SAS® Econometrics and Time Series Analysis 1.1 for JMP®

SAS® Documentation

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2011. *SAS® Econometrics and Time Series Analysis 1.1 for JMP®* . Cary, NC: SAS Institute Inc.

SAS® Econometrics and Time Series Analysis 1.1 for JMP®

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st electronic book, August 2011

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at **support.sas.com/publishing** or call 1-800-727-3228.

Contents

Credits

Documentation

Writing	Jun Ge, Ruth Baldasaro
Editing	Anne Baxter
Documentation Support	Tim Arnold, Sharad Prabhu

Software

SAS Econometrics and Time Series Analysis for JMP

Interface Design	Mark Little, Jan Chvosta, Sharad Prabhu
Interface Coding	Jun Ge, David DeNardis, Joseph Morgan, Sharad Prabhu

SAS/ETS Procedures

PROC AUTOREG	Xilong Chen, Jan Chvosta, Richard Potter, Jason Qiao, John P. Sall
PROC PANEL	Jan Chvosta, Greg Sterijevski
PROC UCM	Rajesh Selukar

Support Groups

Software Testing	Jeanne Martin
Technical Support	Duane Hayes

Chapter 1

About This Book

This book describes the features of SAS Econometrics and Time Series Analysis for JMP and includes instructions for performing analyses. This book does not contain a comprehensive treatment of econometric and time series analyses or more technical details regarding the SAS/ETS® procedures. For more information about econometric and time series analyses, see Greene (1999). For more information about the SAS/ETS procedures, see the *SAS/ETS User's Guide*.

This book is organized as follows:

- This chapter describes the organization of the book.
- Chapter 2 provides a brief description of SAS Econometrics and Time Series Analysis for JMP.
- Chapter 3 gets you started with a simple example of how to use SAS Econometrics and Time Series Analysis for JMP.
- Chapter 4 provides an overview of how to perform a **Fit Y by X with Autocorrelation and Heteroscedastic Errors** analysis, a detailed description of the available features, and a comparison to the JMP Time Series platform.
- Chapter 5 provides an overview of how to perform a **Fit Panel Data Regression** analysis and a detailed description of the available features.
- Chapter 6 provides an overview of how to perform a **Fit Unobserved Components Models** analysis and a detailed description of the available features.

References

Greene, W. H. (1999), *Econometric Analysis*, Fourth Edition, Macmillan.

Chapter 2

Introduction to SAS Econometrics and Time Series Analysis for JMP

Contents

Overview

SAS Econometrics and Time Series Analysis for JMP is a graphical user interface that provides easy access to SAS/ETS procedures. You can specify a model in JMP, and SAS Econometrics and Time Series Analysis for JMP uses the relevant SAS/ETS procedures for model estimation. The results from the SAS/ETS procedures are returned to JMP and displayed in a JMP window. SAS Econometrics and Time Series Analysis for JMP provides three analyses:

- Fitting Y by X with autocorrelation and heteroscedastic errors enables you to fit models to time series data when the errors are autocorrelated or heteroscedastic.
- Fitting panel data regression enables you to perform panel analysis when you have both time series and cross-sectional data.
- Fitting unobserved components models enables you to fit unobserved components models to equally spaced univariate time series data.

Each analysis has a similar structure with four types of windows: initial model, results, dialog, and settings windows. This chapter contains an overview of each of these window types. The subsequent chapters provide detailed descriptions of these windows and instructions for how to specify each analysis.

You can access SAS Econometrics and Time Series Analysis for JMP by selecting **Analyze ► Econometrics and Time Series** from the JMP Home window. The **Econometrics and Time Series** menu contains three analysis commands and the help documentation. The analysis commands are described in detail in Chapter 4, Chapter 5, and Chapter 6.

Specify an Analysis

Creating an analysis consists of the following basic steps (step 1 through 3 for each analysis type are described in detail in Chapter 4, Chapter 5, and Chapter 6):

1. In the initial model window, specify and estimate a base model.
2. In the results window, examine the base model results.
3. In the dialog window, create and estimate an alternative model.
4. In the results window, examine the results of the alternative model and compare the model to the base model.
5. Continue creating alternative models and comparing them with previous models until you are satisfied with the model fit.

Initial Model Window

The initial model window appears when you select an analysis from the **Econometrics and Time Series** menu. In this window, you specify the roles for the variables you want in the base model, which represents default parameter selections. You can change the base model later.

The name of the initial model window and the options it contains depend on which analysis you select. See Chapter 4, Chapter 5, and Chapter 6 for the name of and the options in the initial model window for each type of analysis.

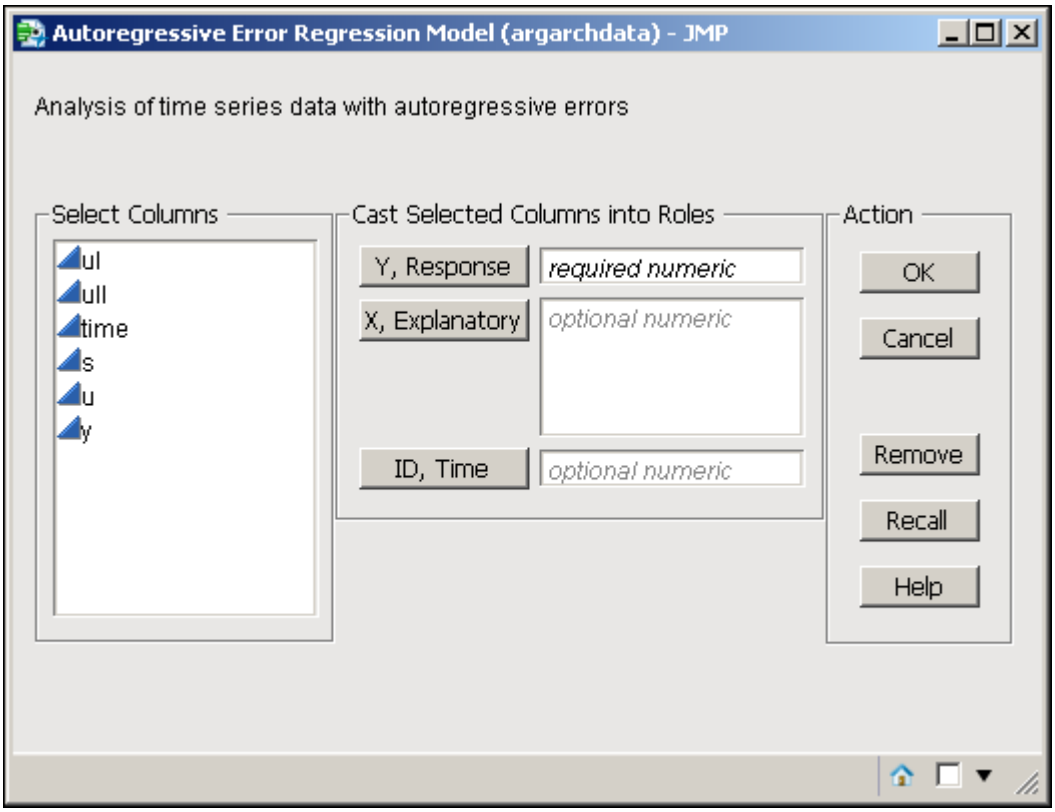
All of the initial model windows also have the following buttons:

- Click **Cancel** to close the window without saving the roles.
- Click **Remove** to remove highlighted roles.
- Click **Recall** to restore the saved roles that were selected in the most recently used model.

- Click **Help** to open the help manual.

Figure 2.1 shows an example of the initial model window for **Fit Y by X with Autocorrelation and Heteroscedastic Errors**.

Figure 2.1 Example Initial Model Window



Results Window

The results window contains plots and tables of results for the models you fit. The name of the results window depends on which analysis option is selected. Each results window contains four areas:

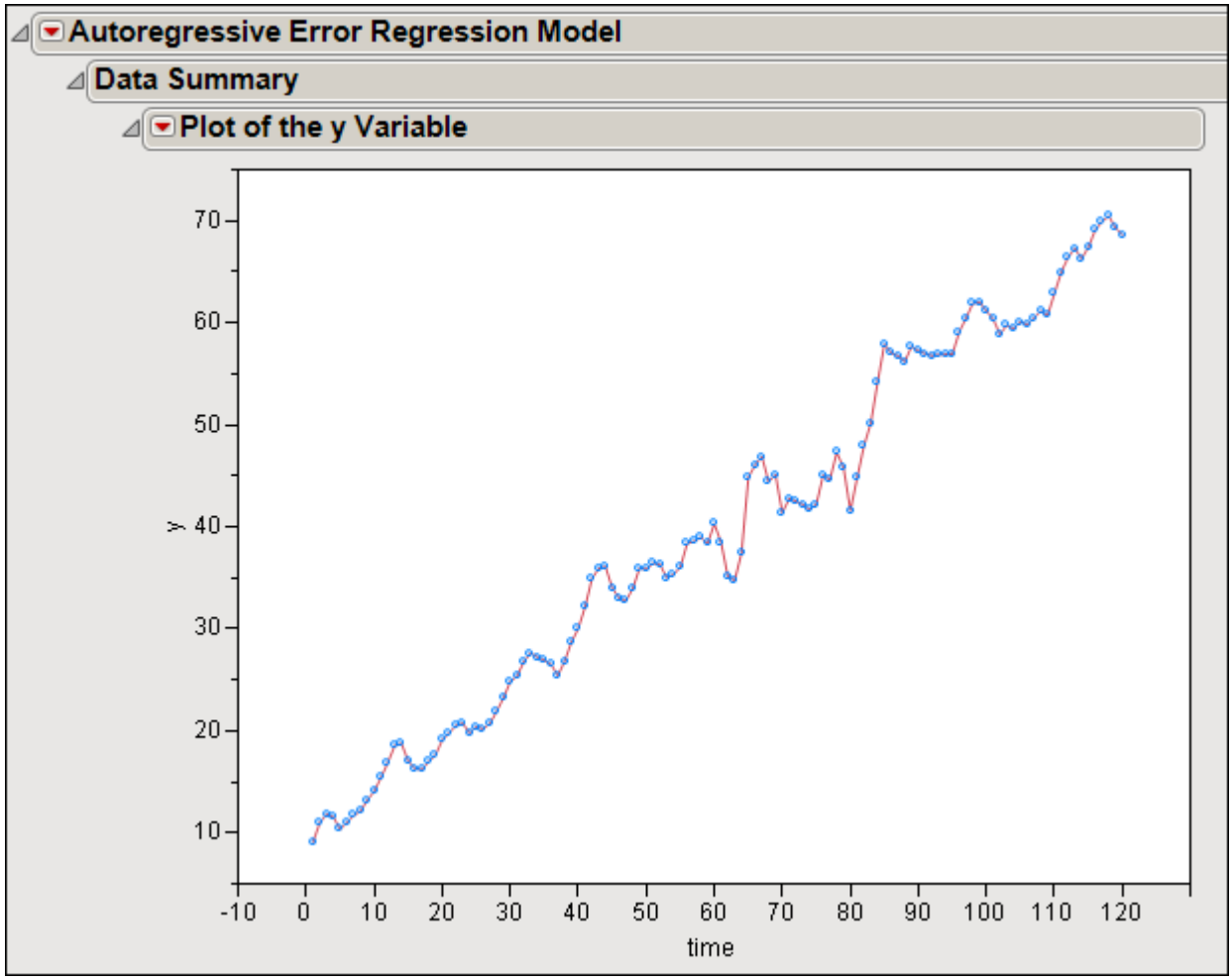
- **Data Summary**
- **Fitted Models**
- **Model Comparison**
- **Model Estimates**

Each of these areas is described in detail in the following sections.

Data Summary Area

The **Data Summary** area includes plots and tables which depend on the selected analysis. An example of the **Data Summary** area for **Fit Y by X with Autocorrelation and Heteroscedastic Errors** is shown in Figure 2.2.

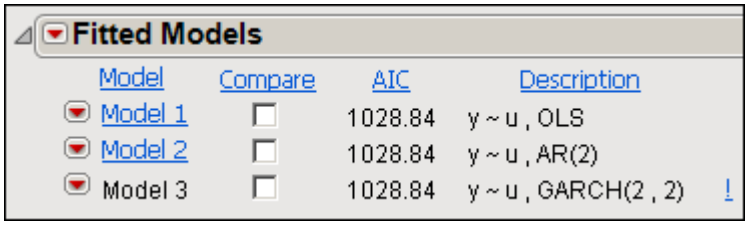
Figure 2.2 Example Results Window with Data Summary Area



Fitted Models Area

The **Fitted Models** area contains a list of the models that have been estimated. For each model, the list has a model name, a check box for including the model in the **Model Comparison** area, model fit statistics, and a description of the model. Click a model name to view the parameter estimates in the **Model Estimates** area. Click the column names to sort the models. For example, click a fit statistic title such as **AIC** (Akaike's information criterion) to sort the models by model fit values. Figure 2.3 shows an example of the **Fitted Models** area for **Fit Y by X with Autocorrelation and Heteroscedastic Errors**.

Figure 2.3 Example Results Window with Fitted Model Area



The **Fitted Models** area has two types of red triangle menus: one type is next to the **Fitted Models** heading, and the other type is next to each model's name.

Fitted Models Red Triangle Menu

The following commands are available in the red triangle menu next to the **Fitted Models** heading for all Results windows:

- **New Model** opens a dialog window. See the section "Dialog Window" on page 10 for more information.
- **Save** saves the model results in the current window. The saved results can be reproduced on another computer. All of the results from each model in the list of models are saved.
- **Rerun all** reruns all of the models. When the model properties in the Settings windows are changed, the changes are not implemented immediately. Select **Rerun all** to implement any changes to the settings for all models.
  NOTE: It takes longer to rerun all of the models if there are several models to rerun.
- **SAS Graphs Settings** opens the SAS Graphs Settings window. See the section "SAS Graphs Settings Window" on page 11 for more information.
- **Display Settings** opens the Display Settings window. See the section"Display Settings Window" on page 12 for more information.

The menu might also contain commands that are associated with a specific analysis option. More details about these menu commands can be found in the chapter for each analysis.

Model Name Red Triangle Menu

The following commands are available in the red triangle menu next to each model name:

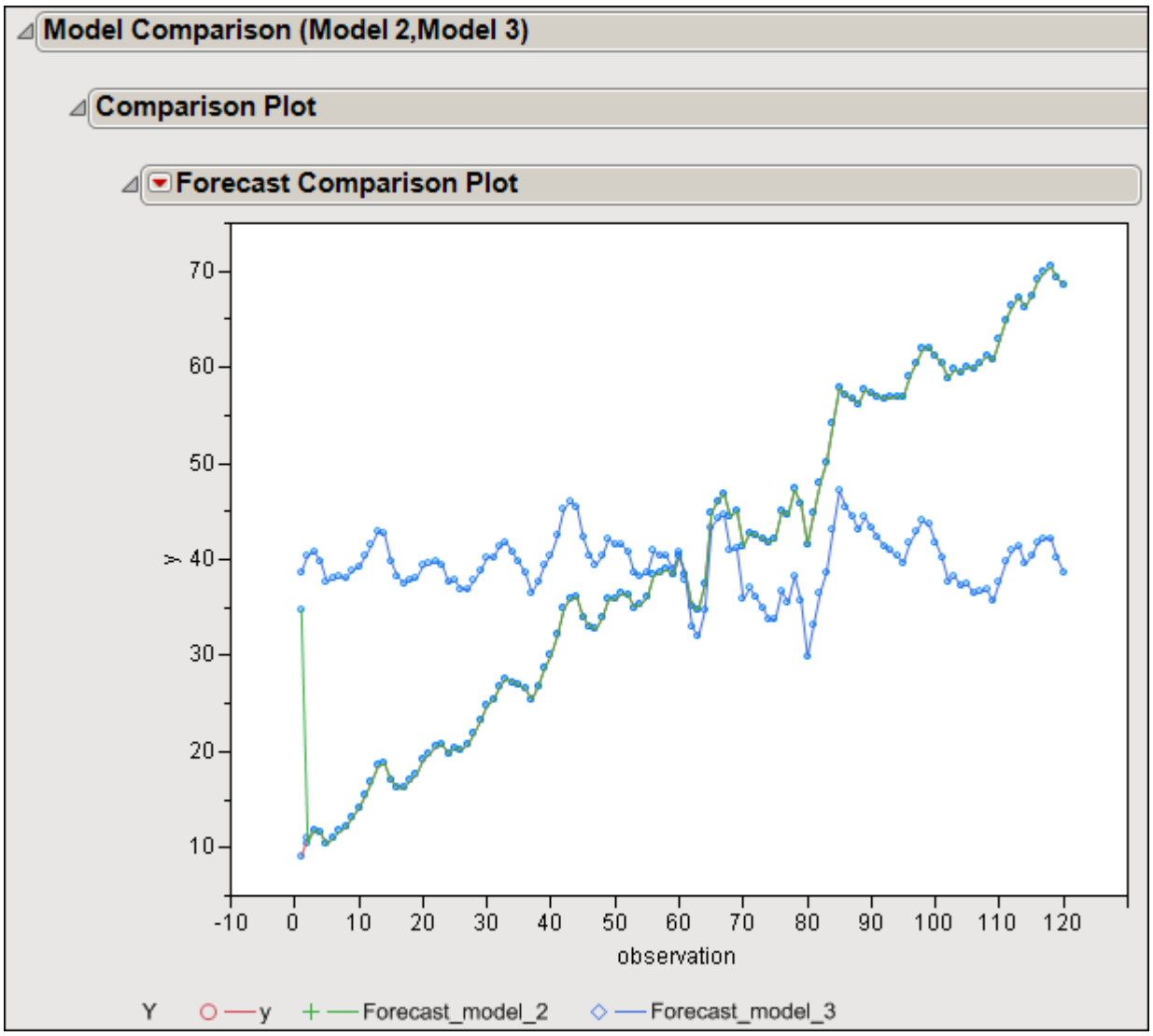
- **Edit** opens a dialog window with the current model's settings.
- **Delete** deletes the model.
- **Rename** renames the model.
- **Copy** copies the model.

- **Rerun** reruns the model.
- **View SAS Log** opens a window with the SAS Log for the model.
- **View/edit SAS code** opens a SAS Program Editor window with the SAS code for the model. You can edit the SAS code for the model in this window.

Model Comparison Area

Figure 2.4 shows an example of the **Model Comparison** area. The **Model Comparison** area is closed by default. To open this area, select one or more of the **Compare** boxes next to the model names in the **Fitted Models** area. If only one model is selected, the original data and the fitted data are plotted in an overlay plot for comparison and a plot of the residuals is provided. If several models are selected, the fitted data for each model are plotted in one overlay plot with a different color for each model for both the fitted values and residuals plots.

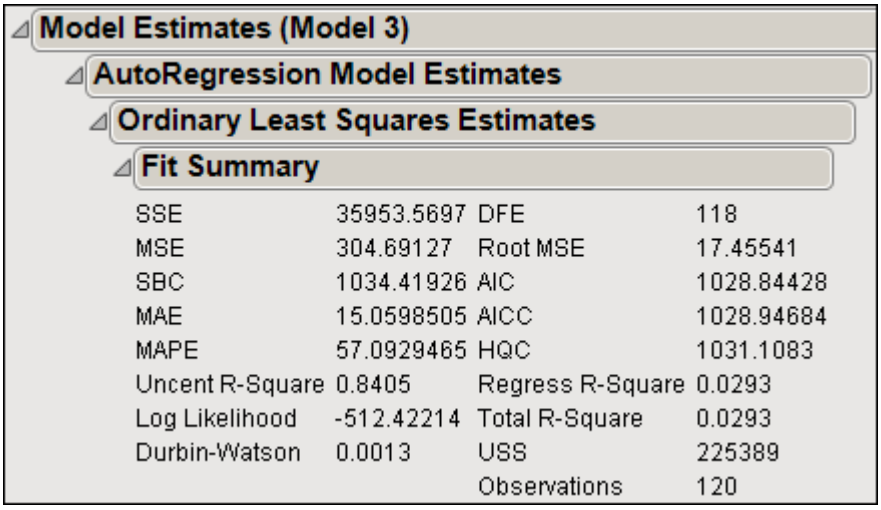
Figure 2.4 Example Results Window with Model Comparison Area



Model Estimates Area

The **Model Estimates** area displays the results generated by the SAS/ETS procedure. Figure 2.5 shows an example of the **Model Estimates** area.

Figure 2.5 Example Results Window with Model Estimates Area



Dialog Window

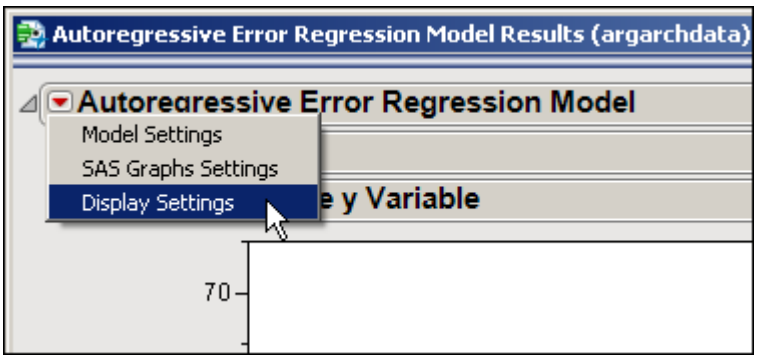
When you create a new model or edit an existing model, you specify the model in the dialog window. When you create a new model, the window starts with default values; when you edit an existing model, the saved values of the current model appear in the window. The dialog windows are different for each analysis. See the following sections for more information:

- "Autoregressive Error Regression Model Parameters Dialog Window" on page 36
- "Panel Data Regression Parameters Dialog Window" on page 54
- "The Unobserved Components Model Parameters Dialog Window" on page 68

Settings Windows

Settings windows save preferences about graphics, results display, and other model settings. The settings are saved locally. Once the settings are changed, the changes remain until you specify other settings. To access settings windows, click on the red triangle in the upper left corner of the results window. See an example of a settings menu in Figure 2.6. All analyses have a SAS Graphs Settings window and a Display Settings window. Some analyses provide additional settings windows (see the sections "Model Settings Window" on page 42 and "Forecast Settings Window" on page 78).

Figure 2.6 Settings Menu

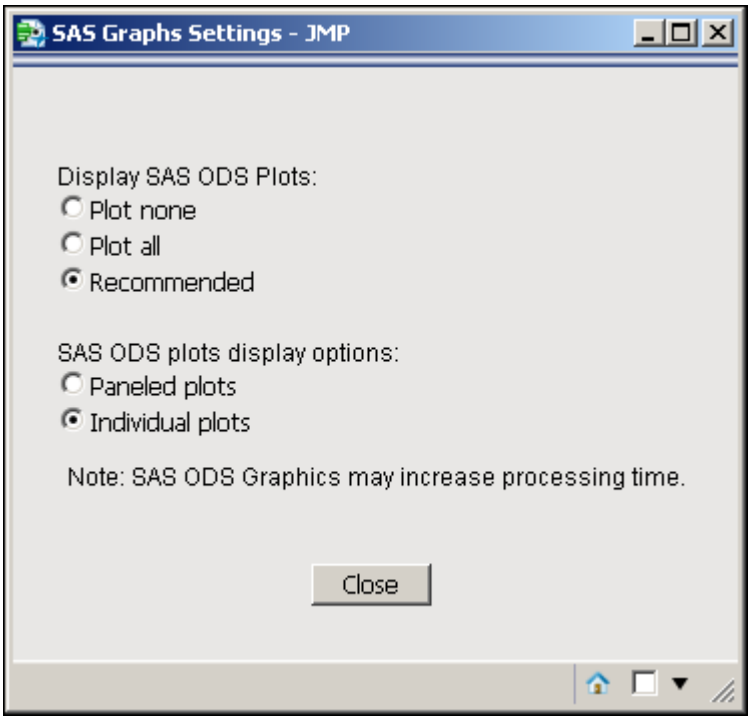


SAS Graphs Settings Window

In the SAS Graphs Settings window, you specify which SAS graphs you want to be created. Figure 2.7 shows an example of the SAS Graphs Settings window. You can select paneled or individual plots. You can also select to plot none, all, or only recommended SAS ODS plots.

