrane-Orcutt method. Note the addition of the letters CO following **AUTO**. The results are shown in Figure 5-10.


```
Parameter Estimation                                                    _ □ ×
 Estimate  Methods  View  Links  Help

Equation Number:    5
Sample Period: 196501-198402
Number of Observations:    78
Ordinary Least Squares
Autoregressive Corrections: Cochrane-Orcutt Technique

IMP=C1*GNP+C2*CAPI+C3

         Parameter    T-Statistic     Std Error       Mean
C1)       0.10146      3.69519         0.02746      1261.404
C2)       0.32930      2.52097         0.13062       135.864
C3)      -87.604      -3.74549        23.38927         1.000

         Rho(  1)= 0.70514      Accurate   3   Digits
         Rho(  2)= 0.18995      Accurate   3   Digits

   Variance=         10.226935      Dependent Mean=       84.372
   Standard Error=    3.197958       Percent Error=        3.79%
   R-Square=    0.9863              R-Bar-Square=        0.9852
   Max LLF=   -198.43169    F-Statistic: F( 4, 71)=    1280.1843
   DW Statistic=   1.99850    Wallis-DW( 4) Statistic=    2.0472
                        I
   Residuals Properties Tests:
    Jarque-Bera Test:                     Chi-Sqr( 2) =     1.7010
    Box-Pierce Test:                      Chi-Sqr( 4) =     0.8351
    Ljung-Box Test:                       Chi-Sqr( 4) =     0.9005

        No Model Attached        1 Bank Open    Dates: 196501-198402    Frequency: Q
```
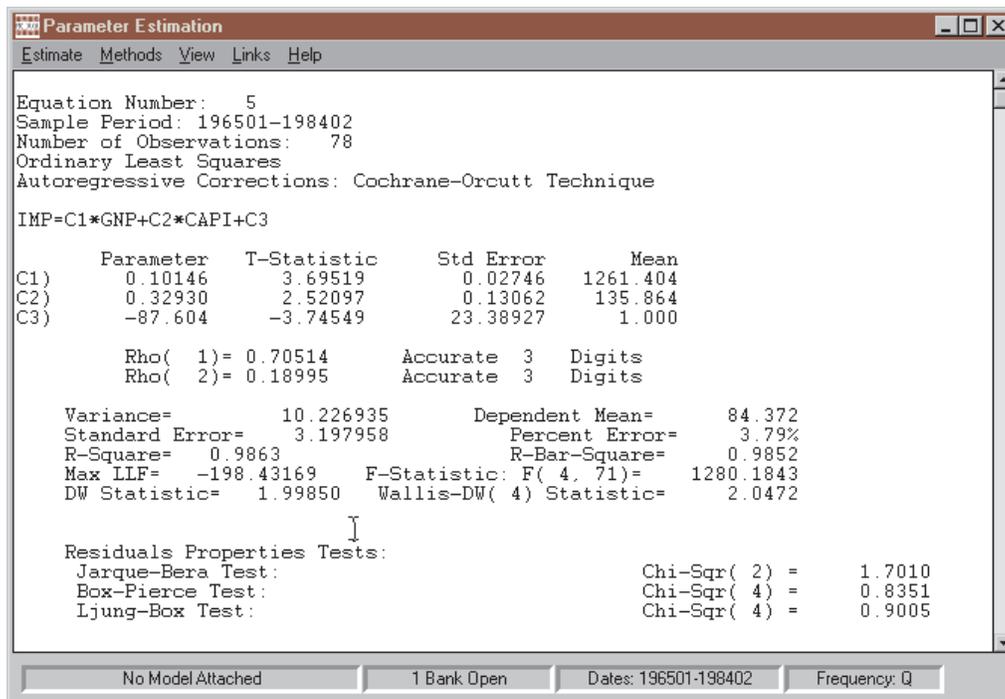
Figure 5-10. Autoregressive Corrections: CO, Order=2

Warning: If you use the AUTO option with others that follow the ; in a regression command, AUTO and its subsidiary commands must be last—otherwise the others will be ignored. For example to combine PRN, and AUTO:CO, type: ;PRN,AUTO:CO **not** the reverse ;AUTO:CO,PRN.

Distributed Lags

 It is often plausible to suppose that variables influence one another with quite long delays. Particularly in the case of quarterly or higher-frequency series this may imply that the effects of changes in an independent variable are distributed over a number of time periods. DataView provides techniques for estimating patterns of distributed lag coefficients represented by polynomial curves.

  As an example, consider the possible influence of import prices on the volume of imports. When the prices of imports are high relative to those of domestic products, the volume of imports is likely to be depressed as purchasers switch to domestic sources of supply. However, the influence of changes in relative import prices is not immediate, in part because of the existence of contracts; if they make themselves felt, these changes may cumulate over several quarters or even over years.

  To calculate an index of the relative price of imports using the data in DEMOBANK, specify:

>**PDIMP=IMP$/IMP**
>**PDGNP=GNP$/GNP**
>**RPIMP=PDIMP/PDGNP**

An unlagged regression for imports as a function of GNP and relative prices can then be specified as:

LN(IMP)=F(LN(GNP),LN(RPIMP))

In contrast, to define a distributed lag, it is necessary to state the total number of lag terms, L, and the degree of the polynomial approximation, D. Suppose that we expect relative import prices to continue having some effect for up to three years (L=12) and want to allow at least some curvature in the pattern of coefficients (D=2). An appropriate regression command might be:

LN(IMP)=F(LN(GNP),LN(RPIMP) <D=2,L=12>)

which yields the results shown in Figure 5-11.
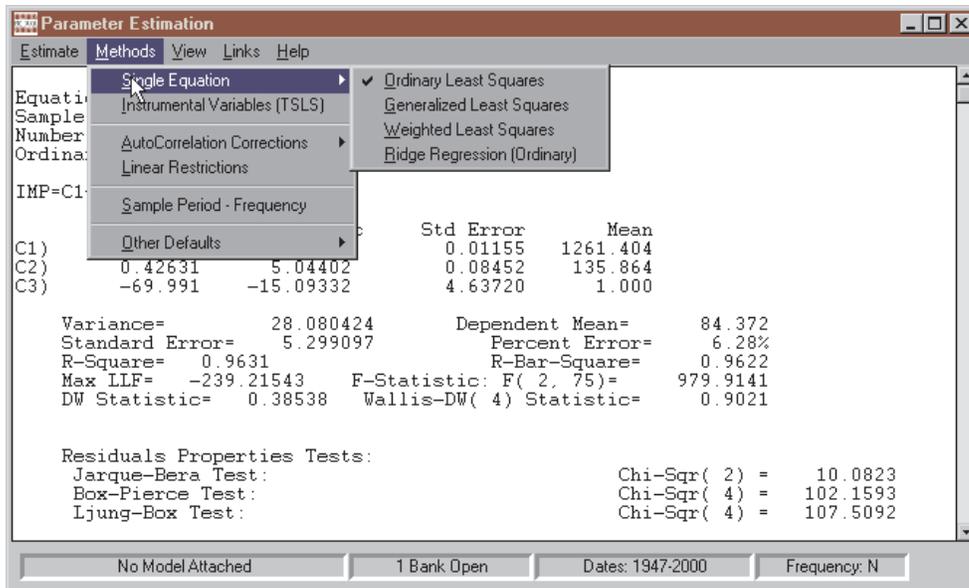
Figure 5-11. Regression Methods and Options

However, DataView also allows you to insist that the coefficients start from zero, NEAR, or end at zero, FAR, or BOTH. It may be reasonable instead to specify "FAR" restrictions:

LN(IMP)=F(LN(GNP),LN(RPIMP)<D=2,L=12,FAR>)

Additional Regression Facilities

 The focus so far in this chapter has been possible variations on the basic regression command. Almost without exception, the techniques considered can be interpreted as simply involving variable transformations and under this interpretation the discussion arguably has been limited essentially to Ordinary Least Squares.

 Five additional, frequently-used, DataView regression options are:

 (i) Restricted Least Squares

 (ii) Weighted Least Squares

 (iii) Ordinary Ridge Regression

 (iv) Logistic Curve

The commands to execute each of these still involve only slight modifications of the basic regression command.

 The stress, as before, will be the commands themselves. The reason is that equations are the stuff of econometric models and therefore their specification must be understood. However, Data-View also provides additional support via menu options. Observe

that clicking on the Methods menu category of the Parameter Estimation screen displays a choice of techniques, as partially shown in Figure 5-8.   Furthermore, given the basic equation specification, making choices among these techniques has the immediate effect of automatically modifying DataView regression command(s). Thus it is not necessary in all cases to modify these commands manually when switching between techniques.

Restricted Least Squares

In order to perform Restricted Least Squares, you must begin by specifying the restrictions. The command is:

RESTRICTIONS:R=r11,r12,...,r1k;r21,..,r2k;r n1 ,..,rnk;C=c1,..,cn

where r_{ij} denotes values in the restrictions matrix, R, and the c_i are elements of the C vector, using the Theil-Johnston notation. Note that semi-colons separate both rows of the R matrix and the elements of the R matrix from those of the C vector. Up to 15 restriction rows may be specified per regression—these must be independent linear restrictions on the parameters.

Once the RESTRICTIONS command is issued, the characteristics of the restrictions have been specified, but to actually perform a Restricted Least Squares regression, you must also insert ;RB=C into the regression command; for example:

$$Y=F(X_1,X_2,..,X_{K-1};RB=C)$$

Regressions performed without this insert will ignore the restrictions.

Weighted Least Squares

   The command to perform Weighted Least Squares also involves the insertion of a term: WLS:Z, where Z is the name of the weight series.   For example:

$$Y=F(X_1,X_2,..,X_{K-1};WLS:Z)$$

   The observations on Z will be divided into the observations on all the regression terms, making this command in the simple case equivalent to:

$$Y/Z=F(X_1/Z,X_2/Z,..,X_{K-1}/Z,1/Z;NOCON)$$

   In order to perform Weighted Least Squares where each term in the regression is weighted by a separate variable, you simply need to divide each term by the appropriate weight variable. But bear in mind that in the simple case, if $Z_i,i=1,2,..,T$ denotes the time series elements of Z, the use of Weighted Least Squares implies the hypothesis that the covariance matrix of the estimator, V, takes the form:

$$V = \text{diag } (s_1{}^2,...,s_T{}^2)$$

In the general case, the elements of V are not so easily specified.

Ordinary Ridge Regression

The Ridge regression command involves the specification of the Ridge Trace constant, c:

$$Y=F(X_1,X_2,..,X_{K-1};RIDGE=c)$$

where c is a real value scalar number.

Logistic Curve Model

The Logistic Curve Model can be defined in terms of DataView's LGT function:

$$LGT(Y)=F(X_1,X_2,..,X_{K-1})$$

where:

$$LGT(Y) = \ln\left(\frac{Y}{1-Y}\right)$$

defines the LGT function; Y **must** take a value between 0 and 1. The Logistic Model can be used in conjunction with Weighted Least Squares, in order to provide a GLS estimator. The argument term of the LGT function can be a single variable or an expression. Estimated equations can be normalized in terms of the variables that appear in the argument provided that the expression in the argument is an analytic function, and thus that an inverse function exists.