SAS/ETS® 14.2 User's Guide
The SIMLIN Procedure

This document is an individual chapter from *SAS/ETS® 14.2 User's Guide*.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2016. *SAS/ETS® 14.2 User's Guide*. Cary, NC: SAS Institute Inc.

SAS/ETS® 14.2 User's Guide

Copyright © 2016, SAS Institute Inc., Cary, NC, USA

All Rights Reserved. Produced in the United States of America.

For a hard-copy book: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government License Rights; Restricted Rights: The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication, or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a), and DFAR 227.7202-4, and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

November 2016

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

SAS software may be provided with certain third-party software, including but not limited to open-source software, which is licensed under its applicable third-party software license agreement. For license information about third-party software distributed with SAS software, refer to http://support.sas.com/thirdpartylicenses.

Chapter 31
The SIMLIN Procedure

Contents

Overview: SIMLIN Procedure

The SIMLIN procedure reads the coefficients for a set of linear structural equations, which are usually produced by the SYSLIN procedure. PROC SIMLIN then computes the reduced form and, if input data are given, uses the reduced form equations to generate predicted values. PROC SIMLIN is especially useful when dealing with sets of structural difference equations. The SIMLIN procedure can perform simulation or forecasting of the endogenous variables.

The SIMLIN procedure can be applied only to models that are as follows:

- linear with respect to the parameters
- linear with respect to the variables
- square (as many equations as endogenous variables)
- nonsingular (the coefficients of the endogenous variables form an invertible matrix)

Getting Started: SIMLIN Procedure

The SIMLIN procedure processes the coefficients in a data set created by the SYSLIN procedure using the OUTEST= option or by another regression procedure such as PROC REG. To use PROC SIMLIN you must first produce the coefficient data set and then specify this data set in the EST= option of the PROC SIMLIN statement. You must also tell PROC SIMLIN which variables are endogenous and which variables are exogenous. List the endogenous variables in an ENDOGENOUS statement, and list the exogenous variables in an EXOGENOUS statement.

The following example illustrates the creation of an OUTEST= data set with PROC SYSLIN and the computation and printing of the reduced form coefficients for the model with PROC SIMLIN:

```
proc syslin data=in outest=e;
   model y1 = y2 x1;
   model y2 = y1 x2;
run;

proc simlin est=e;
   endogenous y1 y2;
   exogenous x1 x2;
run;
```

If the model contains lagged endogenous variables you must also use a LAGGED statement to tell PROC SIMLIN which variables contain lagged values, which endogenous variables they are lags of, and the number of periods of lagging. For dynamic models, the TOTAL and INTERIM= options can be used in the PROC SIMLIN statement to compute and print total and impact multipliers. (For an explanation of multipliers, see the section “Dynamic Multipliers” on page 2327.)

In the following example, the variables Y1LAG1, Y2LAG1, and Y2LAG2 contain lagged values of the endogenous variables Y1 and Y2. Y1LAG1 and Y2LAG1 contain values of Y1 and Y2 for the previous observation, while Y2LAG2 contains 2 period lags of Y2. The LAGGED statement specifies the lagged relationships, and the TOTAL and INTERIM= options request multiplier analysis. The INTERIM=2 option prints matrices showing the impact that changes to the exogenous variables have on the endogenous variables after 1 and 2 periods.

```
data in; set in;
  y1lag1 = lag(y1);
  y2lag1 = lag(y2);
  y2lag2 = lag2(y2);
run;

proc syslin data=in outest=e;
   model y1 = y2 y1lag1 y2lag2 x1;
   model y2 = y1 y2lag1 x2;
run;

proc simlin est=e total interim=2;
   endogenous y1 y2;
   exogenous x1 x2;
   lagged y1lag1 y1 1 y2lag1 y2 1 y2lag2 y2 2;
run;
```

After the reduced form of the model is computed, the model can be simulated by specifying an input data set in the PROC SIMLIN statement and using an OUTPUT statement to write the simulation results to an output data set. The following example modifies the PROC SIMLIN step from the preceding example to simulate the model and stores the results in an output data set:

```
proc simlin est=e total interim=2 data=in;
   endogenous y1 y2;
   exogenous x1 x2;
   lagged y1lag1 y1 1 y2lag1 y2 1 y2lag2 y2 2;
   output out=sim predicted=y1hat y2hat
                  residual=y1resid y2resid;
run;
```

Prediction and Simulation

If an input data set is specified with the DATA= option in the PROC SIMLIN statement, the procedure reads the data and uses the reduced form equations to compute predicted and residual values for each of the endogenous variables. (If no data set is specified with the DATA= option, no simulation of the system is performed, and only the reduced form and multipliers are computed.)