NOTE: Generating all SAS graphs might take longer.

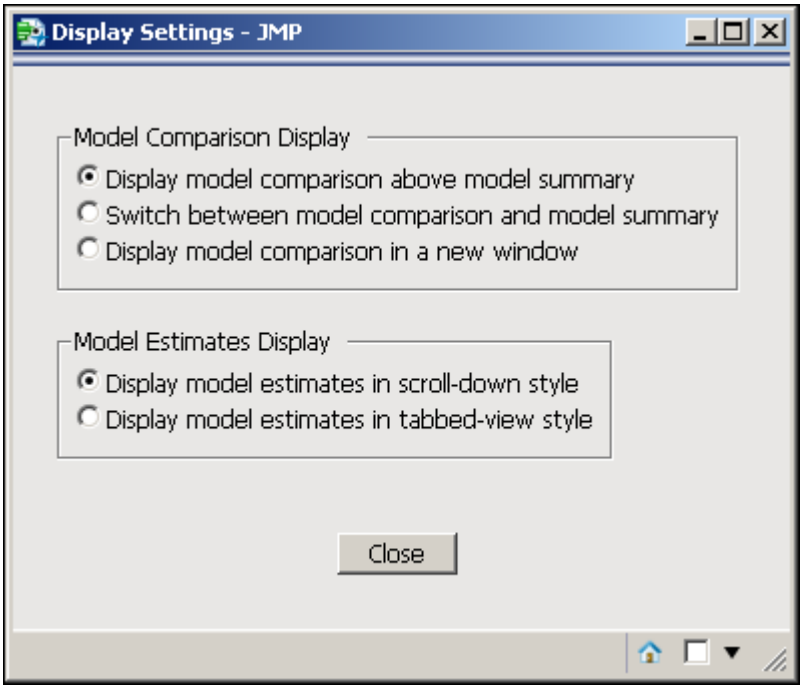
Figure 2.7 SAS Graphics Settings Window



Display Settings Window

In the Display Settings window, you specify the layout of the **Model Comparison** and **Model Estimates** areas of the Results window. Figure 2.8 shows the Display Settings window.

Figure 2.8 Display Settings Window



Chapter 3

Getting Started with SAS Econometrics and Time Series Analysis for JMP

Contents

Overview of Getting Started Example

This getting started example shows you how to create and perform a panel data analysis in SAS Econometrics and Time Series Analysis for JMP. This example begins with a description of the data and the example models; then it shows you the steps for performing an analysis, including how to establish a connection to

SAS, open a data set, create the panel data analysis base model, analyze the model, view the results, create an alternative model, and compare the models. This example also shows you how to save the analysis and open the results later.

This example does not describe all the models you can create or all of the features of SAS Econometrics and Time Series Analysis for JMP. Throughout this example are directions about where to find more information about other features of SAS Econometrics and Time Series Analysis for JMP.

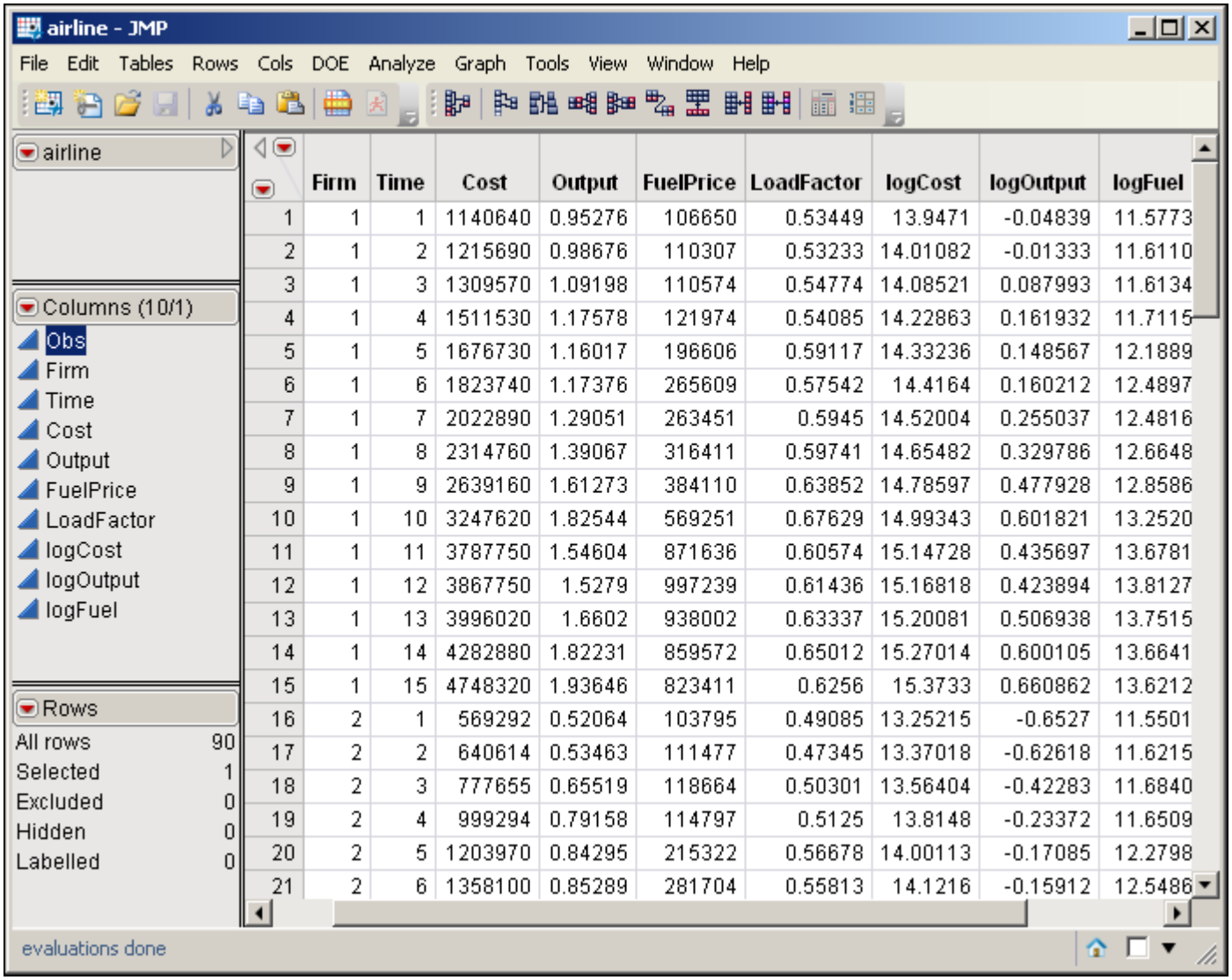
Data Description

This example uses the *airline.jmp* data set, which is the frequently cited Christenson Associates airline data (Greene 1999). The data measure costs, prices of inputs, and utilization rates for six airlines over the 15 years (the observations span 1970–1984). The variables include:

- Firm: firm ID
- Time: time period ID
- Cost: total cost, in thousands
- Output: output in revenue passenger miles (index)
- FuelPrice: fuel price
- LoadFactor: load factor (utilization index)
- logCost: log transformation of costs
- logOutput: log transformation of quantity
- logFuel: log transformation of price of fuel

Figure 3.1 shows part of the data set.

Figure 3.1 Getting Started Example: Airline Data



	Firm	Time	Cost	Output	FuelPrice	LoadFactor	logCost	logOutput	logFuel
1	1	1	1140640	0.95276	106650	0.53449	13.9471	-0.04839	11.5773
2	1	2	1215690	0.98676	110307	0.53233	14.01082	-0.01333	11.6110
3	1	3	1309570	1.09198	110574	0.54774	14.08521	0.087993	11.6134
4	1	4	1511530	1.17578	121974	0.54085	14.22863	0.161932	11.7115
5	1	5	1676730	1.16017	196606	0.59117	14.33236	0.148567	12.1889
6	1	6	1823740	1.17376	265609	0.57542	14.4164	0.160212	12.4897
7	1	7	2022890	1.29051	263451	0.5945	14.52004	0.255037	12.4816
8	1	8	2314760	1.39067	316411	0.59741	14.65482	0.329786	12.6648
9	1	9	2639160	1.61273	384110	0.63852	14.78597	0.477928	12.8586
10	1	10	3247620	1.82544	569251	0.67629	14.99343	0.601821	13.2520
11	1	11	3787750	1.54604	871636	0.60574	15.14728	0.435697	13.6781
12	1	12	3867750	1.5279	997239	0.61436	15.16818	0.423894	13.8127
13	1	13	3996020	1.6602	938002	0.63337	15.20081	0.506938	13.7515
14	1	14	4282880	1.82231	859572	0.65012	15.27014	0.600105	13.6641
15	1	15	4748320	1.93646	823411	0.6256	15.3733	0.660862	13.6212
16	2	1	569292	0.52064	103795	0.49085	13.25215	-0.6527	11.5501
17	2	2	640614	0.53463	111477	0.47345	13.37018	-0.62618	11.6215
18	2	3	777655	0.65519	118664	0.50301	13.56404	-0.42283	11.6840
19	2	4	999294	0.79158	114797	0.5125	13.8148	-0.23372	11.6509
20	2	5	1203970	0.84295	215322	0.56678	14.00113	-0.17085	12.2798
21	2	6	1358100	0.85289	281704	0.55813	14.1216	-0.15912	12.5486

Example Models Description

For this example, you create two models using **Fit Panel Data Regression**: a base model and an alternative model. Both models have LoadFactor, logOutput, and logFuel as explanatory variables and logCost as the response variable. The base model is a two-way fixed-effects model, which is the default model for **Fit Panel Data Regression**. The alternative model is a one-way, cross-sectional fixed-effects model, which ignores any time effects. One goal of this analysis is to determine which model is a better representation of the data. More specifically, does a model that includes a fixed effect of time fit better than a model that ignores time? A second goal is to determine how the explanatory variables are related to cost and which variables are significant predictors of cost.

Perform an Analysis with SAS Econometrics and Time Series Analysis for JMP

Establish a Connection to SAS

To use SAS Econometrics and Time Series Analysis for JMP, you need access to the SAS/ETS software. To create a connection to the SAS System:

1. Open JMP 9.0.2 or later.
2. Select **File ► SAS► Server Connection**. The SAS Server Connections window appears.
3. In the SAS Server Connections window, connect to the SAS System as described in one of the following subsections and then click **Close** to close the SAS Server Connections window.

Connect to SAS on This Machine

1. Select **Connect to SAS on this machine** in the **Establish New Workspace Server Connection** area.
2. Click **Connect**. When the **Open Workspace Server Connections** area displays "Current active connection: Local," you are connected to the SAS System.

Connect to a SAS Metadata Server

1. Click **Metadata Server Profiles**. The Manage Metadata Server Profiles window appears.
2. Specify the appropriate metadata server information in the Manage Metadata Server Profiles window.
3. Click **Connect** to connect to the SAS Metadata Server. When the **Status** displays "Connected," you are connected to the SAS Metadata Server.
4. Click **Close** to close the Manage Metadata Server Profiles window.
5. Verify that the **Connect to metadata-defined SAS server** option is selected in the **Establish New Workspace Server Connection** area.
6. Click **Connect**. When the **Open Workspace Server Connections** area displays "Current active connection" followed by the name of the metadata server, you are connected to the SAS System.

Connect to a Remote SAS Server

1. Select **Connect to remote SAS Server on** in the **Establish New Workspace Server Connection** area.
2. Enter the machine name and port number in the appropriate boxes.

3 Click **Connect**. When the **Open Workspace Server Connections** area displays "Current active connection" followed by the remote SAS server information, you are connected to the SAS System.

Open the Data Set

Before you can begin an analysis, you need to open a data file for analysis. To open the *airline.jmp* data file:

1 Select **File ► Open**.

2 Select *airline.jmp* from the appropriate location.

3 Click **Open**.

Create and Estimate the Base Model

Open a New Analysis

To start a panel data analysis, select **Analyze ► Econometrics and Time Series ► Fit Panel Data Regression**. A Panel Data Analysis window appears. You specify the base model in this window.

Assign Variables to Roles

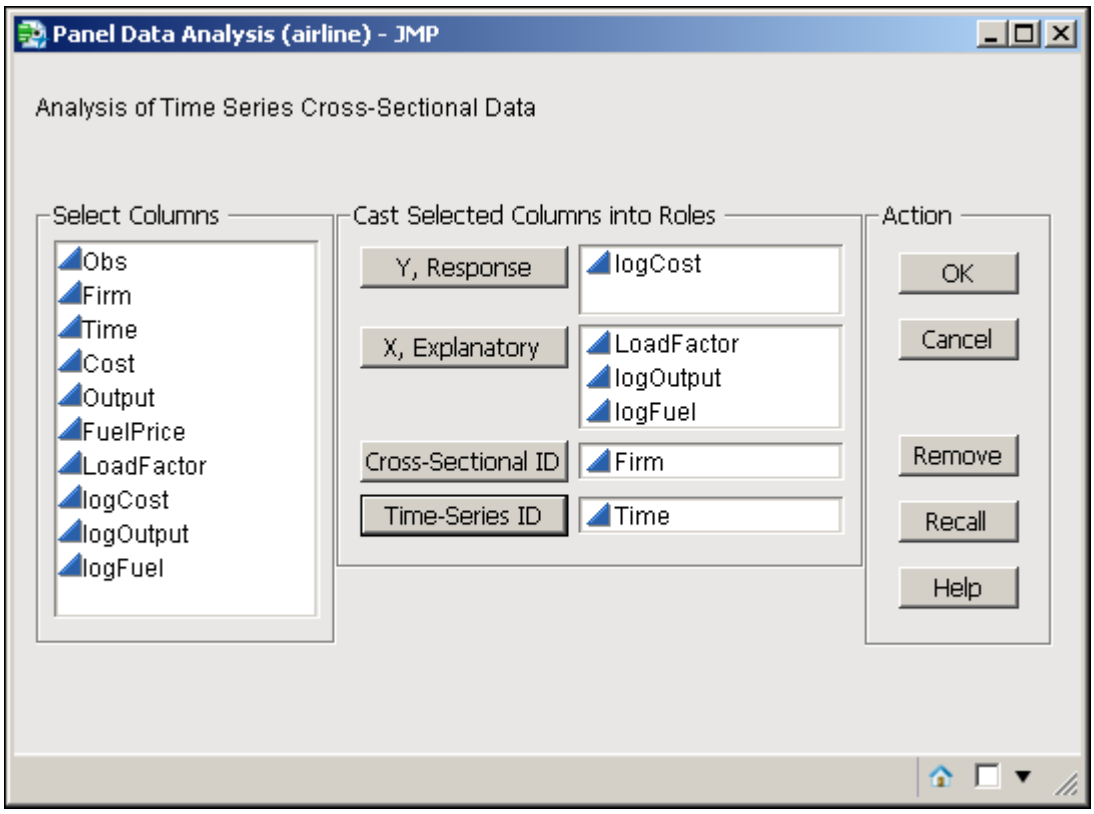
To specify a base model, you need to assign variables to roles. To assign a variable to a role:

1 Select the variable from the **Select Columns** list.

2 Click the appropriate role button in the **Cast Selected Columns into Roles** area.

For this example, you need to assign logCost as the response variable, Firm as the cross-sectional ID variable, and Time as the time series ID variable. You also need to assign logOuput, logFuel, and LoadFactor as the explanatory variables. These model specifications are shown in Figure 3.2.

Figure 3.2 Getting Started Example Base Model Specifications



Estimate the Base Model

After you have assigned all the relevant variables to roles, click **OK** to estimate the base model.

Examine the Base Model Results

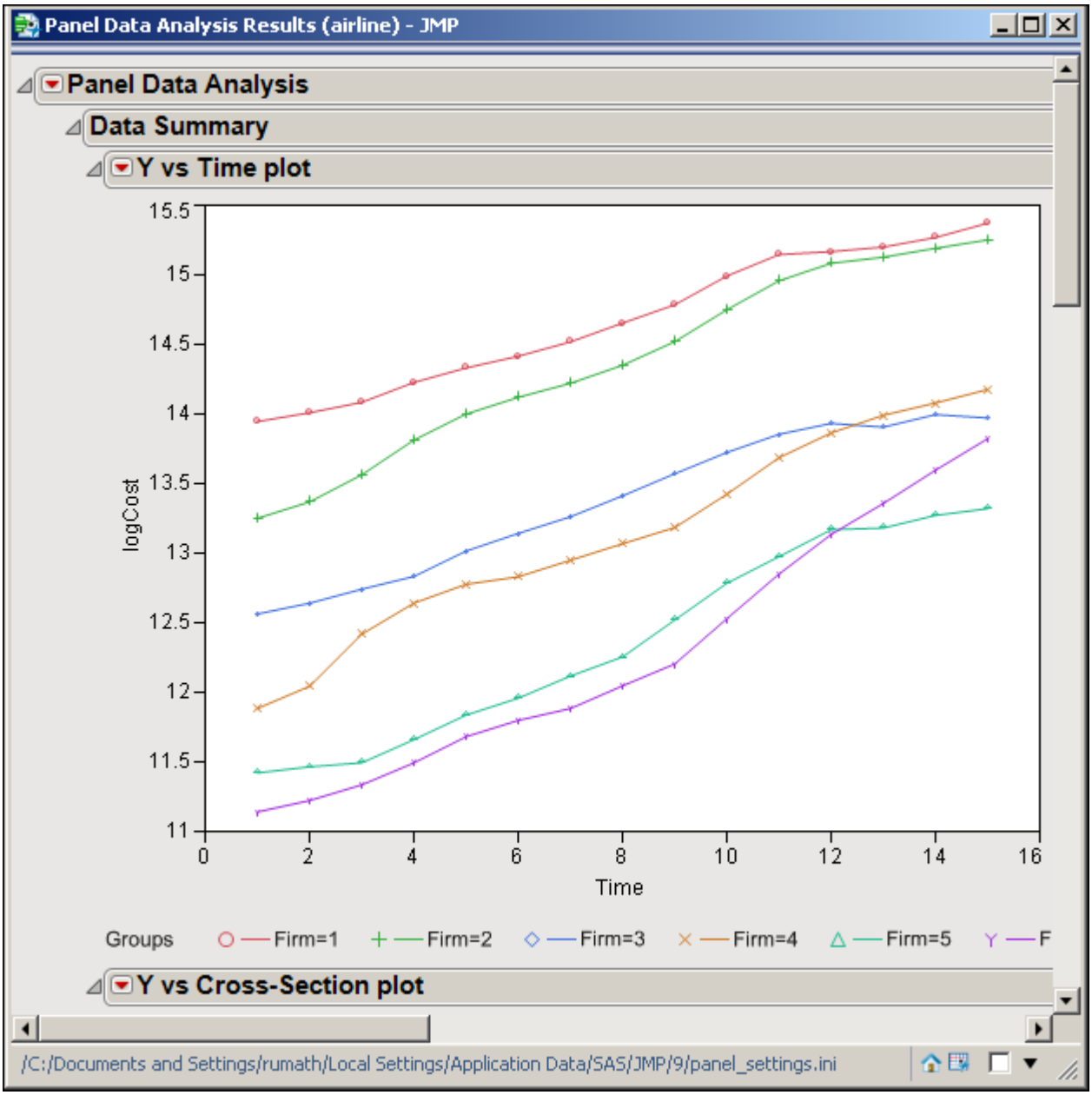
After you specify and estimate the base model, the results window appears. You use the information in the results window to evaluate model fit and decide how to create alternative models.

Data Summary

The **Data Summary** area is at the top of the results window. It contains plots and tables of the observed data. In this area, you can examine the data for any trends or unusual observations. The information that appears in the **Data Summary** area depends on which analysis you select; this example shows the plots and tables included for a panel analysis. You can modify the view of the plots and tables in the **Data Summary** area by clicking on the red triangles next to the headings. You can also add plots or tables to the **Data Summary** area by clicking on the red triangles next to the headings.

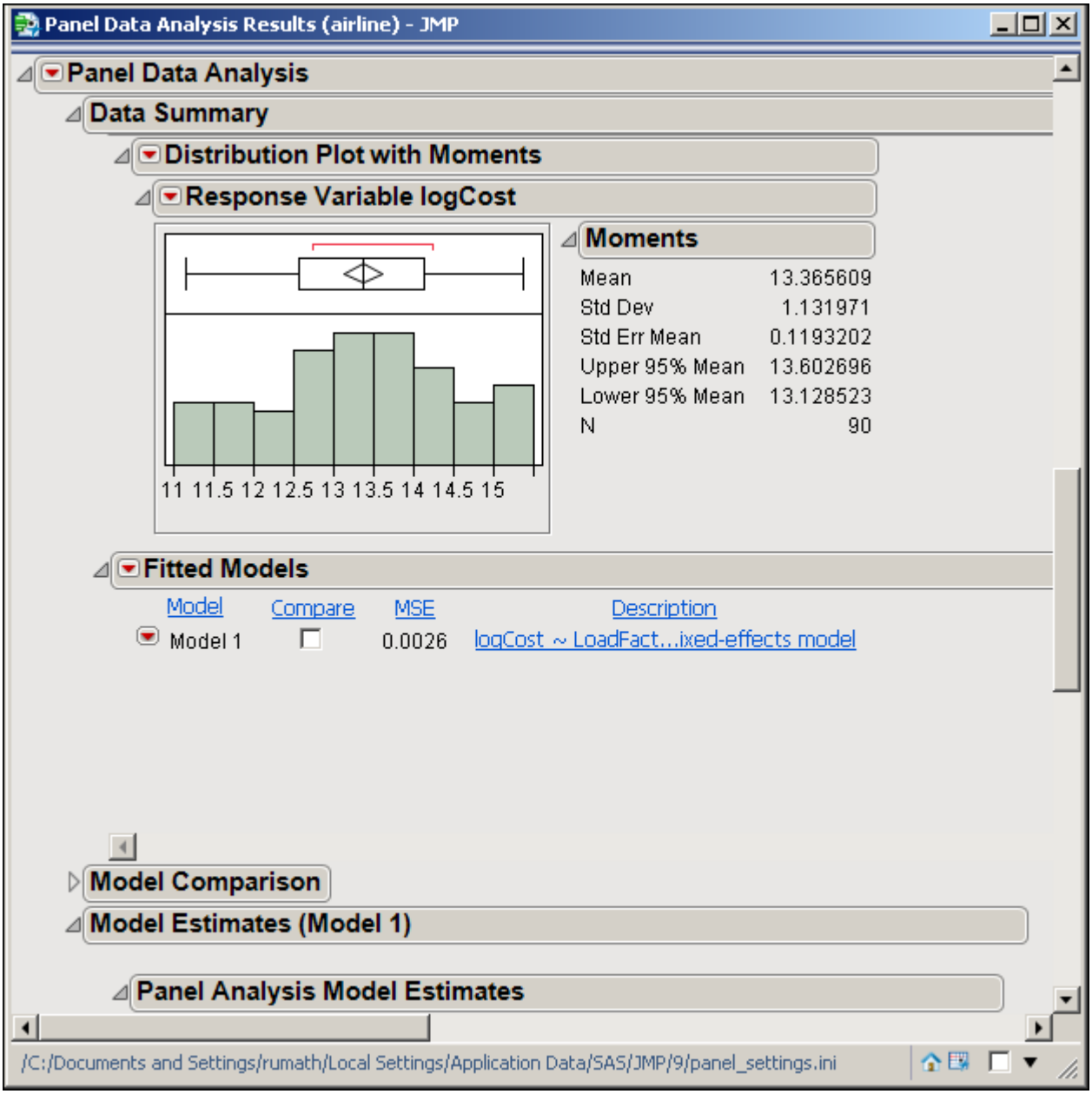
A **Y vs Time plot** of logCost over time for each airline shows that overtime costs are increasing for all of the airlines. See Figure 3.3.

Figure 3.3 Y By Time Plot in the Results Window



Next is a **Y vs Cross-Section** plot of logCost by airline for each time. Following these plots is a distribution plot and a table of moments for logCost. See Figure 3.4.

Figure 3.4 Distribution Plot and Fitted Models Area of the Results Window



Fitted Models

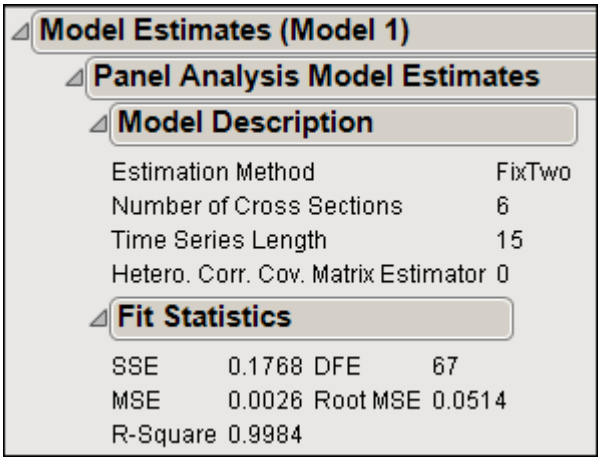
The **Fitted Models** area follows the **Data Summary** area. It contains a list of the models that have been estimated. The model you estimated is called Model 1. Each model in the **Fitted Models** area has a model name, a check box for including the model in the **Model Comparison** area, the mean squared error (MSE) value for the model, and a description of the model. When you click a model name, you can view the parameter estimates for that model in the **Model Estimates** area. Because only one model is estimated, the estimates for that model appear in the **Model Estimates** area.

Model Estimates

The last section of the results window is the **Model Estimates** area. The tables in the Model Estimates area depend on which analysis you select. For a panel analysis, the **Model Estimates** area contains tables for the **Model Description**, **Fit Statistics**, **Parameter Estimates**, **Additional Results**, and **Panel Analysis Model Graphics**. If you have also selected an option to produce SAS ODS graphs in the SAS Graphs Settings window (see the section "SAS Graphs Settings Window" on page 11), those graphs also appear.

To evaluate the base model, you first examine the **Fit Statistics** table. Figure 3.5 shows the **Fit Statistics** table for the base model.

Figure 3.5 Fit Statistics for the Base Model



You see a large R-square in the **Fit Statistics** table, so there is evidence for good model fit.

After examining the **Fit Statistics** table, you examine the parameter estimates in the **Parameter Estimates** area. See Figure 3.6.

Figure 3.6 Parameter Estimates for the Base Model

Parameter Estimates

Variable	DF	Estimate	Standard Error	t Value	Pr > |t|	Label
CS1	1	0.17424	0.06648	2.62083	0.0108*	Cross Sectional Effect 1
CS2	1	0.11141	0.05756	1.9357	0.0571	Cross Sectional Effect 2
CS3	1	-0.1435	0.04271	-3.3612	0.0013*	Cross Sectional Effect 3
CS4	1	0.18019	0.02752	6.54839	<.0001*	Cross Sectional Effect 4
CS5	1	-0.0467	0.02237	-2.0884	0.0406*	Cross Sectional Effect 5
TS1	1	-0.6929	0.21389	-3.2393	0.0019*	Time Series Effect 1
TS2	1	-0.6382	0.21038	-3.0334	0.0034*	Time Series Effect 2
TS3	1	-0.5955	0.20937	-2.8445	0.0059*	Time Series Effect 3
TS4	1	-0.5419	0.20238	-2.6777	0.0093*	Time Series Effect 4
TS5	1	-0.4729	0.1496	-3.1609	0.0024*	Time Series Effect 5
TS6	1	-0.4271	0.12232	-3.4911	0.0009*	Time Series Effect 6
TS7	1	-0.3959	0.11294	-3.5051	0.0008*	Time Series Effect 7
TS8	1	-0.3397	0.0987	-3.4421	0.0010*	Time Series Effect 8
TS9	1	-0.2718	0.09119	-2.9808	0.0040*	Time Series Effect 9
TS10	1	-0.2273	0.05619	-4.0462	0.0001*	Time Series Effect 10
TS11	1	-0.1118	0.03066	-3.6462	0.0005*	Time Series Effect 11
TS12	1	-0.0337	0.03494	-0.9633	0.3389	Time Series Effect 12
TS13	1	-0.0177	0.03222	-0.5507	0.5837	Time Series Effect 13
TS14	1	-0.0187	0.02997	-0.6225	0.5357	Time Series Effect 14
Intercept	1	12.9383	1.36598	9.47182	<.0001*	Intercept
LoadFactor	1	-0.8827	0.24638	-3.5826	0.0006*	
logOutput	1	0.81726	0.02237	36.5372	<.0001*	
logFuel	1	0.16873	0.10289	1.63999	0.1057	

Additional Results

Panel Analysis Model Graphics

There are several important results in the **Parameter Estimates** area:

- Most of the cross-sectional effects are highly significant (with the exception of CS2). This means that the cross sections are significantly different from the sixth cross section (which is the reference section).
- Many of the time effects show significance, but not uniformly.
- The time dummy variables taper off in size and lose significance from time period 12 onward.
- As for the regression parameters:
  - Output affects cost positively.
  - Fuel price affects cost positively.
  - Load factor output affects cost negatively.
- The unusual result is that the fuel cost is not significant.

Based on these results, you decide to examine whether the results change if you ignore the effect of time and fit a one-way, cross-sectional fixed-effects model.

Create an Alternative Model

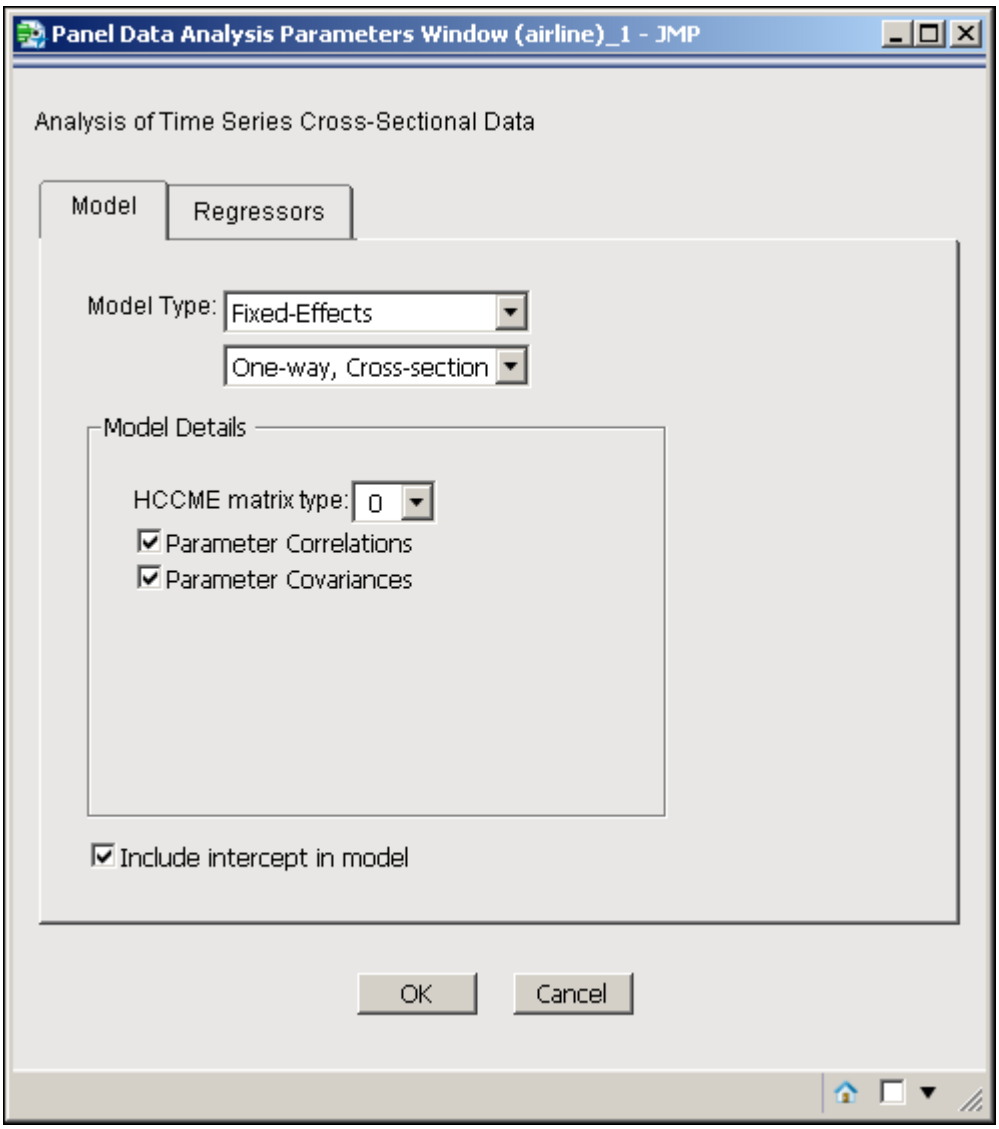
To create a one-way fixed-effects model, you open a dialog window called the Panel Data Analysis Parameters window.

There are several ways to open a dialog window:

- Click the red triangle next to **Fitted Models**, and select **New Model**.
- Click the red triangle next to Model 1, and select **Edit**. If you open the dialog this way, change values, and click **Ok** in the dialog window, any results from the base model are erased, and you can no longer compare the base model with the new alternative model.
- Click the red triangle next to Model 1, and select **Copy**. A new model appears on the **Fitted Models** list; it is called Copy of Model 1. This model is identical to the base model, Model 1. To change this copy of Model 1 into a one-way fixed-effects model that ignores time, click the red triangle next to Copy of Model 1 and select **Edit**.

After you open a dialog window, you change the model to a one-way, cross-sectional fixed-effects model by clicking the arrow next to **Two-way** and selecting **One-Way, Cross-section**. See Figure 3.7. Then click **OK** to estimate the model.

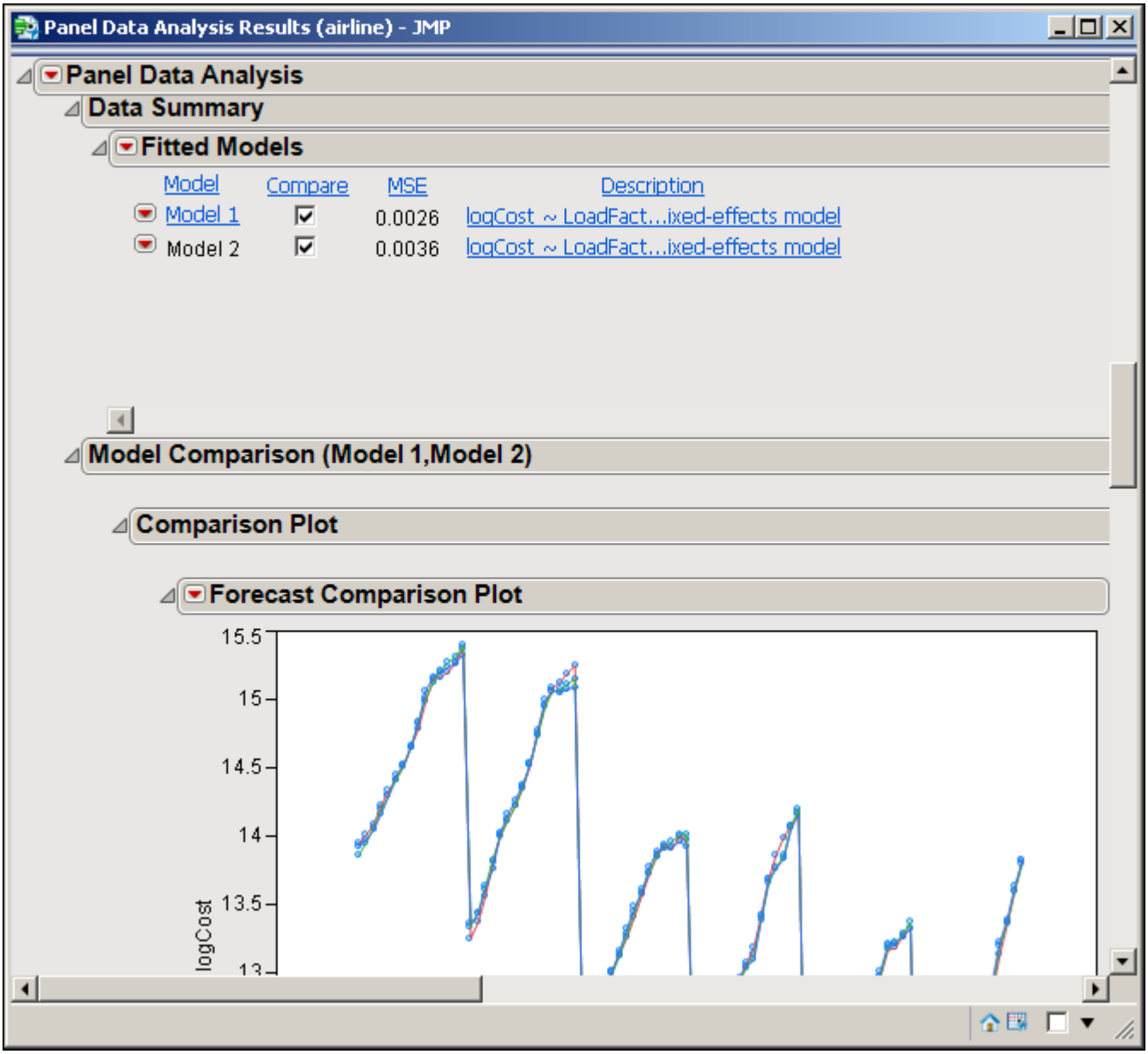
Figure 3.7 Alternative Model Specifications



Compare the Base and Alternative Models

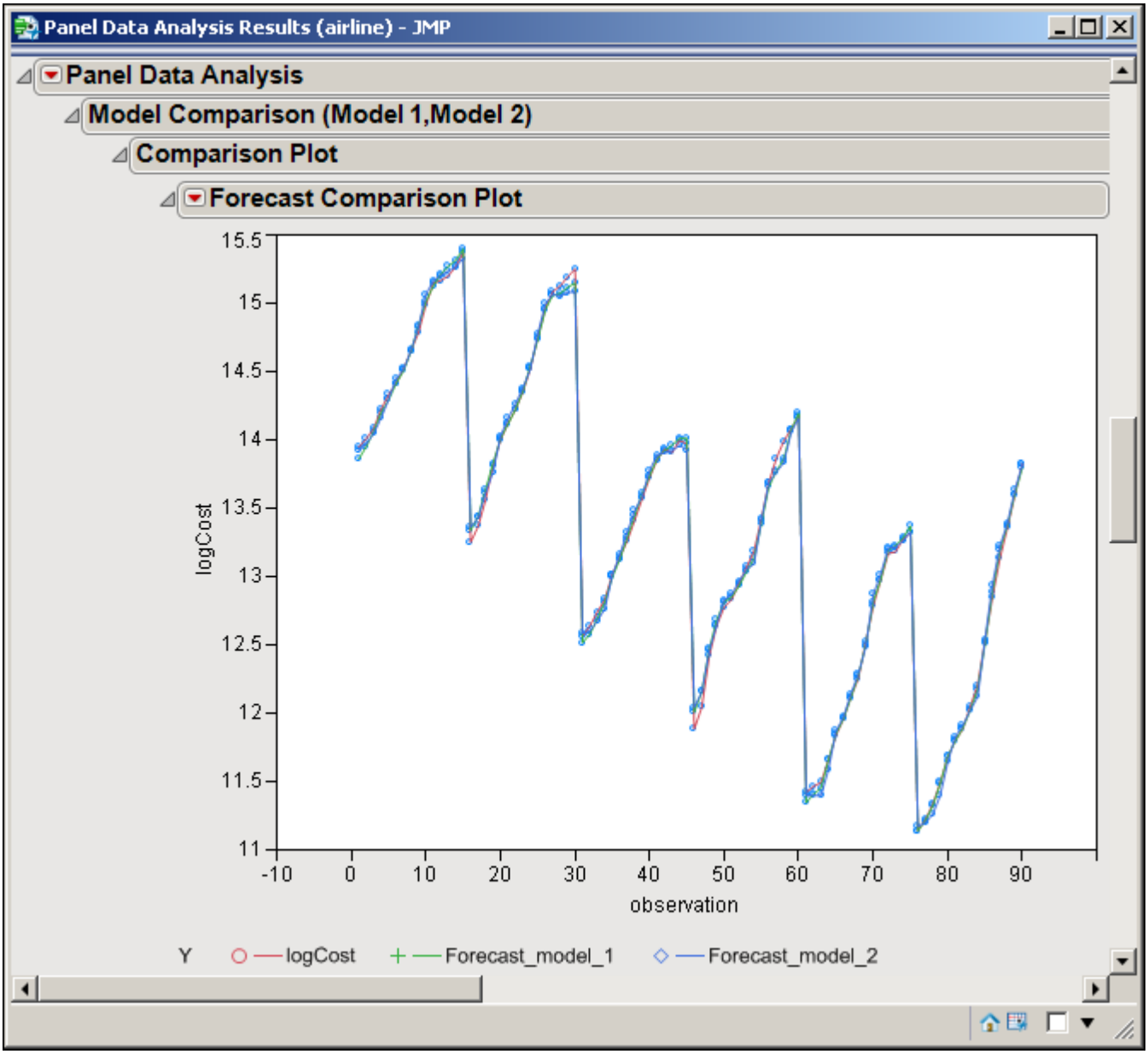
After you estimate the alternative model, the results window is updated with the results for the alternative model. In the **Fitted Models** area, the list of the models contains Model 1 and Model 2. See Figure 3.8.

Figure 3.8 Fitted Model Area with the Base and Alternative Models



Each model in the **Fitted Models** area has a model name, a check box for including the model in the **Model Comparison** area, the MSE value for the model, and a description of the model. You can see that the MSE for Model 1 is lower, suggesting that model fit is better for Model 1 than Model 2. You select both check boxes in the **Compare** column to include both models in the **Model Comparison** area. You examine the **Model Comparison** area to decide which model is a better representation for the data. Figure 3.9 shows the **Forecast Comparison Plot**.

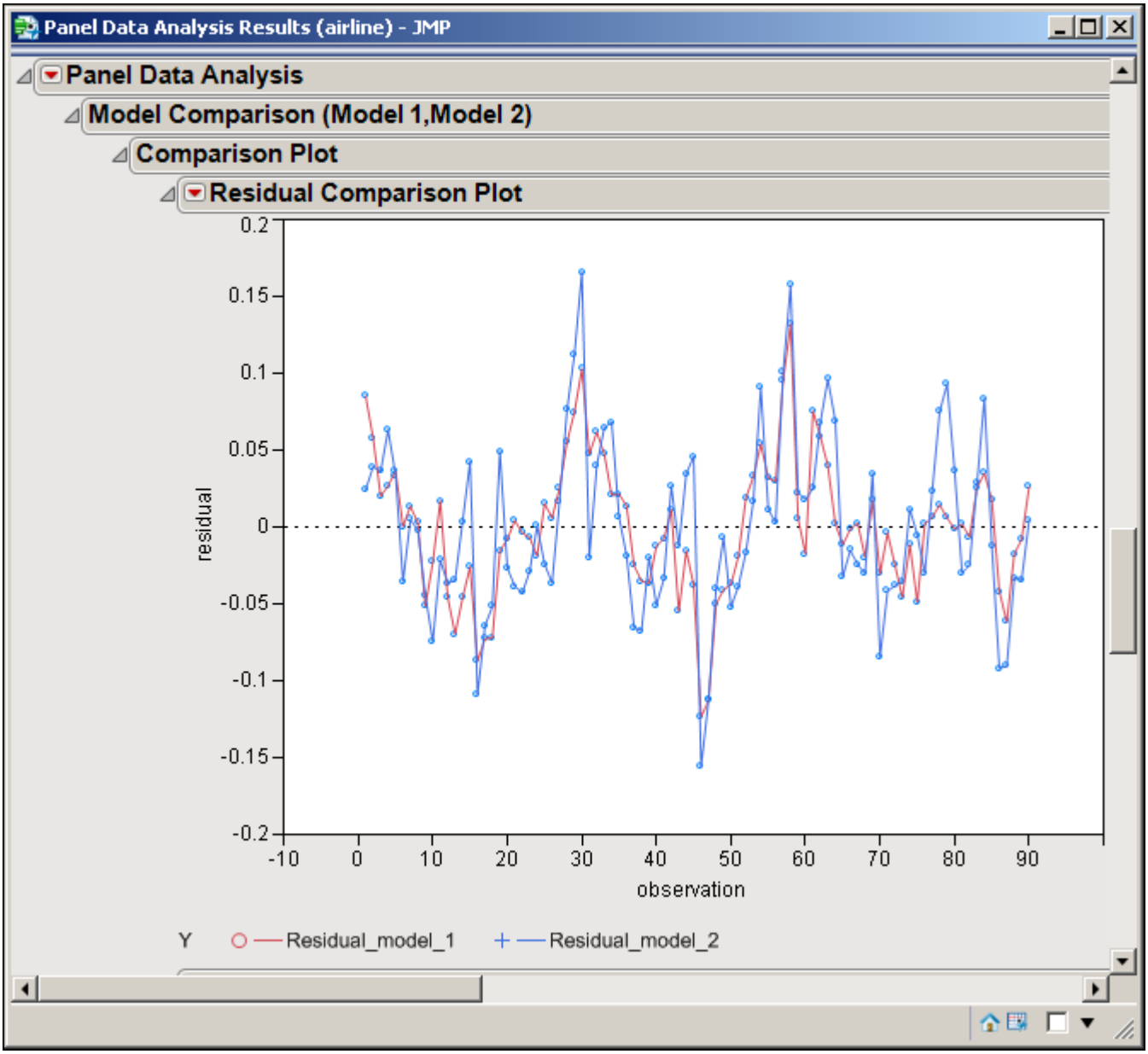
Figure 3.9 Forecast Comparison Plot with the Base and Alternative Models



In this plot, the red line represents the observed logCost, the green line represents the forecast from the base model, and the blue line represents the forecast for the alternative model. You can see that the forecasts for both models are nearly the same.

Below the **Forecast Comparison Plot** is a **Residual Comparison Plot**, as shown in Figure 3.10.