The character of the prediction is based on the START= value. Until PROC SIMLIN encounters the START= observation, actual endogenous values are found and fed into the lagged endogenous terms. Once the START= observation is reached, dynamic simulation begins, where predicted values are fed into lagged endogenous terms until the end of the data set is reached.

The predicted and residual values generated here are different from those produced by the SYSLIN procedure since PROC SYSLIN uses the structural form with actual endogenous values. The predicted values computed by the SIMLIN procedure solve the simultaneous equation system. These reduced-form predicted values are functions only of the exogenous and lagged endogenous variables and do not depend on actual values of current period endogenous variables.

Syntax: SIMLIN Procedure

The following statements can be used with PROC SIMLIN:

PROC SIMLIN *options* ;
BY *variables* ;
ENDOGENOUS *variables* ;
EXOGENOUS *variables* ;
ID *variables* ;
LAGGED *lag-var endogenous-var number . . .* ;
OUTPUT OUT=*SAS-data-set options* ;

Functional Summary

The statements and options controlling the SIMLIN procedure are summarized in Table 31.1.

Table 31.1 Functional Summary

Description	Statement	Option
Data Set Options		
Specify input data set containing structural coefficients	PROC SIMLIN	EST=
Specify type of estimates read from EST= data set	PROC SIMLIN	TYPE=
Write reduced form coefficients and multipliers to an output data set	PROC SIMLIN	OUTEST=
Specify the input data set for simulation	PROC SIMLIN	DATA=
Write predicted and residual values to an output data set	OUTPUT	
Printing Control Options		
Print the structural coefficients	PROC SIMLIN	ESTPRINT
Suppress printing of reduced form coefficients	PROC SIMLIN	NORED
Suppress all printed output	PROC SIMLIN	NOPRINT
Dynamic Multipliers		
Compute interim multipliers	PROC SIMLIN	INTERIM=
Compute total multipliers	PROC SIMLIN	TOTAL
Declaring the Role of Variables		
Specify BY-group processing	BY	
Specify the endogenous variables	ENDOGENOUS	
Specify the exogenous variables	EXOGENOUS	
Specify identifying variables	ID	
Specify lagged endogenous variables	LAGGED	

Table 31.1 *continued*

Description	Statement	Option
Controlling the Simulation		
Specify the starting observation for dynamic simulation	PROC SIMLIN	START=

PROC SIMLIN Statement

PROC SIMLIN *options* **;**

The following options can be used in the PROC SIMLIN statement:

DATA=*SAS-data-set*
specifies the SAS data set containing input data for the simulation. If the DATA= option is used, the data set specified must supply values for all exogenous variables throughout the simulation. If the DATA= option is not specified, no simulation of the system is performed, and only the reduced form and multipliers are computed.

EST=*SAS-data-set*
specifies the input data set containing the structural coefficients of the system. If EST= is omitted the most recently created SAS data set is used. The EST= data set is normally a "TYPE=EST" data set produced by the OUTEST= option of PROC SYSLIN. However, you can also build the EST= data set with a SAS DATA step. For more information, see the section "EST= Data Set" on page 2328.

ESTPRINT
prints the structural coefficients read from the EST= data set.

INTERIM=*n*
requests that interim multipliers be computed for interim numbers 1 through n. If not specified, no interim multipliers are computed. This feature is available only if there are no lags greater than 1.

NOPRINT
suppresses all printed output.

NORED
suppresses the printing of the reduced form coefficients.

OUTEST=*SAS-data-set*
specifies an output SAS data set to contain the reduced form coefficients and multipliers, in addition to the structural coefficients read from the EST= data set. The OUTEST= data set has the same form as the EST= data set. If the OUTEST= option is not specified, the reduced form coefficients and multipliers are not written to a data set.

START=*n*
specifies the observation number in the DATA= data set where the dynamic simulation is to be started. By default, the dynamic simulation starts with the first observation in the DATA= data set for which all variables (including lags) are not missing.

TOTAL

requests that the total multipliers be computed. This feature is available only if there are no lags greater than 1.

TYPE=*value*

specifies the type of estimates to be read from the EST= data set. The TYPE= value must match the value of the _TYPE_ variable for the observations that you want to select from the EST= data set (TYPE=2SLS, for example).

BY Statement

BY *variables* **;**

A BY statement can be used with PROC SIMLIN to obtain separate analyses for groups of observations defined by the BY variables.

The BY statement can be applied to one or both of the EST= and DATA= input data sets. When a BY statement is used and both an EST= and a DATA= input data set are specified, PROC SIMLIN checks to see if one or both of the data sets contain the BY variables.

Thus, there are three ways of using the BY statement with PROC SIMLIN:

1. If the BY variables are found in the EST= data set only, PROC SIMLIN simulates over the entire DATA= data set once for each set of coefficients read from the BY groups in the EST= data set.

2. If the BY variables are found in the DATA= data set only, PROC SIMLIN performs separate simulations over each BY group in the DATA= data set, using the single set of coefficients in the EST= data set.