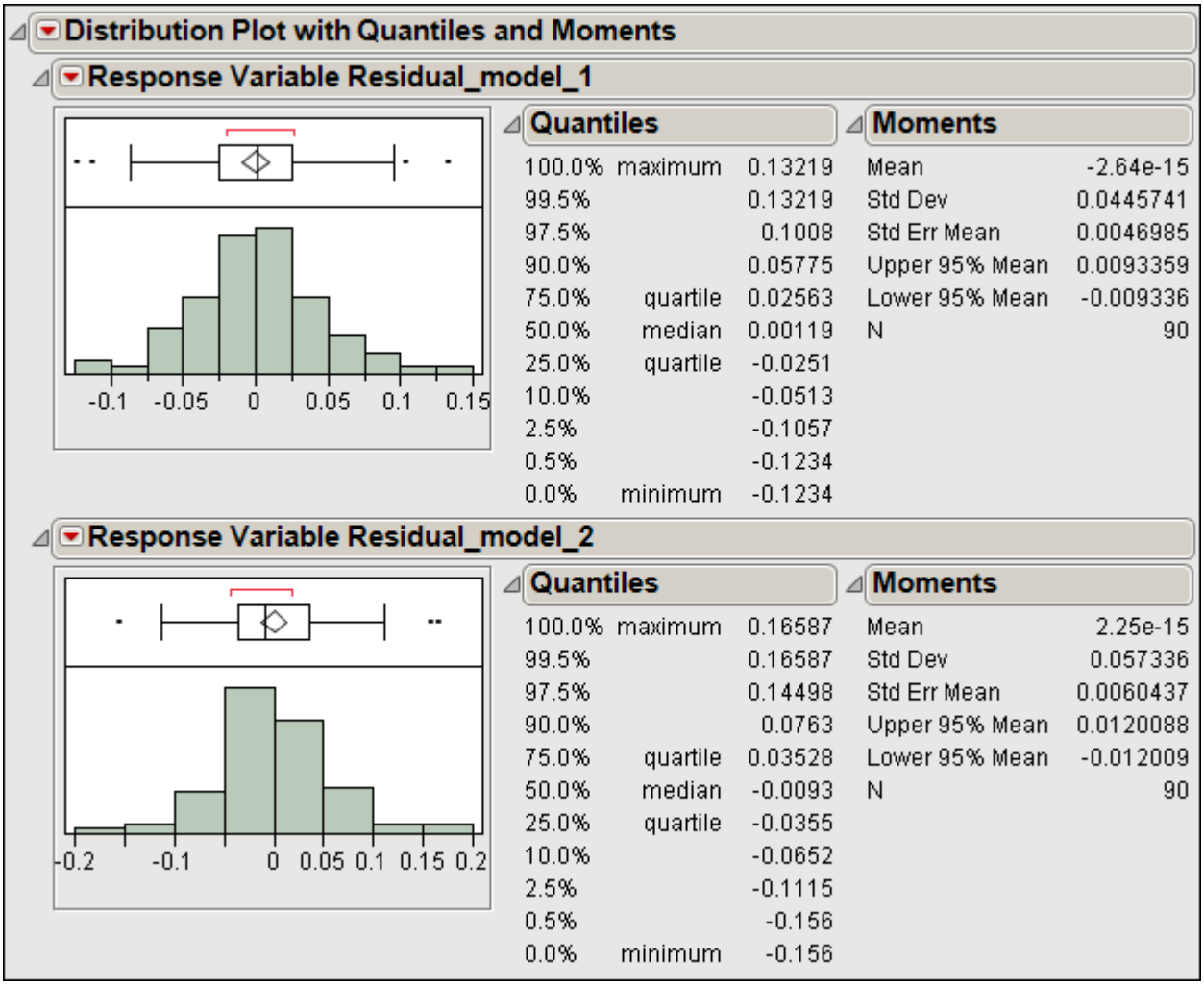
Figure 3.10 Residual Comparison Plot with the Base and Alternative Models



This plot contains the residuals for the base model (in red) and the alternative model (in blue). The red line for the base model appears to be more frequently closer to 0, which suggests that the forecast from base model is closer to the observed data.

The **Distribution Plot with Quantiles and Moments** area contains plots and tables of the residuals for the base and alternative models. See Figure 3.11. These tables and plots also indicate that the residuals for the alternative model tend to have greater variance and a wider range than those for the base model.

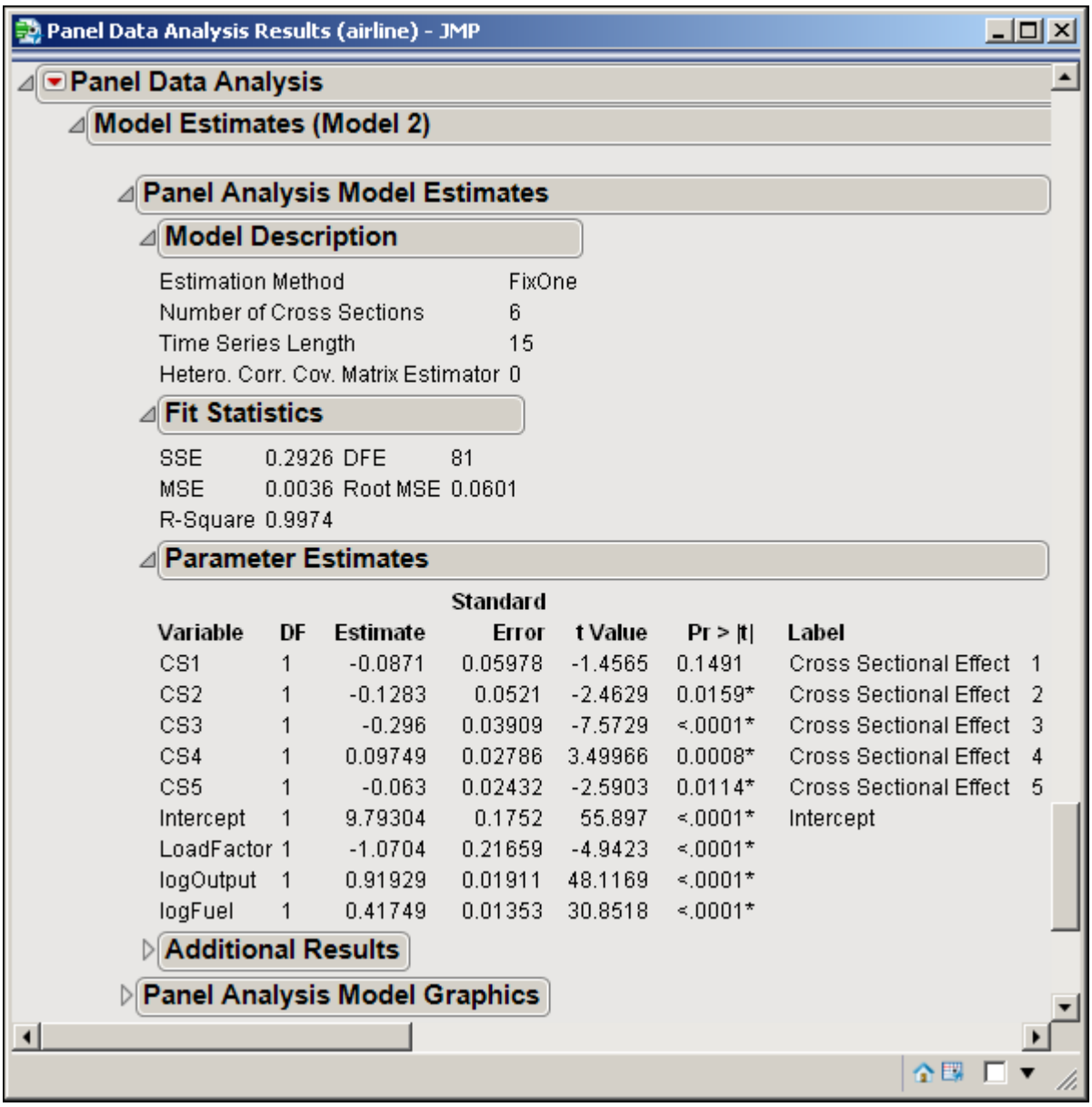
Figure 3.11 Response Residual Distribution Comparison for the Base and Alternative Models



In addition to the **Model Comparison** area, you can also compare the base and alternative models by comparing model fit statistics and the model parameter estimates, which can be found in the **Model Estimates** area. To view the alternative model estimates in the **Model Estimates** area, click Model 2 in the **Fitted Models** area.

As shown in Figure 3.12, the R-square statistic for the alternative model is still large, but the parameters change. The effect of fuel costs (logFuel) comes in as very strong and significant. In addition, the coefficient for LoadFactor increases, although not as dramatically.

Figure 3.12 Fit Statistics and Parameter Estimates for the Alternative Model



Overall, the model fit information and residual analysis suggest that base model has better fit.

Save and Open the Results

Save Results

After you build your models, you can save your work by using the following steps:

1. Click the red triangle next to the **Fitted Models** header.
2. Select **Save**. A SAS Program Editor window appears.
3. In the SAS Program Editor window, specify the name of the model and where you want the model to be saved.
4. Click **Save**.

Open Results

To open results:

1. Open JMP.
2. Open the data file:
    1. In the JMP Home window, select **File► Open**.
    2. Locate and select the data set you used for the analysis, in this case the airline.jmp data.
    3. Click **Open**.
4. Open the saved SAS Program Editor file:
    1. In the JMP Home window, select **File► Open**.
    2. Locate and select the saved SAS Program Editor file.
    3. Click **Open**.
4. In the SAS Program Editor window, select **Edit► Run Script**. The results window appears.

References

Greene, W. H. (1999), *Econometric Analysis*, Fourth Edition, Macmillan.

Chapter 4

Fit Y by X with Autocorrelation and Heteroscedastic Errors

Contents

Overview of Fit Y by X with Autocorrelation and Heteroscedastic Errors

You select **Fit Y by X with Autocorrelation and Heteroscedastic Errors** to fit autocorrelated regression models. Autocorrelated regression models examine time series data when the errors are autocorrelated or heteroscedastic. The autoregressive error model corrects for autocorrelation. The generalized autoregressive conditional heteroscedasticity (GARCH) model and its variants model and correct for heteroscedasticity.

Fit Y by X with Autocorrelation and Heteroscedastic Errors produces test statistics that can be used to diagnose the autocorrelation and heteroscedasticity of errors. After one or both of these aspects are identified, you can fit a model with autoregressive errors or a model that can account for heteroscedasticity. **Fit Y by X with Autocorrelation and Heteroscedastic Errors** supports the following models:

- generalized ARCH (GARCH)

- integrated GARCH (IGARCH)
- exponential GARCH (EGARCH)
- GARCH-in-mean (GARCH-M)

Fit Y by X with Autocorrelation and Heteroscedastic Errors can also analyze models that combine autoregressive errors and GARCH-type heteroscedasticity by setting one or more variables as heteroscedastic.

The first part of this chapter describes how to a **Fit Y by X with Autocorrelation and Heteroscedastic Errors** analysis. The second and third parts provide a detailed description of the associated dialog and settings windows. The final part of this chapter provides a comparison of **Fit Y by X with Autocorrelation and Heteroscedastic Errors** in SAS Econometrics and Time Series Analysis for JMP to the JMP Time Series platform.

Specify Autocorrelated Regression Analysis

Step 1. Specify and Estimate a Base Model in the Initial Model Window

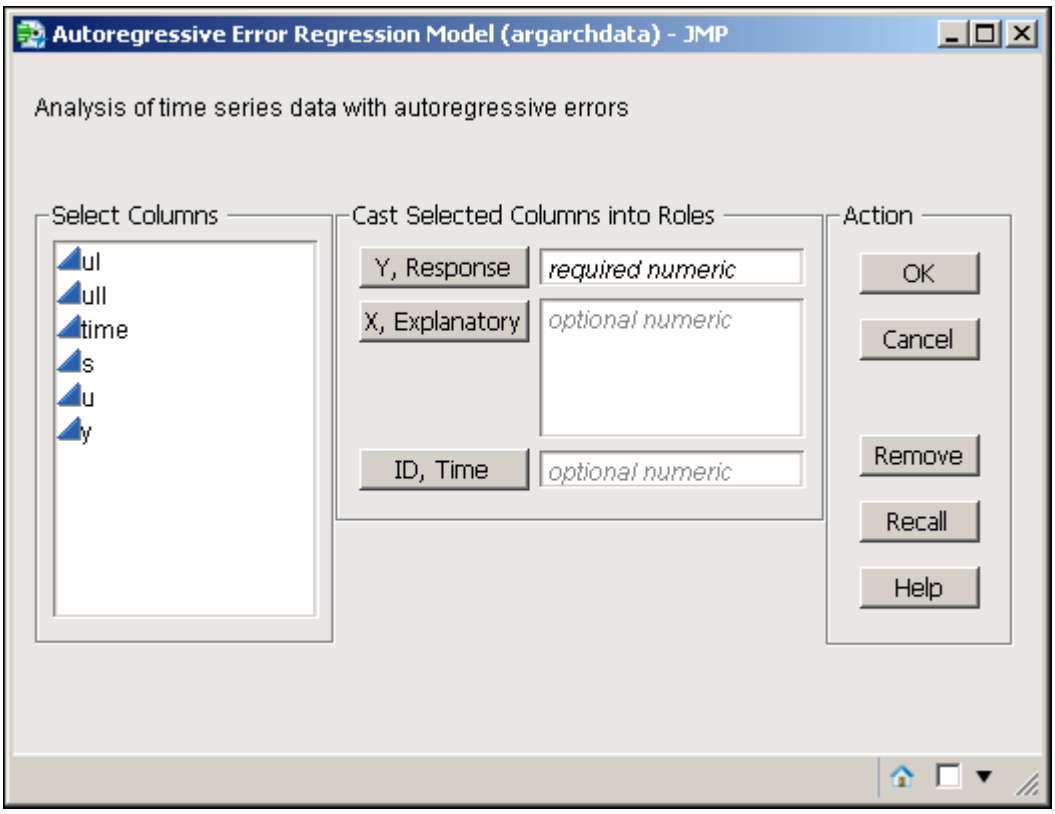
The base model for **Fit Y by X with Autocorrelation and Heteroscedastic Errors** has no autoregressive lags, no GARCH terms, and no heteroscedasticity variables.

To specify and estimate a base model:

1 Open a data file.
NOTE: This chapter contains example windows that use the *argarchdata.jmp* data file.

2 From the JMP Home window, select **Analyze ► Econometrics and Time Series ► Fit Y by X with Autocorrelation and Heteroscedastic Errors**. The Autoregressive Error Regression Model initial model window appears (see Figure 4.1.)

Figure 4.1 Example Autoregressive Error Regression Model Window



3 Specify the Y (response) variable by clicking the variable in the **Select columns** list and then clicking **Y, Response**.

4 (Optional) Specify any X (explanatory) variables by clicking the variables in the **Select columns** list and then clicking **X, Explanatory**. You can select more than one variable by holding CTRL and clicking each variable you want to select.

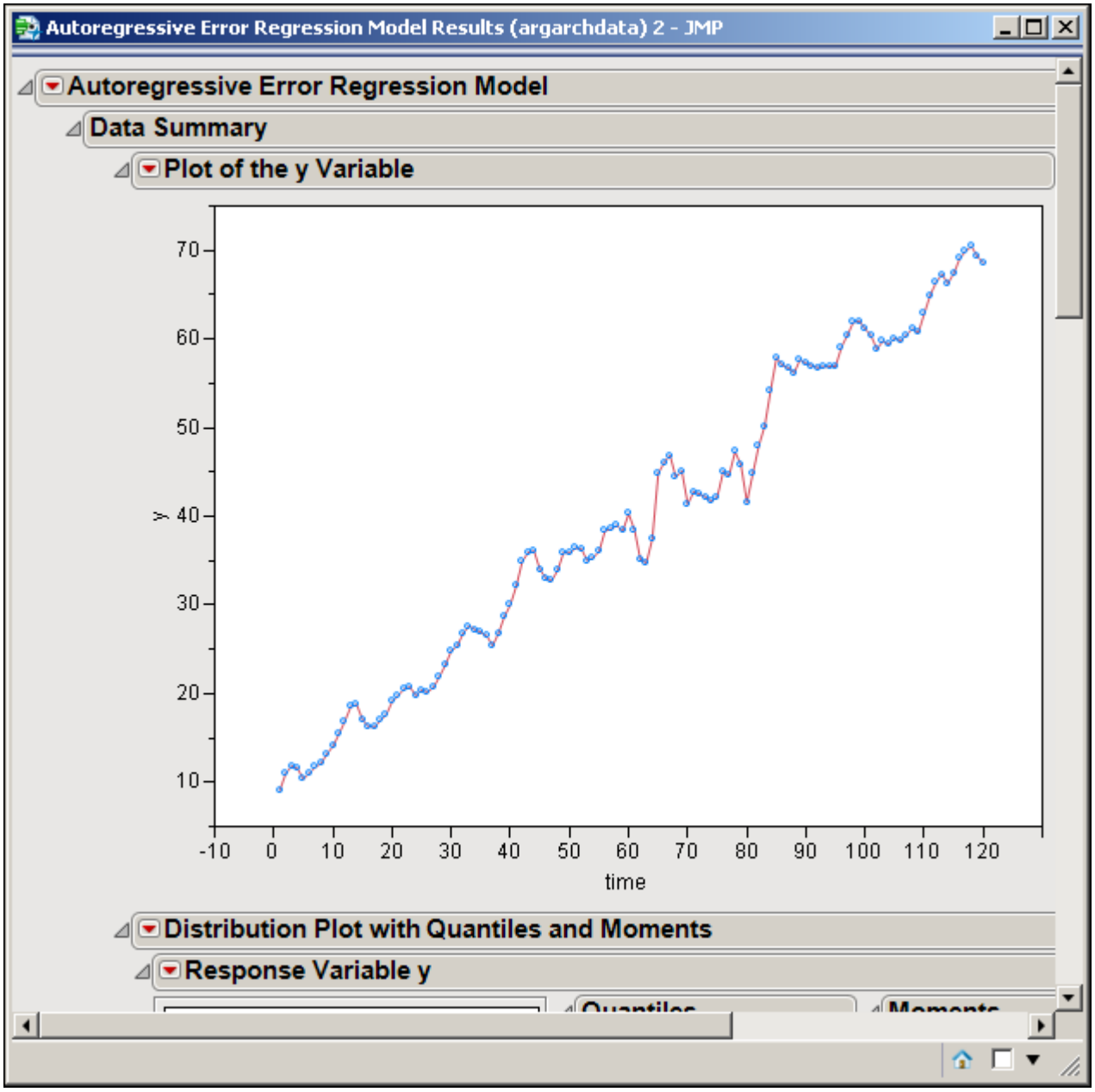
5 (Optional) Specify the ID (time) variable by clicking the variable in the **Select columns** list and then clicking **ID, Time**.

6 Click **OK** to estimate the model.

Step 2. Examine Base Models Results in the Results Window

After you create and estimate the base model, the Autoregressive Error Regression Model Results window appears. Figure 4.2 shows an example.

Figure 4.2 Example Autoregressive Error Regression Model Results Window



The Autoregressive Error Regression Model Results window contains the following areas:

- The **Data Summary** area contains the original data in X–Y plot and a distribution plot with tables of the quantiles and moments for the observed data.
- The **Fitted Models** area is similar for all of the analysis commands. For more information, see the section "Results Window" on page 5.
- The **Model Comparison** area is similar for all analyses. For more information, see the section "Model Comparison Area" on page 8.

- The **Model Estimates** area contains the following information:
  - autoregression model estimates and tests for AR parameters
  - autoregression model graphics
  - SAS ODS Graphics panel (any SAS ODS Graphics that are selected in the SAS ODS Graphics Settings window)

You examine the Autoregressive Error Regression Model Results window to determine how to modify the base model to create an alternative model that is a better representation of your data.

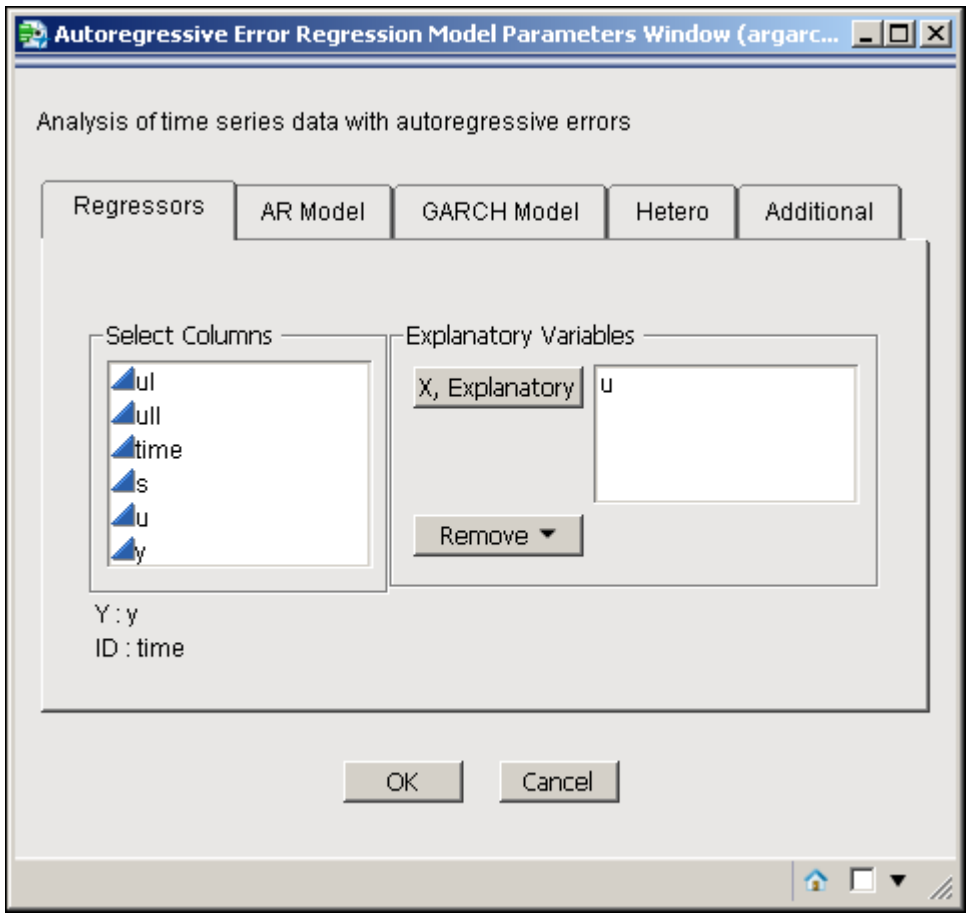
By clicking on the red triangles next to headings in the results window, you can modify the view of the plots and tables shown or you can add more plots or tables to the results window. You can also collapse or expand sections by clicking on the grey triangles next to the headings in the results window.

Step 3. Create and Estimate Alternative Model in the Dialog Window

1 In the Autoregressive Error Regression Model Results window, you can open a Autoregressive Error Regression Model Parameters dialog window (see Figure 4.3) to create a new model by doing one of the following:

- Click the red triangle next to the **Fitted Models** header, and then select **New Model**. The Autoregressive Error Regression Model Parameters window appears with the current model settings.
- Click the red triangle next to **Model 1** in the **Fitted Models** area, and select one of the following commands:
  - Select **Edit** to open the Autoregressive Error Regression Model Parameters window. This window contains the current model settings.
    NOTE: If you choose this option and perform the remaining steps, you can no longer compare the base model to any alternative models.
  - Select **Copy** to create a copy of the base model. Then click the red triangle next to **Copy of Model 1**, and select **Edit** to open the Autoregressive Error Regression Model Parameters window. This window contains the values of the copied model.

Figure 4.3 Regressors Tab of the Autoregressive Error Regression Model Parameters Window



2 In the Autoregressive Error Regression Model Parameters dialog window, specify an alternative model. For more information about this window and its tabs, see the following section.

3 Click **OK** to estimate the alternative model. The Autoregressive Error Regression Model Parameters window closes, and the model results for the alternative model appear in the Autoregressive Error Regression Model Results window. For more information, see the section "Results Window" on page 5.

Autoregressive Error Regression Model Parameters Dialog Window

The Autoregressive Error Regression Model Parameters window contains five tabs, which are described in the following sections. The options on each tab correspond to statements or options in the SAS/ETS AUTOREG procedure.

Regressors Tab

On the **Regressors** tab, you specify the X (explanatory) variables, which are often referred to as regressor variables. See Figure 4.3.

To add a regressor variable to the model:

1 Click the variable in the **Select Columns** list. You can select more than one variable at a time by holding CTRL and clicking each variable you want to include.

2 Click **X, Explanatory**. The variables that you selected appear in the **Explanatory Variables** list.

To remove a regressor variable from the model:

1 Click the variable in the **Explanatory Variables** list.

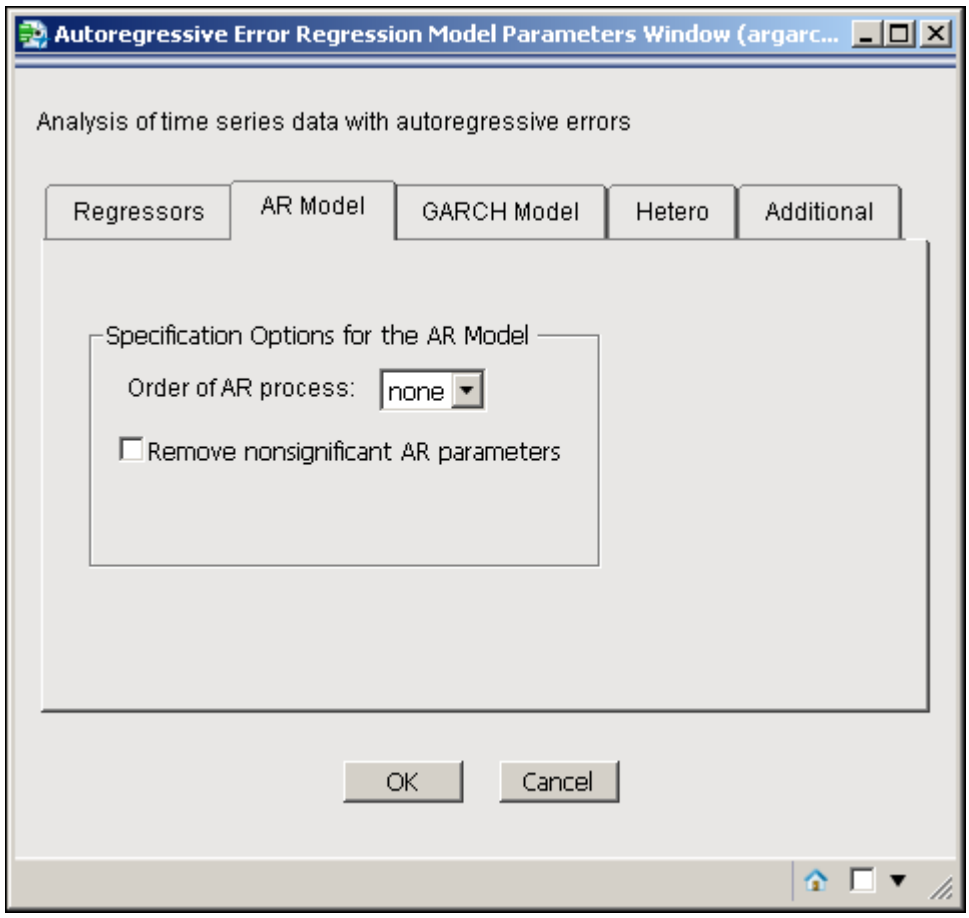
2 Click **Remove**, and click **Selected**. The variable you selected no longer appears in the **Explanatory Variables** list. You can remove all of the variables in the **Explanatory Variables** list by clicking **Remove** and then clicking **All**.

NOTE: Models with different X variables can be compared, either by using the fitted or residual graphs or by using the overall model fit information (for example, Akaike's information criterion or the Bayesian information criterion) for each model.

AR Model Tab

On the **AR Model** tab, you specify the details of the autocorrelated regression process. See Figure 4.4.

Figure 4.4 AR Tab of the Autoregressive Error Regression Model Parameters Window



The following options (which correspond to options in the MODEL statement of the AUTOREG procedure) are available on the **AR Model** tab:

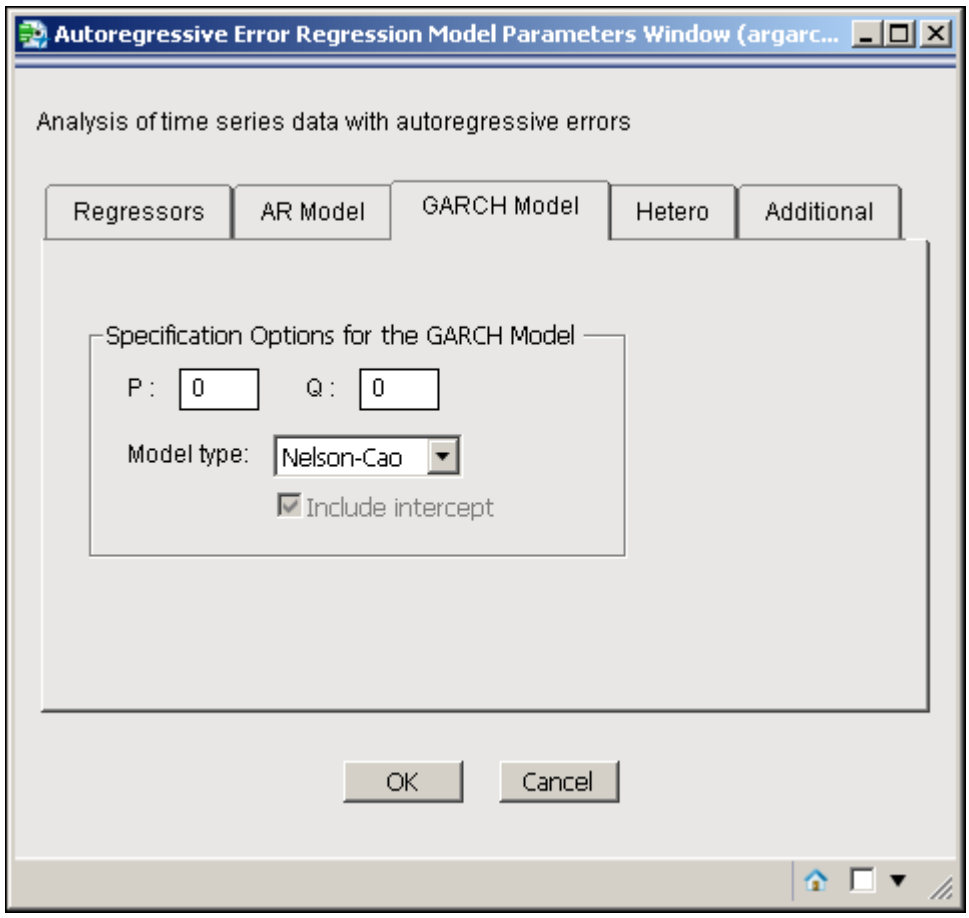
Order of AR process
specifies the order of the autoregressive error process. This option corresponds to the NLAG= option.

Remove nonsignificant AR parameters
specifies whether to remove nonsignificant autoregressive parameters. The least significant parameters are removed first. This option corresponds to the BACKSTEP option.

GARCH Model Tab

On the **GARCH Model** tab, you specify the details of the GARCH model. See Figure 4.5.

Figure 4.5 GARCH Model Tab of the Autoregressive Error Regression Model Parameters Window



The following options (which correspond to options in the MODEL statement of the AUTOREG procedure) are available on the **GARCH Model** tab:

P

specifies the order of the process or the subset of GARCH terms to be fitted. This option corresponds to the P= option.

Q

specifies the order of the process or the subset of ARCH terms to be fitted. This option corresponds to the Q= option.

Model type

specifies the type of GARCH model. This option corresponds to the TYPE= option. The default is Nelson-Cao. Select one of the following from the **Model type** list:

- **Exponential** specifies the exponential GARCH or EGARCH model.
- **Integrated** specifies the integrated GARCH or IGARCH model.
- **Nelson-Cao** specifies the Nelson-Cao inequality constraints.
- **Nonnegative** specifies the GARCH model with nonnegative constraints.

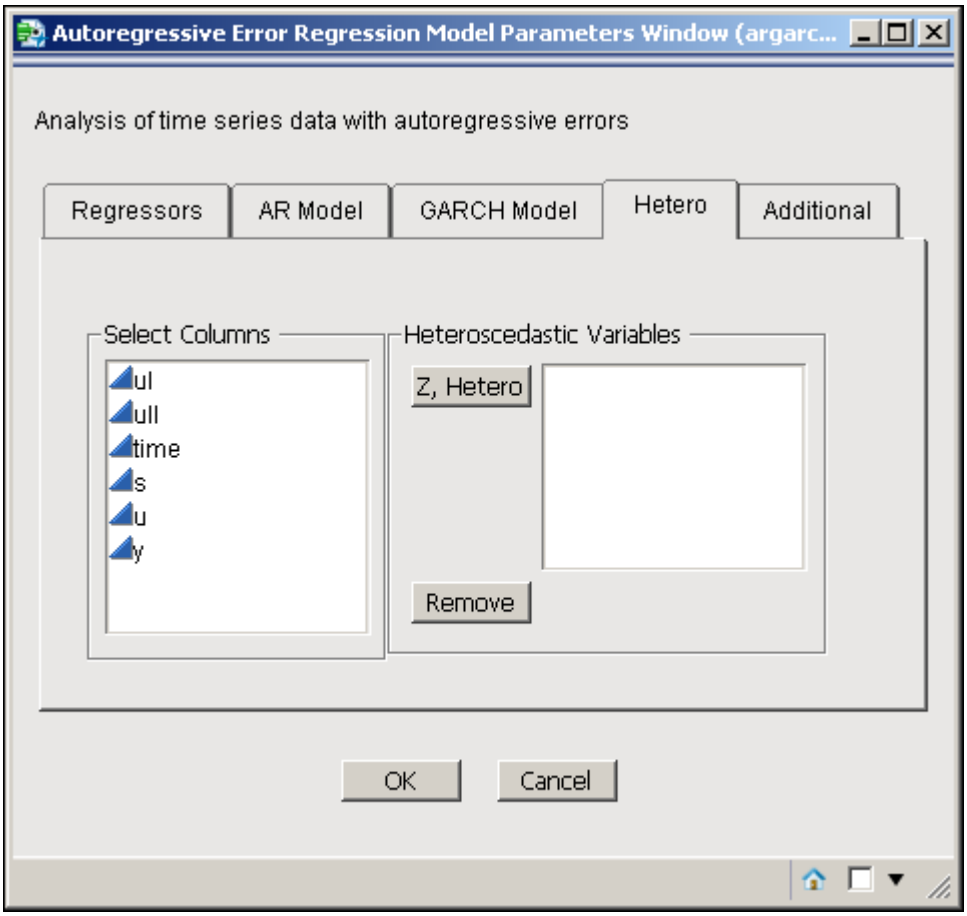
- **Stationary** constrains the sum of GARCH coefficients to be less than 1.

Include Intercept
specifies whether to suppress the intercept parameter. This option is available only when **Integrated** is selected from the **Model type** list. This option corresponds to the NOINT option.

Hetero Tab

On the **Hetero** tab, you specify the heteroscedastic variance model. See Figure 4.6. The variables in the current data tables are displayed in the **Select Columns** list. You can select any of variables that are not used as X (regressor) and Y (dependent) variables.

Figure 4.6 Hetero Tab of the Autoregressive Error Regression Model Parameters Window



To add a heteroscedastic variable to the model:

1 Select the variable in the **Select Columns** list. You can select more than one variable at a time by holding CTRL and clicking each variable you want to include.

2 Click **Z, Hetero**. The variables you selected appears in the **Heteroscedastic Variables** list.

To remove a heteroscedastic variable from the model:

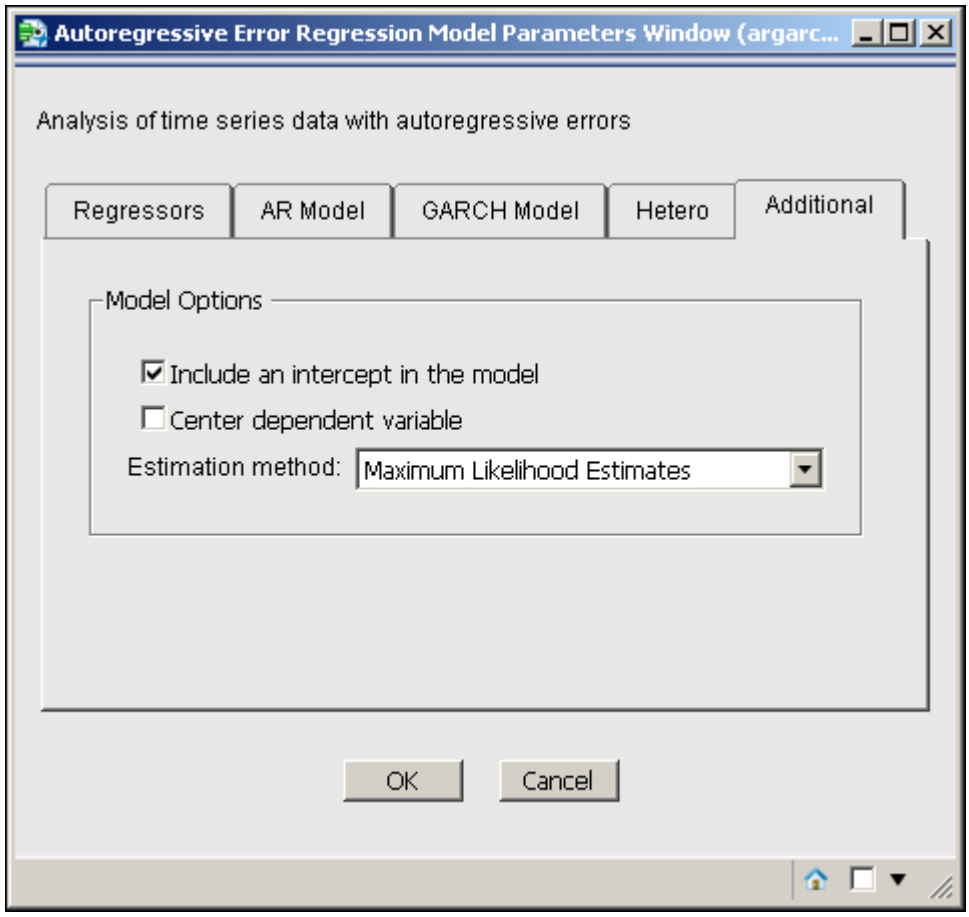
1 Select the variable in the **Heteroscedastic Variables** list.

2 Click **Remove**. The variable you selected no longer appears in the **Heteroscedastic Variables** list.

Additional Tab

On the **Additional** tab, you specify the additional model settings. See Figure 4.7.

Figure 4.7 Additional Tab of the Autoregressive Error Regression Model Parameters Window



The following options (which correspond to options in the MODEL statement of the AUTOREG procedure) are available on the **Additional** tab:

Include an intercept in the model
specifies whether to suppress the intercept parameter. This option corresponds to the NOINT option.

Center dependent variable
specifies whether to center the dependent variable for a model that does not have any explanatory

variables. This option corresponds to the CENTER option. This option is valid only when the model does not have regressors (explanatory variables).

Estimation Method
specifies the estimation method. This option corresponds to the METHOD= option. Select one of the following from the **Model type** list:

- **Maximum Likelihood Estimates** specifies maximum-likelihood estimation.
- **Unconditional Least Squares Estimates** specifies unconditional least squares estimation.
- **Yule-Walker Estimates** specifies Yule-Walker estimation.
- **Iterative Yule-Walker Estimates** specifies iterative Yule-Walker estimation.

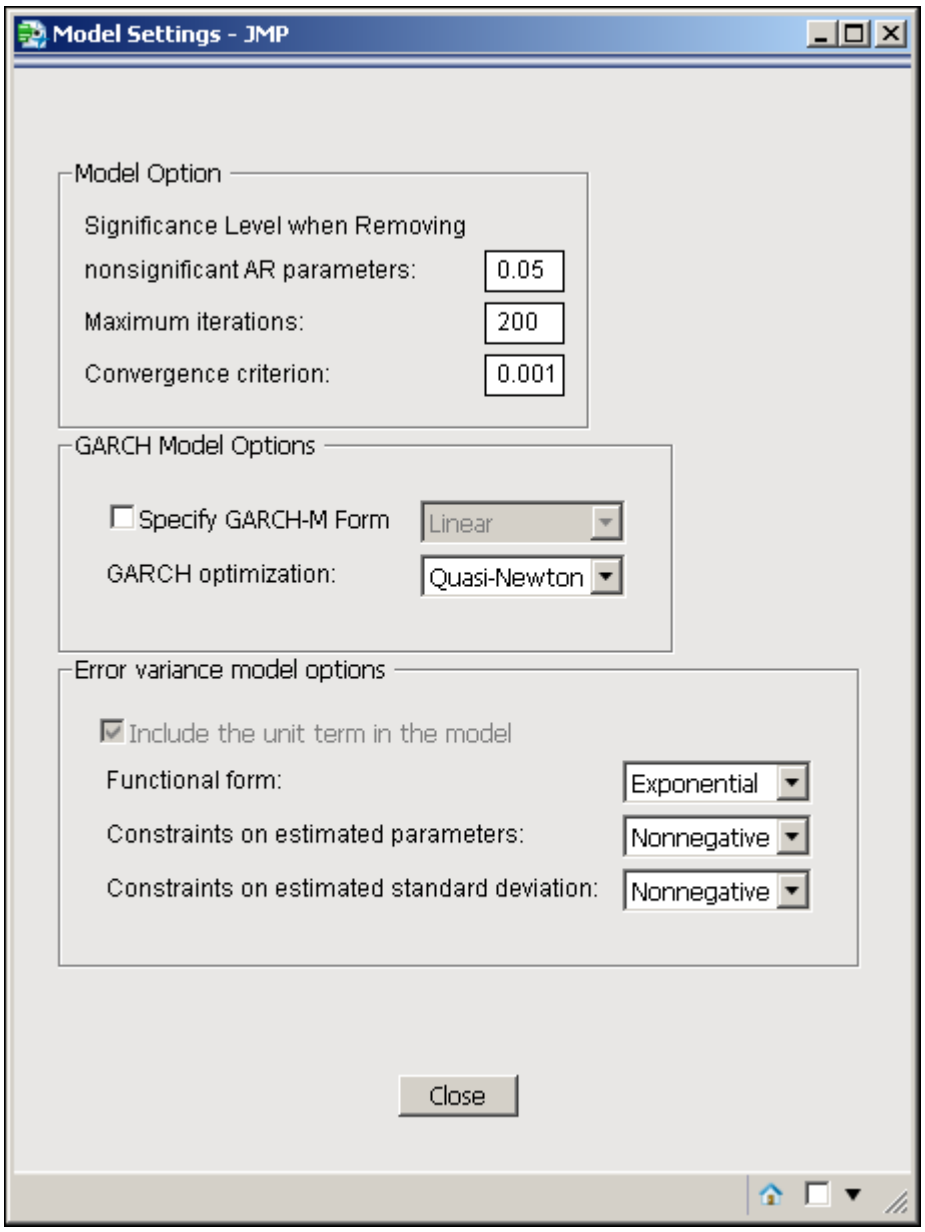
Settings Windows

The **Fit Y by X with Autocorrelation and Heteroscedastic Errors** analysis provides three settings windows. The SAS Graphs Settings window and Display Settings window are described in the section "Settings Windows" on page 10. The Model Settings window is described in the following section.

Model Settings Window

The Model Settings window, shown in Figure 4.8, contains the options for model estimation and model constraints.

Figure 4.8 Model Settings Window



Model Option Area

The following options (which correspond to options in the MODEL statement of the AUTOREG procedure) are available in the **Model Option** area:

Significance level when removing nonsignificant AR parameters
specifies the significance level criterion for removing nonsignificant autoregressive parameters. This option corresponds to the SLSTAY= option.

Maximum iterations
specifies the maximum number of iterations. The default value is 50. This option corresponds to the MAXITER= option.

Convergence Criterion
specifies the convergence criterion. This option corresponds to the CONVERGE= option.

GARCH Model Options Area

The following options (which correspond to options in the MODEL statement of the AUTOREG procedure) are available in the **GARCH Model Options** area:

Specify GARCH-M
specifies the functional form of the GARCH-M model. This option corresponds to the GARCH=(MEAN=) option.

GARCH optimization
specifies the type of GARCH model optimization. You can select the following values from the **GARCH optimization** list:

- **Quasi-Newton** specifies quasi-Newton optimization. This option is the default option.
- **Trust Region** specifies trust region optimization. This option corresponds to the GARCH=(TR) option.

Error Variance Model Options Area

The following options (which correspond to options in the HETERO statement of the AUTOREG procedure) are available in the **Error Variance Model Options** area:

Include the unit term in the model
specifies that the heteroscedasticity model not include the unit term for the functional form option. This option corresponds to the NOCONST option.

Functional form
specifies the functional form of the heteroscedasticity model. The functional form options are **Exponential**, **Square**, and **Linear**. By default, **Exponential** is selected. This option corresponds to the LINK= option.

Constraints on estimated parameters
specifies the constraints on the estimated parameters of the heteroscedastic model. The following options correspond to the values of the COEF= option:

- **Nonnegative** specifies that the heteroscedastic parameters must be nonnegative during model estimation. This option corresponds to COEF=NONNEG.
- **Unit** specifies that the heteroscdastic parameters be fixed to 1. This option corresponds to COEF=UNIT.

- **Zero** specifies that the heteroscdastic parameters be fixed to 0. This option corresponds to COEF=ZERO.
- **Unrestricted** specifies that the heteroscedastic parameters not be restricted during model estimation. This option corresponds to COEF=UNREST.

By default, **Nonnegative** is selected when a GARCH model has been specified and **Unrestricted** is selected when no GARCH terms have been specified.

Constraints on estimated standard deviation

specifies the constraints on the estimated standard deviation of the heteroscedastic model. The following options correspond to the values of the STD= option:

- **Nonnegative** specifies that the estimated standard deviation of the heteroscedastic model be nonnegative during model estimation. This option corresponds to STD=NONNEG.
- **Unit** specifies that the estimated standard deviation of the heteroscedastic model be fixed to 1. This option corresponds to STD=UNIT.
- **Unrestricted** specifies that the estimated standard deviation of the heteroscedastic model not be restricted during model estimation. This option corresponds to STD=UNREST.

By default, **Nonnegative** is selected.

Comparison of Time Series Platform and Autocorrelated Regression Analysis

The JMP Time Series platform also analyzes autocorrelated data. However, there are several differences between the Time Series platform and the **Fit Y by X with Autocorrelation and Heteroscedastic Errors** analysis in the SAS Econometrics and Time Series Analysis for JMP.

The JMP Time Series platform does not call a SAS procedure. It uses an ARIMA (Box-Jenkins) model to model and forecast time series. In contrast, **Fit Y by X with Autocorrelation and Heteroscedastic Errors** calls the SAS/ETS AUTOREG procedure, which performs regression analysis on variables with autocorrelated errors and supports ARCH and GARCH models to deal with heteroscedasticity.

Table 4.1 briefly compares the JMP Time Series platform with the **Fit Y by X with Autocorrelation and Heteroscedastic Errors** analysis.

Table 4.1 Comparison of the JMP Time Series Platform and Fit Y by X with Autocorrelation and Heteroscedastic Errors

Category	JMP Time Series Platform	Fit Y by X with Autocorrelation and Heteroscedastic Errors
Model	Box-Jenkins approach to estimating univariate models. Uses transfer functions when explanatory variables are present.	Linear regression models, corrected for autocorrelated errors. ARCH/GARCH models, corrected for heteroscedasticity.
Unit root tests	Augmented Dickey-Fuller test.	Phillips-Perron unit root test when there are no explanatory variables; the Phillips-Ouliaris cointegration test when there are explanatory variables.
Order of lag	User selects the order of the transfer function.	Stepwise regression, which can automatically select the order of the autoregressive model by removing insignificant autoregressive parameters.
Estimation	Maximum likelihood estimation.	Four different estimation methods.
Residual plots	Plot of AR residuals.	Plots of AR and structural (OLS) residuals.
Modeling options	Time-series modeling options such as prewhitening, spectral density, variograms, and smoothing models.	Does not provide these models.
Profiler analysis	No profiler plots of predicted values.	Profiler plots for structural and AR models.