3. If the BY variables are found in both the EST= and DATA= data sets, PROC SIMLIN performs separate simulations over each BY group in the DATA= data set using the coefficients from the corresponding BY group in the EST= data set.

ENDOGENOUS Statement

ENDOGENOUS *variables* **;**

List the names of the endogenous (jointly dependent) variables in the ENDOGENOUS statement. The ENDOGENOUS statement can be abbreviated as ENDOG or ENDO.

EXOGENOUS Statement

EXOGENOUS *variables* **;**

List the names of the exogenous (independent) variables in the EXOGENOUS statement. The EXOGENOUS statement can be abbreviated as EXOG or EXO.

ID Statement

ID *variables* **;**

The ID statement can be used to restrict the variables copied from the DATA= data set to the OUT= data set. Use the ID statement to list the variables you want copied to the OUT= data set besides the exogenous, endogenous, lagged endogenous, and BY variables. If the ID statement is omitted, all the variables in the DATA= data set are copied to the OUT= data set.

LAGGED Statement

LAGGED *lag-var endogenous-var number . . .* **;**

For each lagged endogenous variable, specify the name of the lagged variable, the name of the endogenous variable that was lagged, and the degree of the lag. Only one LAGGED statement is allowed.

The following is an example of the use of the LAGGED statement:

```
proc simlin est=e;
   endog y1 y2;
   lagged y1lag1 y1 1  y2lag1 y2 1  y2lag3 y2 3;
run;
```

This statement specifies that the variable Y1LAG1 contains the values of the endogenous variable Y1 lagged one period; the variable Y2LAG1 refers to the values of Y2 lagged one period; and the variable Y2LAG3 refers to the values of Y2 lagged three periods.

OUTPUT Statement

OUTPUT OUT=*SAS-data-set options* **;**

The OUTPUT statement specifies that predicted and residual values be put in an output data set. A DATA= input data set must be supplied if the OUTPUT statement is used, and only one OUTPUT statement is allowed. The following options can be used in the OUTPUT statement:

OUT=*SAS-data-set*

names the output SAS data set to contain the predicted values and residuals. If OUT= is not specified, the output data set is named using the DATA*n* convention.

PREDICTED=*names*

P=*names*

names the variables in the output data set that contain the predicted values of the simulation. These variables correspond to the endogenous variables in the order in which they are specified in the ENDOGENOUS statement. Specify up to as many names as there are endogenous variables. If you specify names in the PREDICTED= option for only some of the endogenous variables, predicted values for the remaining variables are not output. The names must not match any variable name in the input data set.

RESIDUAL=*names*

R=*names*

names the variables in the output data set that contain the residual values from the simulation. The residuals are the differences between the actual values of the endogenous variables from the DATA= data set and the predicted values from the simulation. These variables correspond to the endogenous variables in the order in which they are specified in the ENDOGENOUS statement. Specify up to as many names as there are endogenous variables. The names must not match any variable name in the input data set.

The following is an example of the use of the OUTPUT statement. This example outputs predicted values for Y1 and Y2 and outputs residuals for Y1.

```
proc simlin est=e;
   endog y1 y2;
   output out=b predicted=y1hat y2hat
                residual=y1resid;
run;
```

Details: SIMLIN Procedure

The following sections explain the structural and reduced forms, dynamic multipliers, input data sets, and the model simulation process in more detail.

Defining the Structural Form

An EST= input data set supplies the coefficients of the equation system. The data set containing the coefficients is normally a "TYPE=EST" data set created by the OUTEST= option of PROC SYSLIN or another regression procedure. The data set contains the special variables _TYPE_, _DEPVAR_, and INTERCEPT. You can also supply the structural coefficients of the system to PROC SIMLIN in a data set produced by a SAS DATA step as long as the data set is of the form TYPE=EST. For a discussion of the special TYPE=EST type of SAS data set, see SAS/STAT software documentation.

Suppose that there is a $g \times 1$ vector of endogenous variables $\mathbf{y}_t$, an $l \times 1$ vector of lagged endogenous variables $\mathbf{y}_t^L$, and a $k \times 1$ vector of exogenous variables $\mathbf{x}_t$, including the intercept. Then, there are g structural equations in the simultaneous system that can be written

$$\mathbf{G}\mathbf{y}_t = \mathbf{C}\mathbf{y}_t^L + \mathbf{B}\mathbf{x}_t$$

where $\mathbf{G}$ is the matrix of coefficients of current period endogenous variables, $\mathbf{C}$ is the matrix of coefficients of lagged endogenous variables, and $\mathbf{B}$ is the matrix of coefficients of exogenous variables. $\mathbf{G}$ is assumed to be nonsingular.