Chapter 5

Fit Panel Data Regression

Contents

Overview of Fit Panel Data Regression

You select **Fit Panel Data Regression** to perform panel data analysis. Panel data analysis is a class of linear econometric models that commonly arise when time series and cross-sectional data are combined. Typical examples of panel data include observations over time in households, countries, firms, trade, and so on. For example, survey data on household income is panel data when the same households are repeatedly surveyed at multiple times (for example, over several years).

Panel data analysis models can be grouped into several categories depending on the structure of the error term. Some examples of panel data analysis models that can be fit with **Fit Panel Data Regression** are:

- one-way and two-way models
- fixed-effects and random-effects models
- autoregressive models
- moving average models

The first part of this chapter describes how to perform a **Fit Panel Data Regression** analysis. The second part provides a detailed description of the dialog window associated with **Fit Panel Data Regression**. There is no unique Settings window for **Fit Panel Data Regression**; see the section "Settings Windows" on page 10 for all the relevant information about settings windows for this analysis.

Specify a Panel Analysis Model

Step 1. Specify and Estimate a Base Model in the Initial Model Window

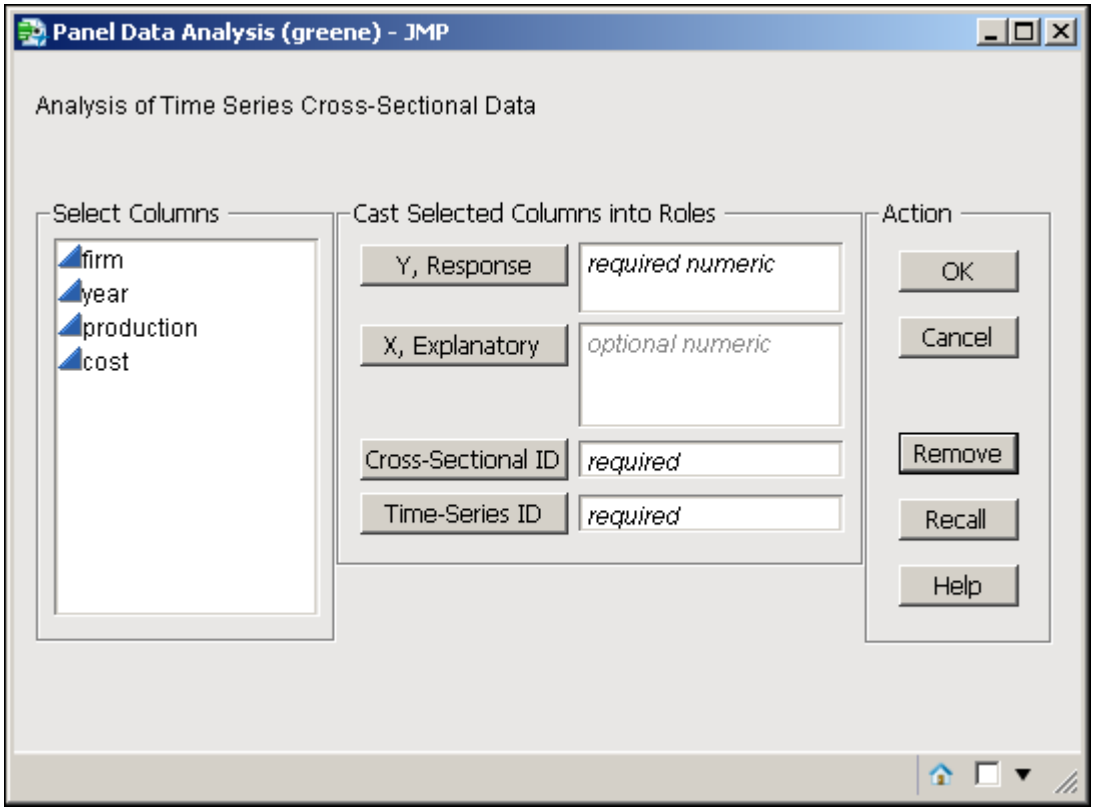
The base model for **Fit Panel Data Regression** is a model with an intercept and two-way fixed effects.

To specify and estimate a base model:

1 Open a data file. This chapter contains example windows that use the *greene.jmp* data file.

2 From the JMP Home window, select **Analyze ► Econometrics and Time Series ► Fit Panel Data Regression**. The Panel Data Analysis initial model window appears. See Figure 5.1.

Figure 5.1 Example Initial Model Window for Fit Panel Data Regression



3 Specify the Y (response) variable by selecting the variable in the **Select columns** list and then clicking **Y, Response**.

4 (Optional) Specify any X (explanatory) variables by selecting the variables in the **Select columns** list and then clicking **X, Explanatory**. You can select more than one variable by holding CTRL and selecting each variable you want to select.

5 Specify a cross-sectional ID variable by selecting the variable in the **Select columns** list and then clicking **Cross-Sectional ID**.

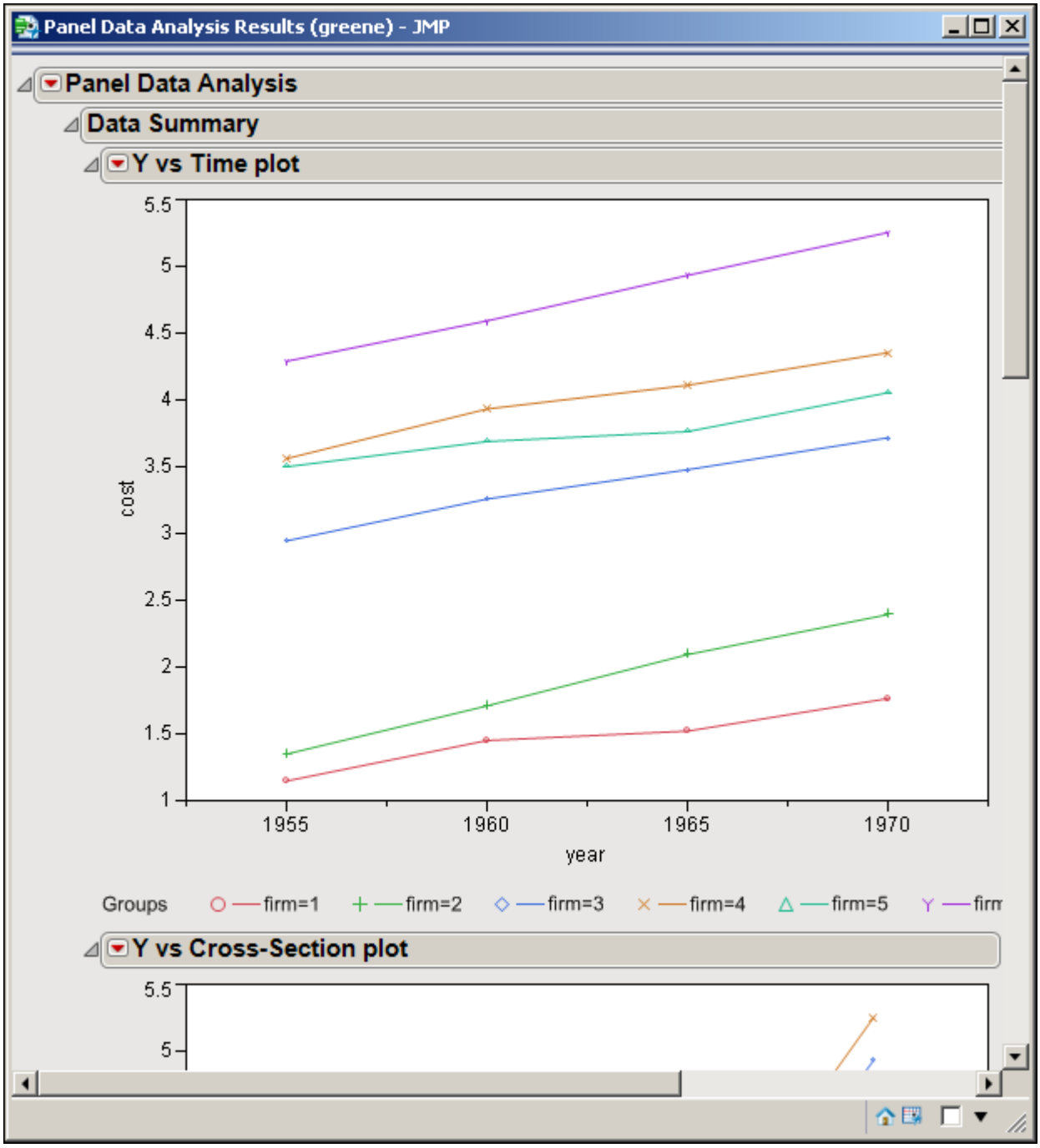
6 Specify a time-series ID variable by selecting the variable in the **Select columns** list and then clicking **Time-Series ID**.

7 Click **OK** to estimate the model.

Step 2. Examine Base Model Results in the Results Window

After you create and estimate the base model, the Panel Data Analysis Results window appears. See Figure 5.2.

Figure 5.2 Example Panel Data Analysis Results Window

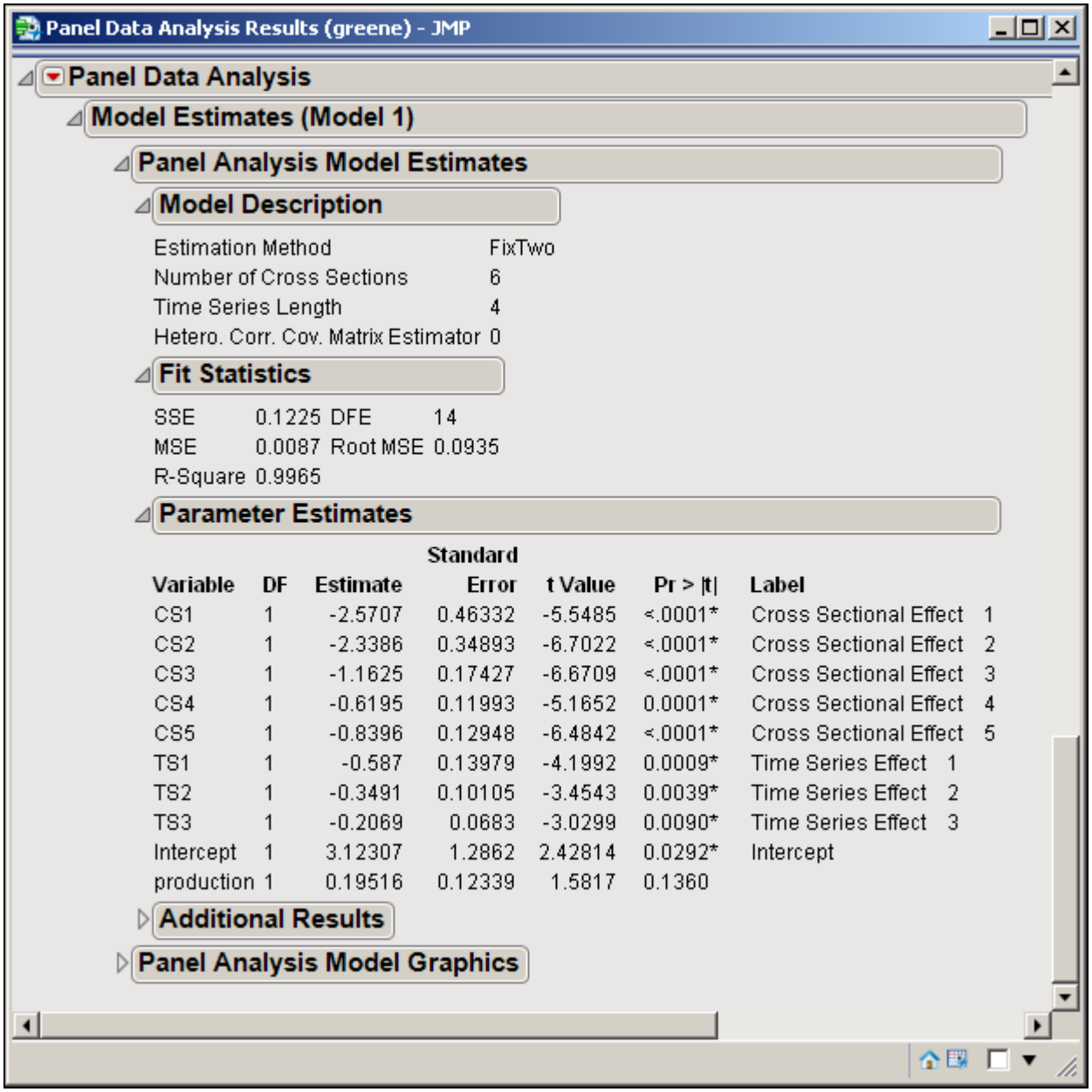


The Panel Data Analysis Results window contains the following areas:

- The **Data Summary** area contains the original data in a Y–Time Plot and a Y–Cross-Section Plot. It also contains tables of the quantiles and moments for the observed data.
- The **Fitted Models** area is similar for all analyses. For more information, see the section "Fitted Models Area" on page 6.
- The **Model Comparison** area is similar for all analyses. For more information, see the section "Model Comparison Area" on page 8.
- The **Model Estimates** area contains the following information:
  - model description
  - fit statistics
  - parameter estimates
  - additional results
  - panel analysis model graphics

See Figure 5.3 for an example of the **Model Estimates** area.

Figure 5.3 Example Model Estimates Area in a Panel Data Analysis Results Window



You examine the Panel Data Analysis Results window to determine how to modify the base model to create an alternative model that is a better representation of your data.

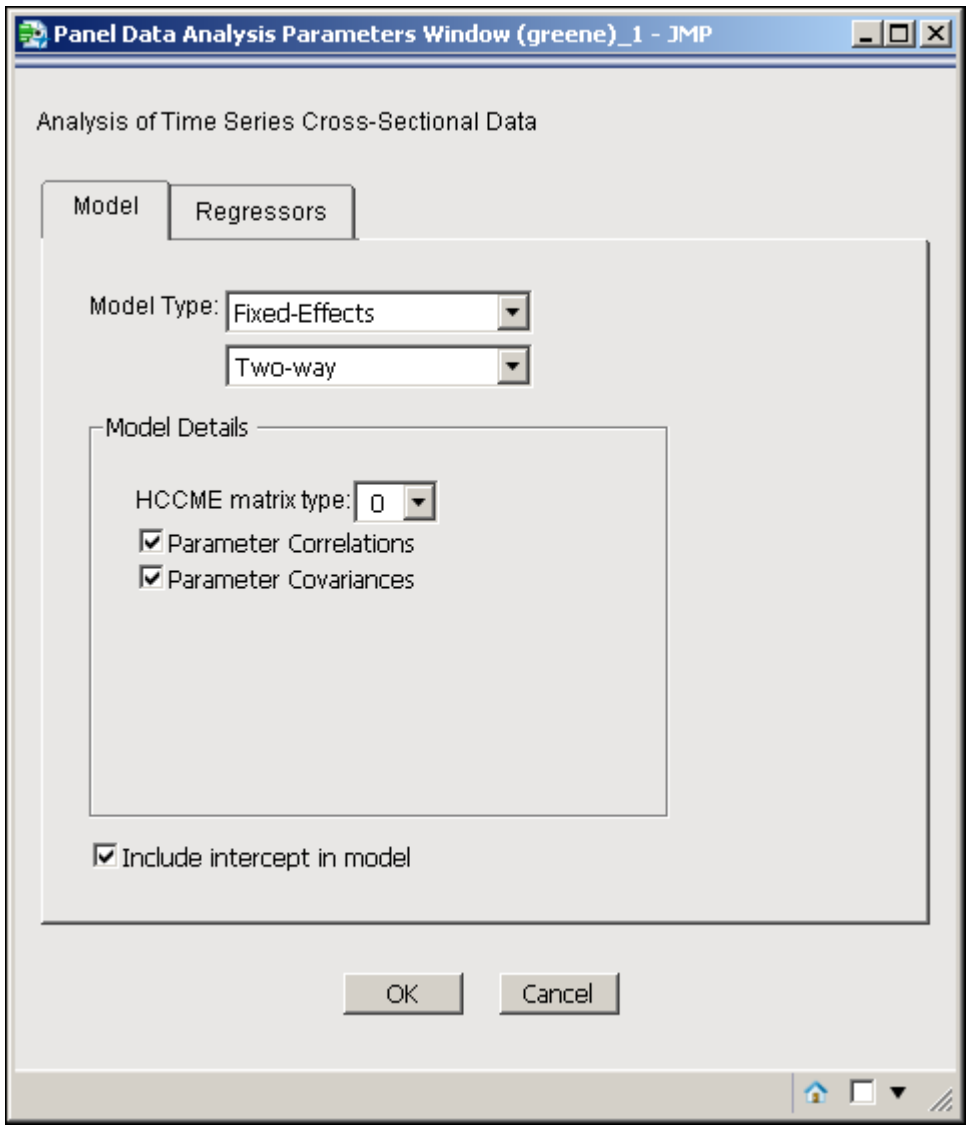
You can modify the plots and tables shown by clicking on the red triangles next to headings in the results window. You can also collapse or expand sections by clicking on the gray triangles next to the headings in the results window.

Step 3. Create and Estimate an Alternative Model in the Dialog Window

1 In the Panel Data Analysis Results window, you can open a Panel Data Analysis Parameters dialog window (see Figure 5.4) to create a new model by doing one of the following:

- Click the red triangle next to the **Fitted Models** header, and then select **New Model**. The Panel Data Analysis Parameters dialog window appears with the current model settings.
- Click the red triangle next to **Model 1** in the **Fitted Models** area, and select one of the following commands:
  - Select **Edit** to open the Panel Data Analysis Parameters dialog window. This window contains the current model settings.
    NOTE: If you choose this option and perform the remaining steps, you can no longer compare the base model to any alternative models.
  - Select **Copy** to create a copy of the base model. Then click the red triangle next to **Copy of Model 1**, and select **Edit** to open the Panel Data Analysis Parameters dialog window. This window contains the values for the copied model.

Figure 5.4 Example Panel Data Analysis Parameters Dialog Window



2 In the Panel Data Analysis Parameters dialog window, specify an alternative model. For more information about this window and its tabs, see the following section.

3 Click **OK** to estimate the alternative model. The Panel Data Analysis Parameters window closes, and the model results for the alternative model appear in the Panel Data Analysis Results window. For more information, see the section "Results Window" on page 5.

Panel Data Regression Parameters Dialog Window

The Panel Data Analysis Parameters dialog window contains two tabs, which are described in the following sections. The options on these tabs correspond to statements or options for the SAS/ETS PANEL procedure.

Model Tab

On the **Model** tab, you specify the model type and options. An example of the **Model** tab is shown in Figure 5.4.

Select one of the following options from the **Model Type** list (these options correspond to the options in the MODEL statement of the PANEL procedure):

Fixed Effects
specifies that a fixed effects model be estimated. Fixed effects can be one-way cross-sectional, one-way time, or two-way. These options correspond to the FIXONE, FIXONETIME, and FIXTWO options, respectively.

Random-Effects
specifies that a random effects model be estimated. Random effects can be one-way or two-way. These options correspond to the RANONE and RANTWO options, respectively.

Average
specifies that a between-groups or between-time-periods model be estimated. This option corresponds to the BTWNG and BTWNT options, respectively.

Da Silva
specifies that the model be estimated using the Da Silva method, which assumes a mixed variance-component moving average model for the error structure. This option corresponds to the DASILVA option.

Parks
specifies that the model be estimated using the Parks method, which assumes a first-order autoregressive model for the error structure. This option corresponds to the PARKS option.

Pooled
specifies that a pooled (OLS) model be estimated. This option corresponds to the POOLED option.

The options in the **Model Details** area depend on which model type you select. These options correspond to the options in the MODEL statement of the PANEL procedure. The following options can appear in the **Model Details** area:

HCCME Matrix type
specifies the type of HCCME variance-covariance matrix requested. The expected range of values is NO or 0 to 4. The default is NO, which specifies no correction. This option corresponds to the HCCME= option.

Parameter Correlation
specifies whether to print the matrix of estimated correlations between the parameter estimates. The default is to print the matrix. This option corresponds to the CORRB option.

Parameter Covariance
specifies whether to print the matrix of estimated covariance between the parameter estimates. The default is to print the matrix. This option corresponds to the COVB option.

Variance Estimation
specifies the type of variance component estimate to use. This option corresponds to the VCOMP= option. The default is Fuller-Battese (VCOMP=FB) for balanced data and Wansbeek-Kapteyn (VCOMP=WK) for unbalanced data.

One-way Breusch-Pagan Test
specifies whether to perform the Breusch-Pagan one-way test for random effects. This option corresponds to the BP option.

Two-way Breusch-Pagan Test
specifies whether to perform the Breusch-Pagan two-way test for random effects. This option corresponds to the BP2 option.

Order of Moving Average
specifies the order of the moving average process in the Da Silva method. This option corresponds to the M= option. The default is 1.

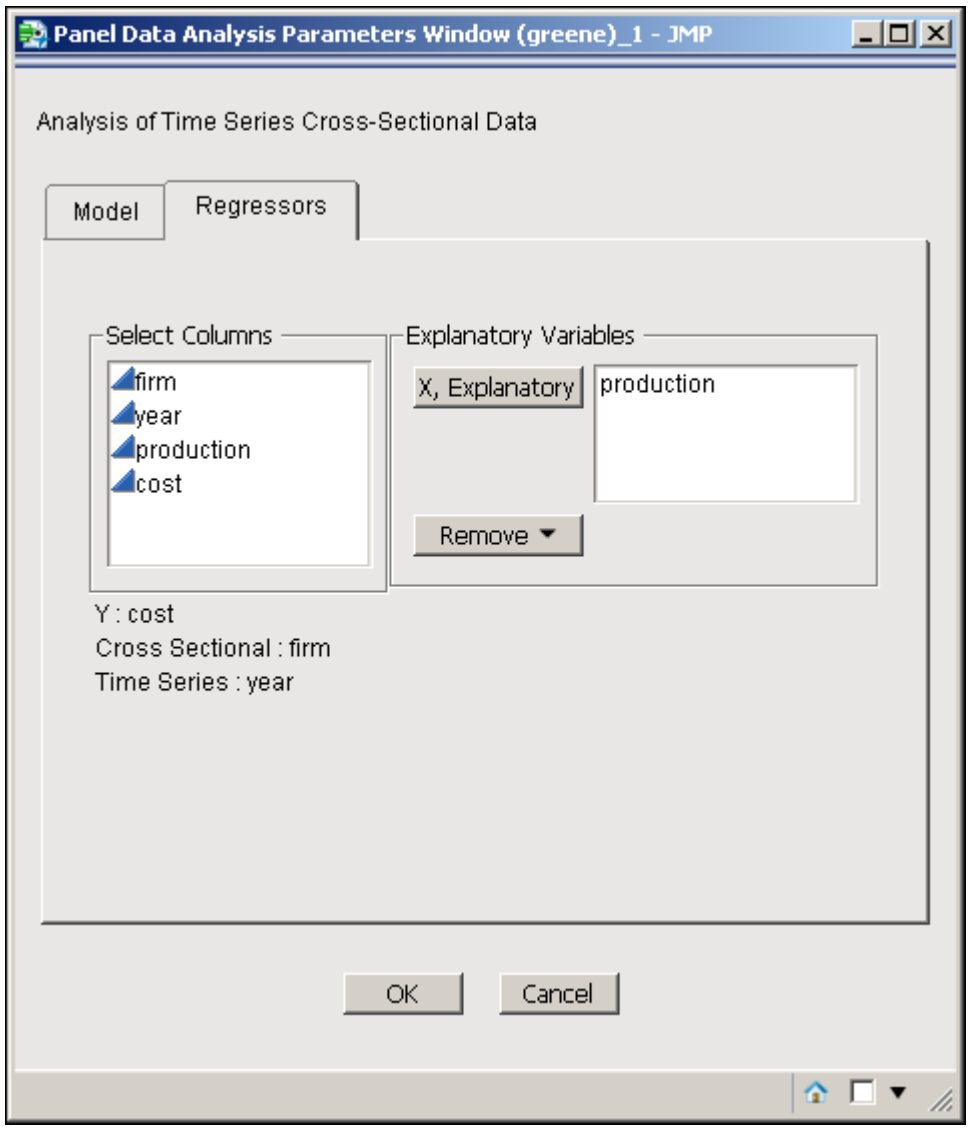
PHI matrix of covariance estimates
specifies whether to print the Φ matrix of estimated covariances of the observations for the Parks method. The PHI option is relevant only when the PARKS option is used. This option corresponds to the PHI option.