Computing the Reduced Form

First, the SIMLIN procedure computes reduced form coefficients by premultiplying by $\mathbf{G}^{-1}$:

$$\mathbf{y}_t = \mathbf{G}^{-1}\mathbf{C}\mathbf{y}_t^L + \mathbf{G}^{-1}\mathbf{B}\mathbf{x}_t$$

This can be written as

$$\mathbf{y}_t = \Pi_1 \mathbf{y}_t^L + \Pi_2 \mathbf{x}_t$$

where $\Pi_1 = \mathbf{G}^{-1}\mathbf{C}$ and $\Pi_2 = \mathbf{G}^{-1}\mathbf{B}$ are the reduced form coefficient matrices.

The reduced form matrices $\Pi_1 = \mathbf{G}^{-1}\mathbf{C}$ and $\Pi_2 = \mathbf{G}^{-1}\mathbf{B}$ are printed unless the NORED option is specified in the PROC SIMLIN statement. The structural coefficient matrices $\mathbf{G}$, $\mathbf{C}$, and $\mathbf{B}$ are printed when the ESTPRINT option is specified.

Dynamic Multipliers

For models that have only first-order lags, the equation of the reduced form of the system can be rewritten

$$\mathbf{y}_t = \mathbf{D}\mathbf{y}_{t-1} + \Pi_2 \mathbf{x}_t$$

$\mathbf{D}$ is a matrix formed from the columns of Π_1 plus some columns of zeros, arranged in the order in which the variables meet the lags. The elements of Π_2 are called *impact multipliers* because they show the immediate effect of changes in each exogenous variable on the values of the endogenous variables. This equation can be rewritten as

$$\mathbf{y}_t = \mathbf{D}^2\mathbf{y}_{t-2} + \mathbf{D}\Pi_2\mathbf{x}_{t-1} + \Pi_2 \mathbf{x}_t$$

The matrix formed by the product $\mathbf{D}\Pi_2$ shows the effect of the exogenous variables one lag back; the elements in this matrix are called *interim multipliers* and are computed and printed when the INTERIM= option is specified in the PROC SIMLIN statement. The ith period interim multipliers are formed by $\mathbf{D}^i\Pi_2$.

The series can be expanded as

$$\mathbf{y}_t = \mathbf{D}^{\infty}\mathbf{y}_{t-\infty} + \sum_{i=0}^{\infty} \mathbf{D}^i \Pi_2 \mathbf{x}_{t-i}$$

A permanent and constant setting of a value for x has the following cumulative effect:

$$\left(\sum_{i=0}^{\infty} \mathbf{D}^i\right) \Pi_2 \mathbf{x} = (\mathbf{I} - \mathbf{D})^{-1}\Pi_2 \mathbf{x}$$

The elements of $(\mathbf{I} - \mathbf{D})^{-1}\Pi_2$ are called the *total multipliers*. Assuming that the sum converges and that $(\mathbf{I} - \mathbf{D})$ is invertible, PROC SIMLIN computes the total multipliers when the TOTAL option is specified in the PROC SIMLIN statement.

Multipliers for Higher-Order Lags

The dynamic multiplier options require the system to have no lags of order greater than one. This limitation can be circumvented, since any system with lags greater than one can be rewritten as a system where no lag is greater than one by forming new endogenous variables that are single-period lags.

For example, suppose you have the third-order single equation

$$y_t = ay_{t-3} + b\mathrm{x}_t$$

This can be converted to a first-order three-equation system by introducing two additional endogenous variables, $y_{1,t}$ and $y_{2,t}$, and computing corresponding first-order lagged variables for each endogenous variable: y_{t-1}, $y_{1,t-1}$, and $y_{2,t-1}$. The higher-order lag relations are then produced by adding identities to link the endogenous and identical lagged endogenous variables:

$$y_{1,t} = y_{t-1}$$

$$y_{2,t} = y_{1,t-1}$$

$$y_t = ay_{2,t-1} + b\mathbf{X}_t$$

This conversion using the SYSLIN and SIMLIN procedures requires three steps:

1. Add the extra endogenous and lagged endogenous variables to the input data set using a DATA step. Note that two copies of each lagged endogenous variable are needed for each lag reduced, one to serve as an endogenous variable and one to serve as a lagged endogenous variable in the reduced system.

2. Add IDENTITY statements to the PROC SYSLIN step to equate each added endogenous variable to its lagged endogenous variable copy.

3. In the PROC SIMLIN step, declare the added endogenous variables in the ENDOGENOUS statement and define the lag relations in the LAGGED statement.

For an illustration of how to convert an equation system with higher-order lags into a larger system with only first-order lags, see Example 31.2.