Autocorrelation coefficient estimates
specifies whether to print the estimated autocorrelation coefficients for the Parks method. This option corresponds to the RHO option.

Regressors Tab

On the **Regressors** tab, you specify the X (explanatory) variables, which are often referred to as regressor variables. An example of the **Regressors** tab is shown in Figure 5.5.

Figure 5.5 Panel Data Analysis Parameters Window with the Regressors Tab



To add a regressor variable to the model:

1 Select the variable in the **Select Columns** list. You can select more than one variable at a time by holding CTRL and clicking each variable you want to select.

2 Click **X, Explanatory**. The variables you selected appear in the **Explanatory Variables** list.

To remove a regressor variable from the model:

1 Select the variable in the **Explanatory Variables** list.

2 Click **Remove**, and click **Selected**. The variable you selected no longer appears in the **Explanatory Variables** list. You can remove all of the variables in the **Explanatory Variables** list by clicking **Remove** and clicking **All**.

NOTE: Models with different X variables can be compared, either by using the fitted or residual graphs or by using the overall model fit information (for example, AIC or BIC) for each model.

Chapter 6

Fit Unobserved Components Models

Contents

Overview of Fit Unobserved Components Models

You select **Fit Unobserved Components Models** to fit unobserved components models. Unobserved components models (UCMs) are used to forecast equally spaced univariate time series data. UCMs are also called structural models in the time series literature. A UCM decomposes the response series into components such as trend, seasonals, cycles, and the regression effects due to predictor series. The components in the model attempt to capture the salient features of the series that are useful in explaining and predicting its behavior.

Fit Unobserved Components Models can be used to fit unobserved components models with one or more of the following components:

- autoregressive
- seasonal
- cycle
- irregular
- level
- regression

- slope

The first part of this chapter describes how to perform a **Fit Unobserved Components Models** analysis. The second and third parts provide a detailed description of the dialog and settings windows associated with **Fit Unobserved Components Models**.

Specify an Unobserved Component Model

Step 1. Specify and Estimate a Base Model

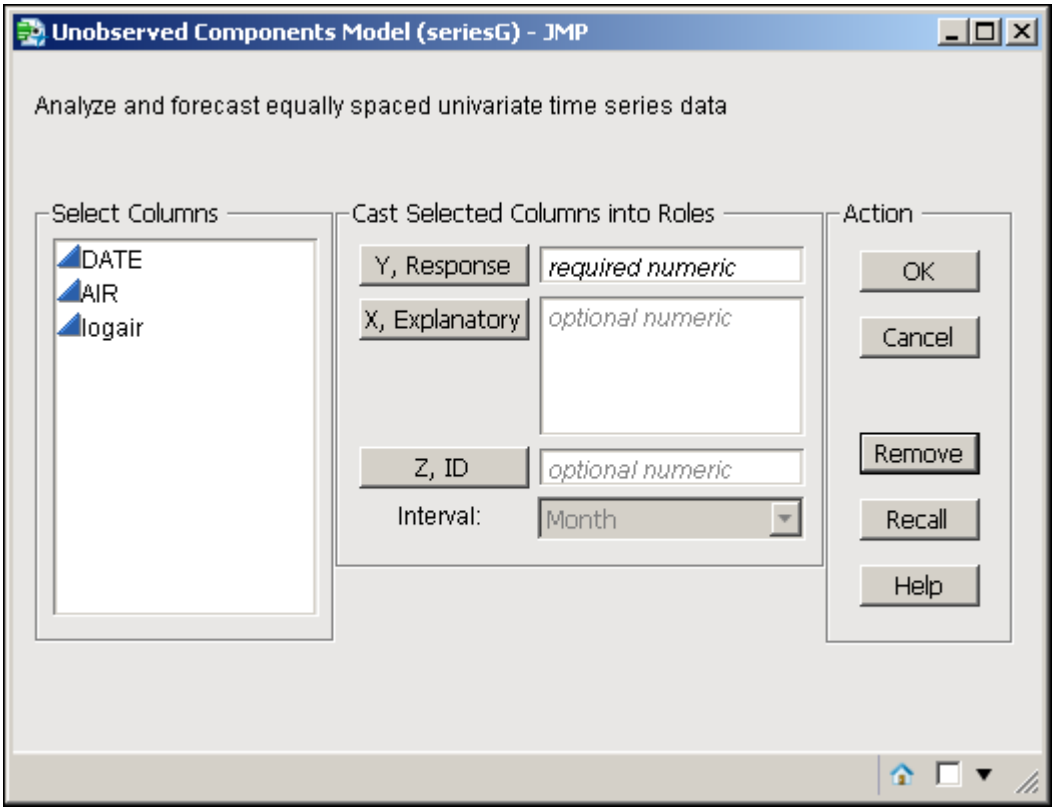
The base model for **Fit Unobserved Components Models** includes level, seasonal, and slope components.

To specify and estimate a base model:

1 Open a data file. This chapter contains example windows that use the *seriesG.jmp* data file.

2 From the JMP Home window, select **Analyze ► Econometrics and Time Series ► Fit Unobserved Components Models**. The Unobserved Components Model initial model window appears. See Figure 6.1.

Figure 6.1 Example Initial Model Window for Fit Unobserved Component Model



3 Specify the Y (response) variable by selecting the variable in the **Select columns** list and then clicking **Y, Response**.

4 (Optional) Specify any X (explanatory) variables by selecting the variables in the **Select columns** list and then clicking **X, Explanatory**. You can select more than one variable by holding CTRL and clicking each variable you want to select.

5 (Optional) Specify the Z (ID) variable by selecting the variable in the **Select columns** list and then clicking **Z, ID**.

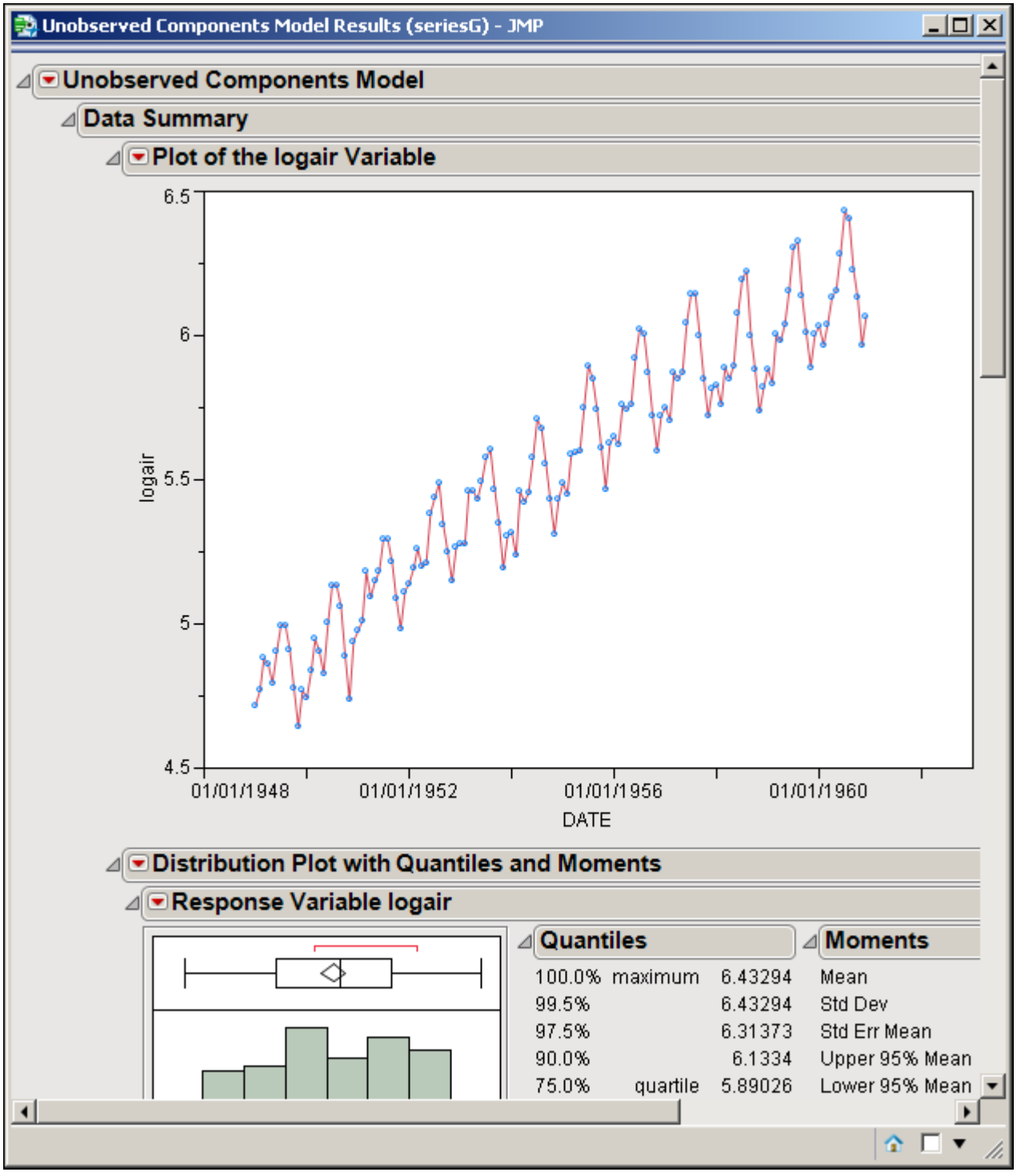
6 (Optional) Specify an interval for the observations by selecting a value from the **Interval** list.

7 Click **OK** to estimate the model.

Step 2. Examine Base Model Results in the Results Window.

After you create the base model, the Unobserved Components Model Results window appears. See Figure 6.2.

Figure 6.2 Example Unobserved Components Models Results Window



The Unobserved Components Model Results window contains the following areas:

- The **Data Summary** area contains the original data in an X–Y plot and a distribution plot with tables of the quantiles and moments for the observed data.
- The **Fitted Models** area is similar for all analyses. For more information, see the section "Fitted Models Area" on page 6.
- The **Model Comparison** area is similar for all analyses. For more information, see the section "Model Comparison Area" on page 8.
- The **Model Estimates** area of the results window contains the following information:
  - parameter estimates
  - estimation span
  - goodness-of-fit statistics
  - residual distribution information
  - outlier summary
  - additional results
  - UCM model graphics
  - UCM model forecast
  - UCM model forecast data table

See Figure 6.3 and Figure 6.4 for examples of the **Model Estimates** area of the Unobserved Components Model Results window.

Figure 6.3 Example Model Estimates Area in a Unobserved Components Model Results Window

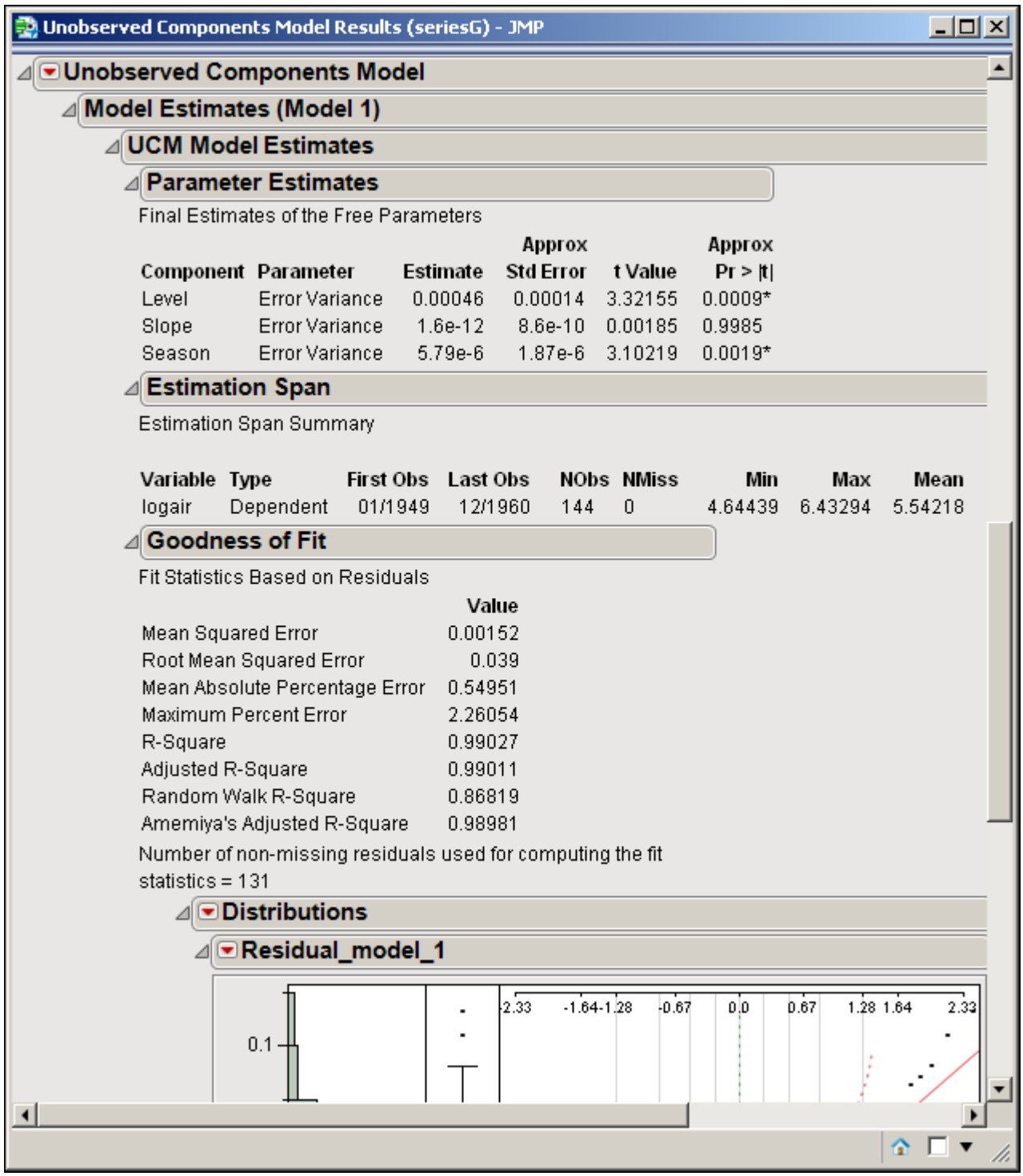
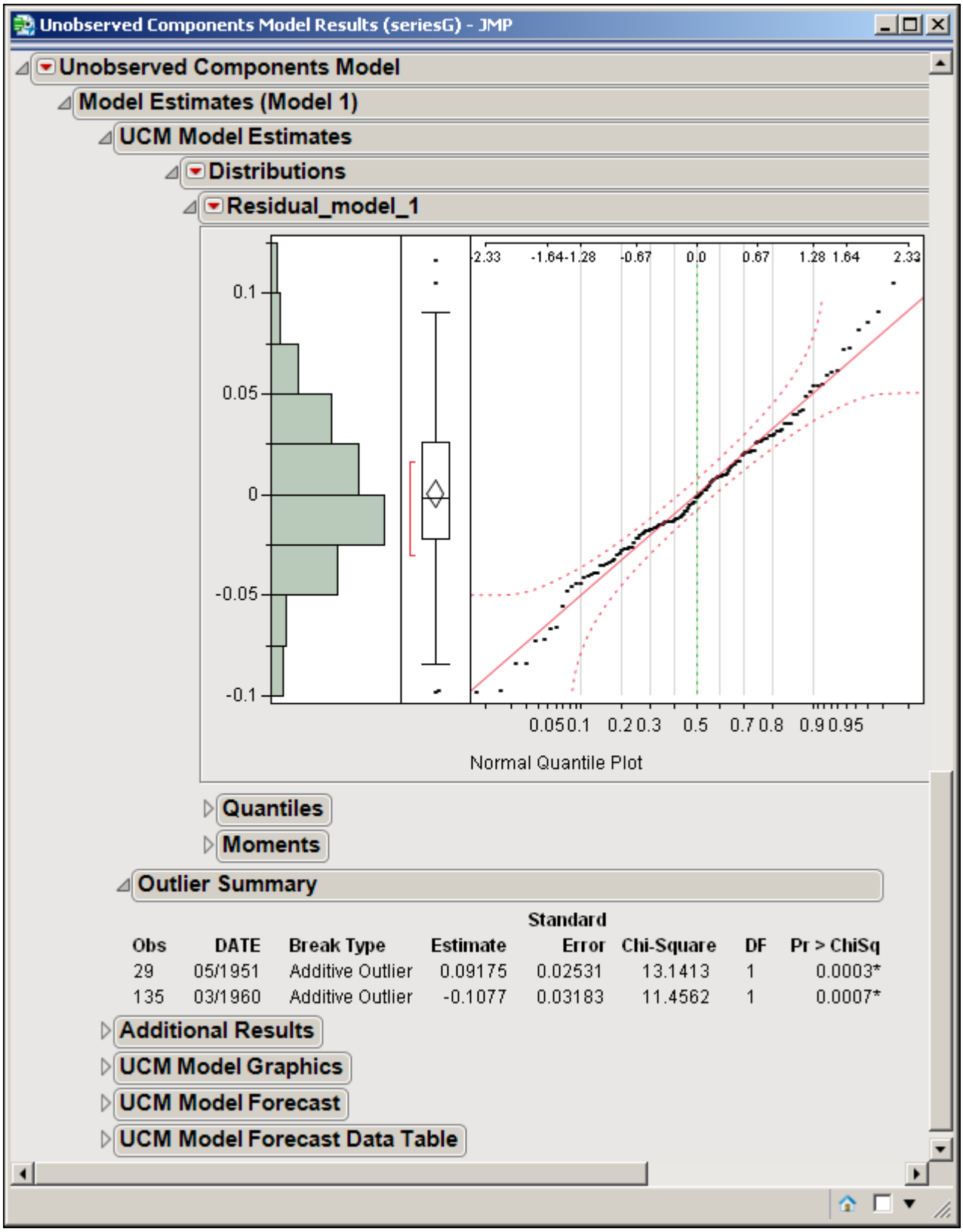


Figure 6.4 Example Model Estimates Area in a Unobserved Components Model Results Window Continued



Obs	DATE	Break Type	Estimate	Standard Error	Chi-Square	DF	Pr > ChiSq
29	05/1951	Additive Outlier	0.09175	0.02531	13.1413	1	0.0003*
135	03/1960	Additive Outlier	-0.1077	0.03183	11.4562	1	0.0007*

You examine the Unobserved Components Model Results window to determine how to modify the base model to create an alternative model that is a better representation of your data.

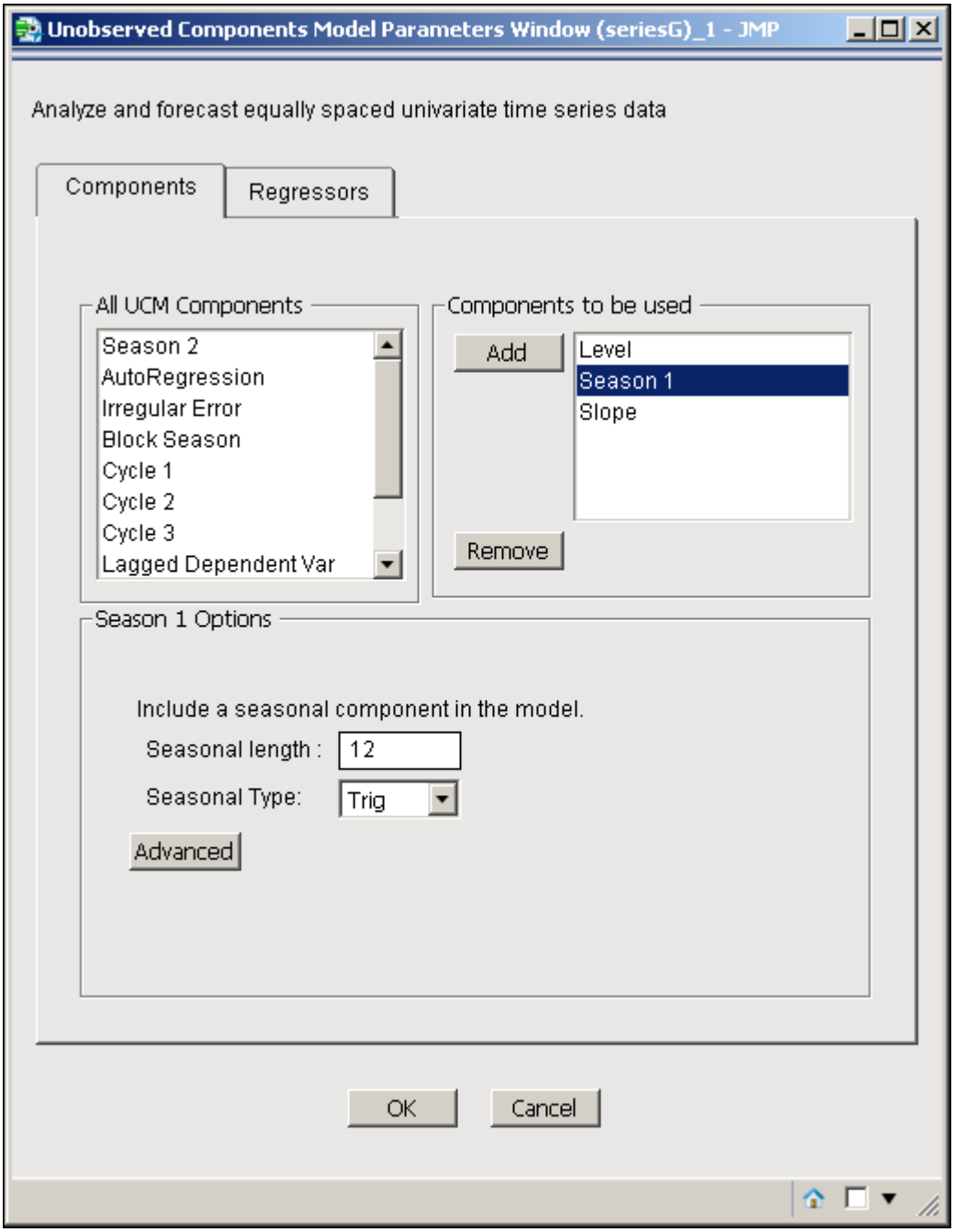
You can modify the plots and tables shown by clicking on the red triangles next to headings in the results window. You can also collapse or expand sections by clicking on the gray triangles next to the headings in the results window.

Step 3. Create and Estimate an Alternative Model in the Dialog Window

1 In the Unobserved Components Model Results window, you can open a Unobserved Components Model Parameters dialog window (see Figure 6.5) to create a new model by doing one of the following:

- Click the red triangle next to the **Fitted Models** header, and then select **New Model**. The Unobserved Components Model Parameters dialog window appears with the current model settings.
- Click the red triangle next to **Model 1** in the **Fitted Models** area, and select one of the following commands:
  - Select **Edit** to open the Unobserved Components Model Parameters dialog window. This window contains the current model settings.
    NOTE: If you choose this option and perform the remaining steps, you can no longer compare the base model to any alternative models.
  - Select **Copy** to create a copy of the base model. Then click the red triangle next to **Copy of Model 1**, and select **Edit** to open the Unobserved Components Model Parameters dialog window. This window contains the values for the copied model.