EST= Data Set

Normally, PROC SIMLIN uses an EST= data set produced by PROC SYSLIN with the OUTEST= option. This data set is in the form expected by PROC SIMLIN. If there is more than one set of estimates produced by PROC SYSLIN, you must use the TYPE= option in the PROC SIMLIN statement to select the set to be simulated. Then PROC SIMLIN reads from the EST= data set only those observations with a _TYPE_ value corresponding to the TYPE= option (for example, TYPE=2SLS) or with a _TYPE_ value of IDENTITY.

The SIMLIN procedure can only solve square, nonsingular systems. If you have fewer equations than endogenous variables, you must specify IDENTITY statements in the PROC SYSLIN step to bring the system up to full rank. If there are g endogenous variables and $m < g$ stochastic equations with unknown parameters, then you use m MODEL statements to specify the equations with parameters to be estimated and you must use $g - m$ IDENTITY statements to complete the system.

You can build your own EST= data set with a DATA step rather than use PROC SYSLIN. The EST= data set must contain the endogenous variables, the lagged endogenous variables (if any), and the exogenous variables in the system (if any). If any of the equations have intercept terms, the variable INTERCEPT must supply these coefficients. The EST= data set should also contain the special character variable comp _DEPVAR_ to label the equations.

The EST= data set must contain one observation for each equation in the system. The values of the lagged endogenous variables must contain the **C** coefficients. The values of the exogenous variables and the INTERCEPT variable must contain the **B** coefficients. The values of the endogenous variables, however, must contain the negatives of the **G** coefficients. This is because the SYSLIN procedure writes the coefficients to the OUTEST= data set in the form

$$0 = \mathbf{H}\mathbf{y}_t + \mathbf{C}\mathbf{y}_t^L + \mathbf{B}\mathbf{x}_t$$

where $\mathbf{H} = -\mathbf{G}$.

For more information about building the EST= data set, see the section "Multipliers for Higher-Order Lags" on page 2327 and Example 31.2.

DATA= Data Set

The DATA= data set must contain all of the exogenous variables. Values for all of the exogenous variables are required for each observation for which predicted endogenous values are desired. To forecast past the end of the historical data, the DATA= data set should contain nonmissing values for all of the exogenous variables and missing values for the endogenous variables for the forecast periods, in addition to the historical data. (For an illustration, see Example 31.1.)

In order for PROC SIMLIN to output residuals and compute statistics of fit, the DATA= data set must also contain the endogenous variables with nonmissing actual values for each observation for which residuals and statistics are to be computed.

If the system contains lags, initial values must be supplied for the lagged variables. This can be done by including either the lagged variables or the endogenous variables, or both, in the DATA= data set. If the lagged variables are not in the DATA= data set or if they have missing values in the early observations, PROC SIMLIN prints a warning and uses the endogenous variable values from the early observations to initialize the lags.

OUTEST= Data Set

The OUTEST= data set contains all the variables read from the EST= data set. The variables in the OUTEST= data set are as follows:

- the BY statement variables, if any
- _TYPE_, a character variable that identifies the type of observation
- _DEPVAR_, a character variable containing the name of the dependent variable for the observation
- the endogenous variables
- the lagged endogenous variables
- the exogenous variables
- INTERCEPT, a numeric variable containing the intercept values

- _MODEL_, a character variable containing the name of the equation
- _SIGMA_, a numeric variable containing the estimated error variance of the equation (output only if present in the EST= data set)

The observations read from the EST= data set that supply the structural coefficients are copied to the OUTEST= data set, except that the signs of endogenous coefficients are reversed. For these observations, the _TYPE_ variable values are the same as in the EST= data set.

In addition, the OUTEST= data set contains observations with the following _TYPE_ values:

REDUCED — the reduced form coefficients. The endogenous variables for this group of observations contain the inverse of the endogenous coefficient matrix $\mathbf{G}$. The lagged endogenous variables contain the matrix $\Pi_1 = \mathbf{G}^{-1}\mathbf{C}$. The exogenous variables contain the matrix $\Pi_2 = \mathbf{G}^{-1}\mathbf{B}$.

IMULT*i* — the interim multipliers, if the INTERIM= option is specified. There are *gn* observations for the interim multipliers, where *g* is the number of endogenous variables and *n* is the value of the INTERIM=*n* option. For these observations the _TYPE_ variable has the value IMULT*i*, where the interim number *i* ranges from 1 to *n*.

The exogenous variables in groups of *g* observations that have a _TYPE_ value of IMULT*i* contain the matrix $\mathbf{D}^i\Pi_2$ of multipliers at interim *i*. The endogenous and lagged endogenous variables for this group of observations are set to missing.

TOTAL — the total multipliers, if the TOTAL option is specified. The exogenous variables in this group of observations contain the matrix $(\mathbf{I}-\mathbf{D})^{-1}\Pi_2$. The endogenous and lagged endogenous variables for this group of observations are set to missing.

OUT= Data Set

The OUT= data set normally contains all of the variables in the input DATA= data set, plus the variables named in the PREDICTED= and RESIDUAL= options in the OUTPUT statement.

You can use an ID statement to restrict the variables that are copied from the input data set. If an ID statement is used, the OUT= data set contains only the BY variables (if any), the ID variables, the endogenous and lagged endogenous variables (if any), the exogenous variables, plus the PREDICTED= and RESIDUAL= variables.

The OUT= data set contains an observation for each observation in the DATA= data set. When the actual value of an endogenous variable is missing in the DATA= data set, or when the DATA= data set does not contain the endogenous variable, the corresponding residual is missing.

Printed Output

Structural Form

The following items are printed as they are read from the EST= input data set. Structural zeros are printed as dots in the listing of these matrices.

1. Structural Coefficients for Endogenous Variables. This is the $\mathbf{G}$ matrix, with g rows and g columns.
2. Structural Coefficients for Lagged Endogenous Variables. These coefficients make up the $\mathbf{C}$ matrix, with g rows and l columns.
3. Structural Coefficients for Exogenous Variables. These coefficients make up the $\mathbf{B}$ matrix, with g rows and k columns.

Reduced Form

1. The reduced form coefficients are obtained by inverting $\mathbf{G}$ so that the endogenous variables can be directly expressed as functions of only lagged endogenous and exogenous variables.
2. Inverse Coefficient Matrix for Endogenous Variables. This is the inverse of the $\mathbf{G}$ matrix.
3. Reduced Form for Lagged Endogenous Variables. This is $\Pi_1 = \mathbf{G}^{-1}\mathbf{C}$, with g rows and l columns. Each value is a dynamic multiplier that shows how past values of lagged endogenous variables affect values of each of the endogenous variables.
4. Reduced Form for Exogenous Variables. This is $\Pi_2 = \mathbf{G}^{-1}\mathbf{B}$, with g rows and k columns. Its values are called *impact multipliers* because they show the immediate effect of each exogenous variable on the value of the endogenous variables.

Multipliers

Interim and total multipliers show the effect of a change in an exogenous variable over time.

1. Interim Multipliers. These are the interim multiplier matrices. They are formed by multiplying Π_2 by powers of $\mathbf{D}$. The dth interim multiplier is $\mathbf{D}^d\Pi_2$. The interim multiplier of order d shows the effects of a change in the exogenous variables after d periods. Interim multipliers are only available if the maximum lag of the endogenous variables is 1.
2. Total Multipliers. This is the matrix of total multipliers, $\mathrm{T} = (\mathbf{I} - \mathbf{D})^{-1}\Pi_2$. This matrix shows the cumulative effect of changes in the exogenous variables. Total multipliers are only available if the maximum lag is one.

Statistics of Fit

If the DATA= option is used and the DATA= data set contains endogenous variables, PROC SIMLIN prints a statistics-of-fit report for the simulation. The statistics printed include the following. (Summations are over the observations for which both y_t and $\hat{y}_t$ are nonmissing.)

1. the number of nonmissing errors. (Number of observations for which both y_t and $\hat{y}_t$ are nonmissing.)
2. the mean error: $\frac{1}{n}\sum(y_t - \hat{y}_t)$
3. the mean percent error: $\frac{100}{n}\sum\frac{(y_t-\hat{y}_t)}{y_t}$
4. the mean absolute error: $\frac{1}{n}\sum|y_t - \hat{y}_t|$
5. the mean absolute percent error $\frac{100}{n}\sum\frac{|y_t-\hat{y}_t|}{y_t}$
6. the root mean square error: $\sqrt{\frac{1}{n}\sum(y_t - \hat{y}_t)^2}$
7. the root mean square percent error: $\sqrt{\frac{100}{n}\sum(\frac{(y_t-\hat{y}_t)}{y_t})^2}$

ODS Table Names

PROC SIMLIN assigns a name to each table it creates. You can use these names to reference the table when using the Output Delivery System (ODS) to select tables and create output data sets. These names are listed in Table 31.2.

Table 31.2 ODS Tables Produced in PROC SIMLIN

ODS Table Name	Description	Option
Endogenous	Structural coefficients for endogenous variables	Default
LaggedEndogenous	Structural coefficients for lagged endogenous variables	Default
Exogenous	Structural coefficients for exogenous variables	Default
InverseCoeff	Inverse coefficient matrix for endogenous variables	Default
RedFormLagEndo	Reduced form for lagged endogenous variables	Default
RedFormExog	Reduced form for exogenous variables	Default
InterimMult	Interim multipliers	INTERIM=
TotalMult	Total multipliers	TOTAL=
FitStatistics	Fit statistics	Default