Figure 6.5 Example Unobserved Components Model Parameters Dialog Window



2 In the Unobserved Components Model Parameters dialog window, specify an alternative model. For more information about this window and its tabs, see the following section.

3 Click **OK** to estimate the alternative model. The Unobserved Components Model Parameters dialog window closes, and the model results for the alternative model appear in the Unobserved Components Model Results window. For more information, see the section "Results Window" on page 5.

The Unobserved Components Model Parameters Dialog Window

The Unobserved Components Model Parameters dialog window contains two tabs, which are described in the following sections. Options on these tabs correspond to options or statements for the SAS/ETS UCM procedure.

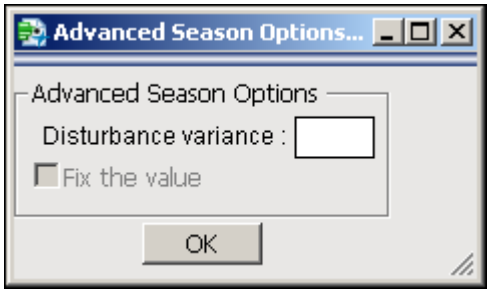
Components Tab

On the **Components** tab, you specify which components to use in an unobserved component model. Figure 6.5 shows an example of the **Components** tab.

By default, the **Components** tab includes level, seasonal, and slope components, but you can add or remove components. To add a component, select the component from the **All UCM Components** list and then click **Add**. To remove a component, select the component from the **Components to be used** list and then click **Remove**.

If you click a component in the **Components to be used** area, a description and required options for the selected component appear in the **Options** area. To specify optional options, click **Advanced** to open the Advanced Options window. See Figure 6.6 for an example of an Advanced Options window.

Figure 6.6 Example Advanced Options Window



Specify any options you want in the Advanced Options window, and click **OK** to save the selected options and close the window. The following sections describe the options associated with each component.

Options for a Level Component

When you select **Level** from the **Components to be used** list, there are no required options. Click **Advanced** to open the Advanced Options window where you can specify the following options, which correspond to options in the LEVEL statement of the UCM procedure:

Disturbance variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any

nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST option.

Check Breaks
specifies whether to check for breaks in the level component. This option corresponds to the CHECKBREAK option.

Options for a Season 1 or Season 2 Component

When you select **Season 1** or **Season 2** from the **Components to be used** list, you can also specify values for the following required options (these options correspond to options in the SEASON statement of the UCM procedure):

Seasonal Length
specifies the season length. The season length can be any integer greater than or equal to 2. Typical examples of season lengths are 12, which corresponds to monthly seasonality, and 4, which corresponds to quarterly seasonality. This option corresponds to the LENGTH= option.

Seasonal Type
specifies the type of the seasonal component. The default type is DUMMY. This option corresponds to the TYPE= option.

Click **Advanced** to specify the following options in the Advanced Options window (these options correspond to options in the SEASON statement of the UCM procedure):

Disturbance variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST=VARIANCE option.

Options for a Slope Component

When you select **Slope** from the **Components to be used** list, there are no required options. Click **Advanced** to open the Advanced Options window where you can specify the following options, which correspond to options in the SLOPE statement of the UCM procedure:

Disturbance variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any

nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST=VARIANCE option.

Options for an Autoregression Component

When you select **Autoregression** from the **Components to be used** list, there are no required options. Click **Advanced** to open the Advanced Options window where you can specify the following options, which correspond to options in the AUTOREG statement of the UCM procedure:

Damping factor RHO
specifies an initial value for the damping factor ρ during the parameter estimation process. The value of ρ must be in the interval $[-1, 1)$ (which includes –1, but excludes 1). This option corresponds to the RHO= option.

Fix the value
specifies whether to fix the ρ to the value specified in the **Damping factor RHO** option. This option corresponds to the NOEST=RHO option.

Disturbance Variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST=VARIANCE option.

Options for an Irregular Error Component

When you select **Irregular Error** from the **Components to be used** list, there are no required options. Click **Advanced** to open the Advanced Options window where you can specify the following options, which correspond to options in the IRREGULAR statement of the UCM procedure:

Error Variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Error variance** option. This option corresponds to the NOEST option.

ARMA Specification
specifies options for an ARMA model for the irregular component:

P

specifies the order of the nonseasonal autoregressive polynomial. The order can be any nonnegative integer; the default value is 0. In practice, the order is a small integer such as 1, 2, or 3. This option corresponds to the P= option.

SP

specifies the order of the seasonal autoregressive polynomial. The order can be any nonnegative integer; the default value is 0. In practice, the order is a small integer such as 1 or 2. This option corresponds to the SP= option.

Q

specifies the order of the nonseasonal moving average polynomial. The order can be any nonnegative integer; the default value is 0. In practice, the order is a small integer such as 1, 2, or 3. This option corresponds to the Q= option.

SQ

specifies the order of the seasonal moving average polynomial. The order can be any nonnegative integer; the default value is 0. In practice, the order is a small integer such as 1 or 2. This option corresponds to the SQ= option.

S

specifies the season length used during the specification of the seasonal autoregressive or seasonal moving average polynomial. The season length can be any positive integer; for example, S=4 might be an appropriate value for a quarterly series. The default value is S=1. This option corresponds to the S= option.

Options for a Block Season Component

When you select **Block Season** from the **Components to be used** list, you can also specify values for the following required options (these options correspond to options in the BLOCKSEASON statement of the UCM procedure):

Block Size

specifies the block size. The block size can be any integer larger than or equal to 2. Typical examples of block sizes are 24 (which corresponds to the hours of the day when a day is used as a block in hourly data), or 60 (which corresponds to the minutes in an hour when an hour is used as a block in data that are recorded by minutes). This option corresponds to the BLOCKSIZE= option.

Number of Blocks

specifies the number of blocks. The number of blocks can be any integer greater than or equal to 2. This option corresponds to the NBLOCKS= option.

Click **Advanced** to specify the following options in the Advanced Options window (these options correspond to options in the BLOCKSEASON statement of the UCM procedure):

Position of the first measurement

specifies the position of the first measurement within the block, if the first measurement is not at the start of a block. The value must be between 1 and the block size. The default value is 1. The first measurement refers to the start of the estimation span and the forecast span. If these spans differ,

their starting measurements must be separated by an integer multiple of the block size. This option corresponds to the OFFSET= option.

Block Seasonal Type
specifies the type of the block seasonal component. The default type is DUMMY. This option corresponds to the TYPE= option.

Disturbance Variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST option.

Options for a Cycle 1, Cycle 2, or Cycle 3 Component

When you select **Cycle 1**, **Cycle 2**, or **Cycle 3** from the **Components to be used** list, you can also specify values for the following required options (these options correspond to options in the CYCLE statement of the UCM procedure):

Cycle period
specifies an initial value for the cycle period during the parameter estimation process. The value must be strictly greater than 2. This option corresponds to the PERIOD= option.

Fix the value
specifies whether to fix the cycle period values to those specified in the **Cycle period** option. This option corresponds to the NOEST=PERIOD option.

Damping factor RHO
specifies an initial value for the damping factor ρ during the parameter estimation process. The value of ρ must be in the interval $[-1, 1)$ (which includes –1, but excludes 1). This option corresponds to the RHO= option.

Fix the value
specifies whether to fix the damping factor ρ to the value specified in the **Damping factor RHO** option. This option corresponds to the NOEST=RHO option.

Disturbance variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST option.

Options for a Lagged Dependent Var Component

When you select **Lagged Dependent Var** from the **Components to be used** list, there are no required options. Click **Advanced** to open the Advanced Options window where you can specify the following options, which correspond to options in the DEPLAG statement of the UCM procedure:

Lags
specify the lags of the dependent variable to be included as predictors in the model. This option corresponds to the LAGS= option.

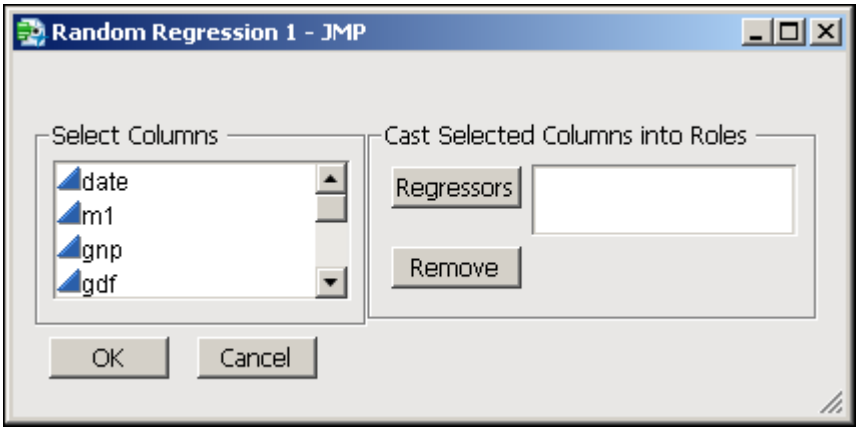
PHI
lists starting values for the coefficients of the lagged dependent variable. The order of the values listed corresponds with the order of the lags specified in the **Lags** option. This option corresponds to the PHI= option.

Fix the coefficient value
specifies whether to fix the coefficients to the values specified in **PHI** option. This option corresponds to the NOEST option.

Options for a Random Regression Component

When you select **Random Regression** from the **Components to be used** list, there are no required options. Click **Edit selections** to open the Random Regression window where you can add or remove regressors with time-varying regression coefficients (these options correspond to options in the RANDREG statement of the UCM procedure). Figure 6.7 shows an example of the Random Regression options window.

Figure 6.7 Example Random Regression Options Window



To add a Regressor (Z) variable to the model:

1. Click **Edit selections**.
2. Select the variable in the **Select Columns** list. You can select more than one variable at a time by holding CTRL and clicking each variable you want to select.
3. Click **Regressors**. The variables you selected appear in the **Cast Selected Columns into Roles** list.

4 Click **OK**.

To remove a Regressor (Z) variable from the model:

1 Click **Edit selections**.

2 Select the variable in the **Cast Selected Columns into Roles** list.

3 Click **Remove**. The variable you selected no longer appears in the **Cast Selected Columns into Roles** list.

4 Click **OK**.

Click **Advanced** to specify the following options in the Advanced Options window (these options correspond to options in the RANDREG statement of the UCM procedure):

Disturbance Variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST option.

Options for a Spline Regression Component

When you select **Spline Regression** from the **Components to be used** list, there are no required options. Click **Edit selections** to open the Spline Regression window where you can add or remove regressors that have a nonlinear relationship with the outcome variable that can be approximated by a given B-spline. These options correspond to the options in the SPLINEREG statement of the UCM procedure.

To add a Regressor (Z) variable to the model:

1 Click **Edit selections**.

2 Select the variable in the **Select Columns** list. You can select more than one variable at a time by holding CTRL and clicking each variable you want to select.

3 Click **Regressors**. The variables you selected appear in the **Cast Selected Columns into Roles** list.

4 Click **OK**.

To remove a Regressor (Z) variable from the model:

1 Click **Edit selections**.

2 Select the variable in the **Cast Selected Columns into Roles** list.

3 Click **Remove**. The variable you selected no longer appears in the **Cast Selected Columns into Roles** list.

4 Click **OK**.

Click **Advanced** to specify the following options in the Advanced Options window (these options correspond to options in the SPLINEREG statement of the UCM procedure):

Degree of the spline
specifies the degree of the spline, *n*. *n* can be any integer greater than or equal to 0. The default value is 3. The polynomial degree should be a small integer, usually 0, 1, 2, or 3. Larger values are rarely useful. If you have any doubt as to what degree to specify, use the default. This option corresponds to the DEGREE= option.

Disturbance Variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST option.

Num of Knots
specifies the number of knots, *m*. The first knot is placed at the $100/(m + 1)$ percentile, the second knot at the $200/(m + 1)$ percentile, and so on. Knots are always placed at data values; there is no interpolation. This option corresponds to the NKNOTS= option.

Knots List
specifies the interior knots or break points. The values in the knot list must be nondecreasing and must lie between the minimum and the maximum of the spline regressor values in the input data set. The first time you specify a value in the knot list, it indicates a discontinuity in the *n*th (from the degree of the spline option) derivative of the transformation function at the value of the knot. The second mention of a value indicates a discontinuity in the $(n–1)$th derivative of the transformation function at the value of the knot. Knots can be repeated any number of times for decreasing smoothness at the break points, but the values in the knot list can never decrease. This option corresponds to the KNOTS= option.

Spline Season

When you select **Spline Season** from the **Components to be used** list, you can also specify values for the following required options (these options correspond to options in the SPLINESEASON statement of the UCM procedure):

Seasonal Length
specifies the season length. The season length can be any integer greater than or equal to 2. Typical examples of season lengths are 12 (which corresponds to monthly seasonality), or 4 (which corresponds to quarterly seasonality). This option corresponds to the LENGTH= option.

Click **Advanced** to specify the following options in the Advanced Options window (these options correspond to options in the SPLINESEASON statement of the UCM procedure):

Degree of the spline
specifies the degree of the spline. The value can be any integer greater than or equal to 0. The default value is 3. The polynomial degree should be a small integer, usually 0, 1, 2, or 3. Larger values are rarely useful. If you have any doubt as to what degree to specify, use the default. This option corresponds to the DEGREE= option.

Position of the first measurement
specifies the position of the first measurement within the season, if the first measurement is not at the start of the season. The value must be between 1 and the season length. The default is 1. The first measurement refers to the start of the estimation span and the forecast span. If these spans differ, their starting measurements must be separated by an integer multiple of the season length. This option corresponds to the OFFSET= option.

Disturbance Variance
specifies an initial value for the disturbance variance during the parameter estimation process. Any nonnegative value, including 0, is an acceptable starting value. This option corresponds to the VARIANCE= option.

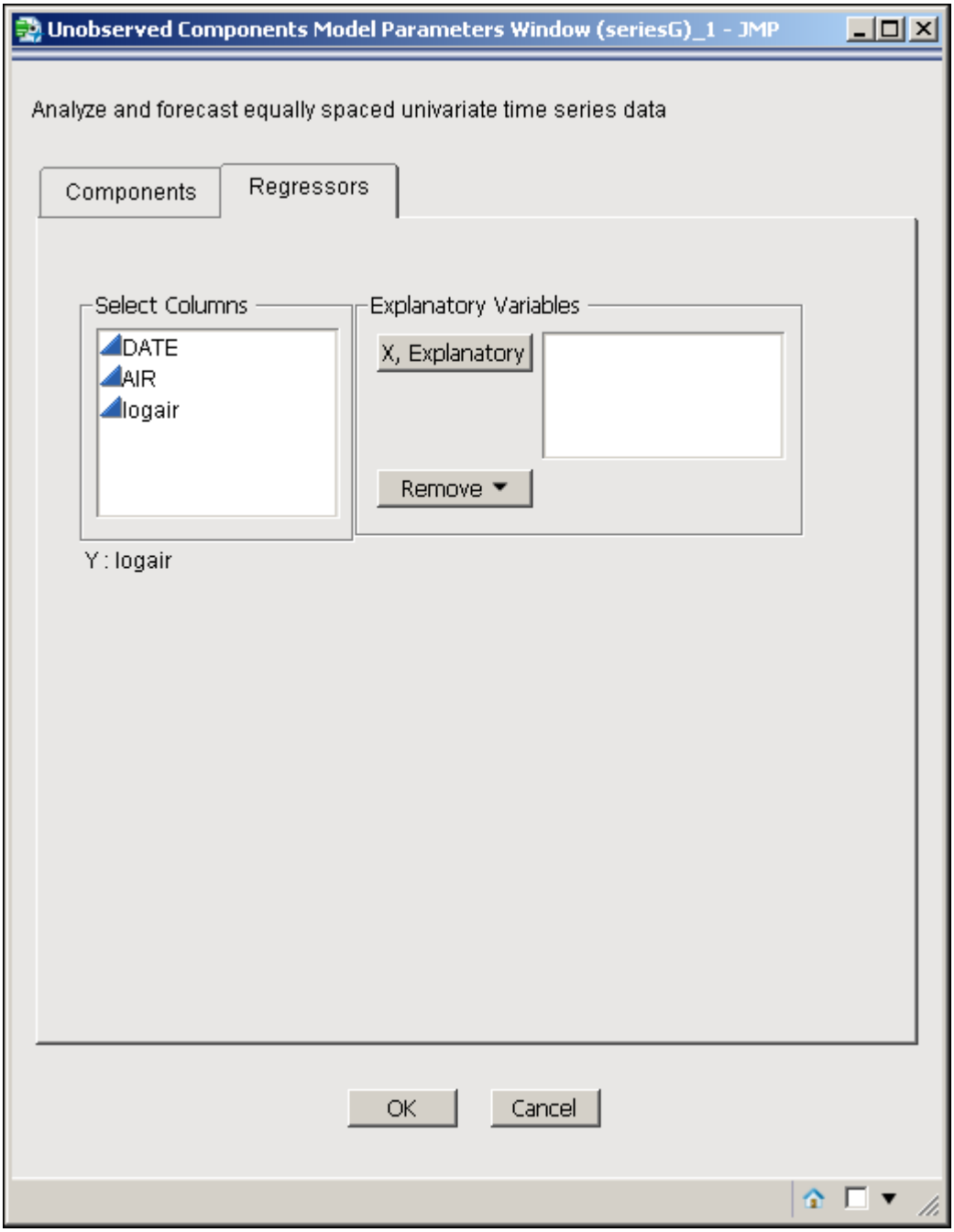
Fix the value
specifies whether to fix the disturbance to the value specified in the **Disturbance variance** option. This option corresponds to the NOEST option.

Knots
lists the internal knots. This list of values must be a nondecreasing sequence of integers within the range of 2 to $s-1$, where s is the season length specified in the **Seasonal Length** option. This option corresponds to the KNOTS= option.

Regressors Tab

On the **Regressors** tab, you specify the X (explanatory) variables, which are sometimes referred to as regressor variables. An example of the **Regressor** tab is shown in Figure 6.8.

Figure 6.8 Unobserved Components Model Parameters Window with the Regressor Tab



To add a regressor variable to the model:

1 Select the variable in the **Select Columns** list. You can select more than one variable at a time by holding CTRL and clicking each variable you want to select.

2 Click **X, Explanatory**. The variables you selected appear in the **Explanatory Variables** list.

To remove a regressor variable from the model:

1 Select the variable in the **Explanatory Variables** list.

2 Click **Remove**, and click **Selected**. The variable you selected no longer appears in the **Explanatory Variables** list. You can remove all of the variables in the **Explanatory Variables** list by clicking **Remove** and selecting **All**.

NOTE: Models with different X variables can be compared, either by using the fitted or residual graphs or by using the overall model fit information (for example, AIC or BIC) for each model.

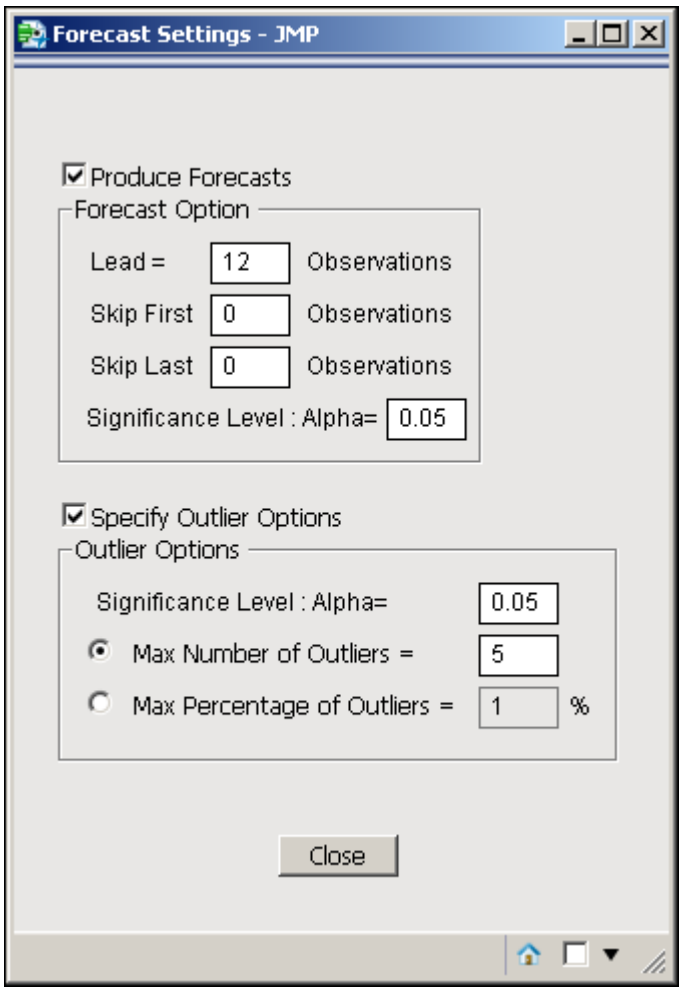
The Settings Window

The **Fit Unobserved Components Models** analysis provides three settings windows: SAS Graphs Settings, Display Settings, and Forecast Settings. The SAS Graphs Settings window and Display Settings window are described in the section "Settings Windows" on page 10. The Forecast Settings window is described in the following section.

Forecast Settings Window

In the Forecast Settings window, you specify the forecast and outlier options. Figure 6.9 shows an example of the Forecast Settings window.

Figure 6.9 Forecast Settings Window for Fit Unobserved Components Models



The Forecast Settings window contains the following options which correspond to options in the FORECAST and OUTLIER statements of the UCM procedure:

Produce Forecasts

specifies the overall forecasting environment for the specified model. This option corresponds to including the FORECAST statement in the UCM procedure. When you select this check box, you can also specify the following options, which correspond to options in the FORECAST statement of the UCM procedure:

Lead=

specifies the number of periods to forecast beyond the historical period defined by the **Skip First** and **Skip Last** options. For example, entering 10 in the **Lead** area results in the forecasting of 10 future values of the response series. The default is 12. This option corresponds to the LEAD= option.

Skip First and Skip Last

specify the holdout sample for the evaluation of the forecasting performance of the model. For example, specifying 10 in the **Skip Last** area results in treating the last 10 observed values

of the response series as unobserved. A post-sample prediction analysis table is produced for comparing the predicted values with the actual values in the holdout period. The default is 0. These options correspond to the SKIPLAST= and SKIPFIRST= options, respectively.

Alpha=
specifies the significance level of the forecast confidence intervals. For example, ALPHA=0.05, which is the default, results in a 95% confidence interval.

Specify Outlier Options
enables you to control the reporting of the additive outliers (AO) and level shifts (LS) in the response series. The AOs are searched by default. This option corresponds to including the OUTLIER statement in the UCM procedure. When you select this check box, you can also specify the following options, which correspond to options in OUTLIER statement of the UCM procedure:

Alpha=
specifies the significance level for reporting outliers. The default is 0.05, which results in a 95% confidence interval. This option corresponds to the ALPHA= option.

Max Number of Outliers
limits the number of outliers to search. The default is 5. This option corresponds to the MAXNUM= option.

Max Percentage of Outliers
limits the number of outliers to search. The default is 1. This option corresponds to the MAXPCT= option.

Your Turn

We welcome your feedback.

- If you have comments about this book, please send them to **yourturn@sas.com**. Include the full title and page numbers (if applicable).
- If you have comments about the software, please send them to **suggest@sas.com**.