Examples: SIMLIN Procedure

Example 31.1: Simulating Klein's Model I

In this example, the SIMLIN procedure simulates a model of the U.S. economy called Klein's Model I. The SAS data set KLEIN is used as input to the SYSLIN and SIMLIN procedures.

```
data klein;
   input year c p w i x wp g t k wsum;
   date=mdy(1,1,year);
   format date year.;
   y    = c + i + g - t;
   yr   = year - 1931;
   klag = lag( k );
   plag = lag( p );
   xlag = lag( x );
   if year >= 1921;
   label c    ='consumption'
         p    ='profits'
         w    ='private wage bill'
         i    ='investment'
         k    ='capital stock'
         y    ='national income'
         x    ='private production'
         wsum='total wage bill'
         wp   ='govt wage bill'
         g    ='govt demand'
         t    ='taxes'
         klag='capital stock lagged'
         plag='profits lagged'
         xlag='private product lagged'
         yr   ='year-1931';
datalines;
1920      .  12.7     .     .  44.9     .     .     .  182.8     .
1921   41.9  12.4  25.5 -0.2  45.6   2.7    3.9   7.7  182.6  28.2

   ... more lines ...
```

First, the model is specified and estimated using the SYSLIN procedure, and the parameter estimates are written to an OUTEST= data set. The printed output produced by the SYSLIN procedure is not shown here; see Example 36.1 in Chapter 36 for the printed output of the PROC SYSLIN step.

```
title1 'Simulation of Klein''s Model I using SIMLIN';
proc syslin 3sls data=klein outest=a;

   instruments klag plag xlag wp g t yr;
   endogenous c p w i x wsum k y;

   consume: model    c = p plag wsum;
   invest:  model    i = p plag klag;
   labor:   model    w = x xlag yr;

   product: identity x = c + i + g;
   income:  identity y = c + i + g - t;
   profit:  identity p = x - w - t;
   stock:   identity k = klag + i;
   wage:    identity wsum = w + wp;
run;
```

The OUTEST= data set A created by the SYSLIN procedure contains parameter estimates to be used by the SIMLIN procedure. The OUTEST= data set is shown in Output 31.1.1.

Output 31.1.1 The OUTEST= Data Set Created by PROC SYSLIN

Simulation of Klein's Model I using SIMLIN

Obs	_TYPE_	_STATUS_	_MODEL_	_DEPVAR_	_SIGMA_	Intercept	klag	plag	xlag	wp
1	INST	0 Converged	FIRST	c	2.11403	58.3018	-0.14654	0.74803	0.23007	0.19327
2	INST	0 Converged	FIRST	p	2.18298	50.3844	-0.21610	0.80250	0.02200	-0.07961
3	INST	0 Converged	FIRST	w	1.75427	43.4356	-0.12295	0.87192	0.09533	-0.44373
4	INST	0 Converged	FIRST	i	1.72376	35.5182	-0.19251	0.92639	-0.11274	-0.71661
5	INST	0 Converged	FIRST	x	3.77347	93.8200	-0.33906	1.67442	0.11733	-0.52334
6	INST	0 Converged	FIRST	wsum	1.75427	43.4356	-0.12295	0.87192	0.09533	0.55627
7	INST	0 Converged	FIRST	k	1.72376	35.5182	0.80749	0.92639	-0.11274	-0.71661
8	INST	0 Converged	FIRST	y	3.77347	93.8200	-0.33906	1.67442	0.11733	-0.52334
9	3SLS	0 Converged	CONSUME	c	1.04956	16.4408	.	0.16314	.	.
10	3SLS	0 Converged	INVEST	i	1.60796	28.1778	-0.19485	0.75572	.	.
11	3SLS	0 Converged	LABOR	w	0.80149	1.7972	.	.	0.18129	.
12	IDENTITY	0 Converged	PRODUCT	x	.	0.0000	.	.	.	.
13	IDENTITY	0 Converged	INCOME	y	.	0.0000	.	.	.	.
14	IDENTITY	0 Converged	PROFIT	p	.	0.0000	.	.	.	.
15	IDENTITY	0 Converged	STOCK	k	.	0.0000	1.00000	.	.	.
16	IDENTITY	0 Converged	WAGE	wsum	.	0.0000	.	.	.	1.00000

Obs	g	t	yr	c	p	w	i	x	wsum	k	y
1	0.20501	-0.36573	0.70109	-1	.	.	.	.	.	.	.
2	0.43902	-0.92310	0.31941	.	-1.00000	.	.	.	.	.	.
3	0.86622	-0.60415	0.71358	.	.	-1	.	.	.	.	.
4	0.10023	-0.16152	0.33190	.	.	.	-1	.	.	.	.
5	1.30524	-0.52725	1.03299	.	.	.	.	-1.00000	.	.	.
6	0.86622	-0.60415	0.71358	.	.	.	.	.	-1.00000	.	.
7	0.10023	-0.16152	0.33190	.	.	.	.	.	.	-1	.
8	1.30524	-1.52725	1.03299	.	.	.	.	.	.	.	-1
9	.	.	.	-1	0.12489	.	.	.	0.79008	.	.
10	.	.	.	.	-0.01308	.	-1	.	.	.	.
11	.	.	0.14967	.	.	-1	.	0.40049	.	.	.
12	1.00000	.	.	1	.	.	1	-1.00000	.	.	.
13	1.00000	-1.00000	.	1	.	.	1	.	.	.	-1
14	.	-1.00000	.	.	-1.00000	-1	.	1.00000	.	.	.
15	.	.	.	.	.	.	1	.	.	-1	.
16	.	.	.	.	.	1	.	.	-1.00000	.	.

Using the OUTEST= data set A produced by the SYSLIN procedure, the SIMLIN procedure can now compute the reduced form and simulate the model. The following statements perform the simulation:

```
title1 'Simulation of Klein''s Model I using SIMLIN';
proc simlin data=klein
            est=a type=3sls
            estprint
            total interim=2
            outest=b;
   endogenous c p w i x wsum k y;
   exogenous  wp g t yr;
   lagged  klag k 1   plag p 1   xlag x 1;
   id year;
   output out=c p=chat phat what ihat xhat wsumhat khat yhat
                r=cres pres wres ires xres wsumres kres yres;
run;
```

The reduced form coefficients and multipliers are added to the information read from EST= data set A and written to the OUTEST= data set B. The predicted and residual values from the simulation are written to the OUT= data set C specified in the OUTPUT statement.

The SIMLIN procedure first prints the structural coefficient matrices read from the EST= data set, as shown in Output 31.1.2 through Output 31.1.4.

Output 31.1.2 SIMLIN Procedure Output — Endogenous Structural Coefficients

Simulation of Klein's Model I using SIMLIN

The SIMLIN Procedure

Structural Coefficients for Endogenous Variables

Variable	c	p	w	i	x	wsum	k	y
c	1.0000	-0.1249	.	.	.	-0.7901	.	.
i	.	0.0131	.	1.0000	.	.	.	.
w	.	.	1.0000	.	-0.4005	.	.	.
x	-1.0000	.	.	-1.0000	1.0000	.	.	.
y	-1.0000	.	.	-1.0000	.	.	.	1.0000
p	.	1.0000	1.0000	.	-1.0000	.	.	.
k	.	.	.	-1.0000	.	.	1.0000	.
wsum	.	.	-1.0000	.	.	1.0000	.	.

Output 31.1.3 SIMLIN Procedure Output — Lagged Endogenous Structural Coefficients

Structural Coefficients for Lagged Endogenous Variables

Variable	klag	plag	xlag
c	.	0.1631	.
i	-0.1948	0.7557	.
w	.	.	0.1813
x	.	.	.
y	.	.	.
p	.	.	.
k	1.0000	.	.
wsum	.	.	.

Output 31.1.4 SIMLIN Procedure Output — Exogenous Structural Coefficients

Structural Coefficients for Exogenous Variables					
Variable	wp	g	t	yr	Intercept
c	.	.	.	.	16.4408
i	.	.	.	.	28.1778
w	.	.	.	0.1497	1.7972
x	.	1.0000	.	.	0
y	.	1.0000	-1.0000	.	0
p	.	.	-1.0000	.	0
k	.	.	.	.	0
wsum	1.0000	.	.	.	0

The SIMLIN procedure then prints the inverse of the endogenous variables coefficient matrix, as shown in Output 31.1.5.

Output 31.1.5 SIMLIN Procedure Output — Inverse Coefficient Matrix

Inverse Coefficient Matrix for Endogenous Variables								
Variable	c	i	w	x	y	p	k	wsum
c	1.6347	0.6347	1.0957	0.6347	0	0.1959	0	1.2915
p	0.9724	0.9724	-0.3405	0.9724	0	1.1087	0	0.7682
w	0.6496	0.6496	1.4406	0.6496	0	0.0726	0	0.5132
i	-0.0127	0.9873	0.004453	-0.0127	0	-0.0145	0	-0.0100
x	1.6219	1.6219	1.1001	1.6219	0	0.1814	0	1.2815
wsum	0.6496	0.6496	1.4406	0.6496	0	0.0726	0	1.5132
k	-0.0127	0.9873	0.004453	-0.0127	0	-0.0145	1.0000	-0.0100
y	1.6219	1.6219	1.1001	0.6219	1.0000	0.1814	0	1.2815

The SIMLIN procedure next prints the reduced form coefficient matrices, as shown in Output 31.1.6.

Output 31.1.6 SIMLIN Procedure Output — Reduced Form Coefficients

Reduced Form for Lagged Endogenous Variables			
Variable	klag	plag	xlag
c	-0.1237	0.7463	0.1986
p	-0.1895	0.8935	-0.0617
w	-0.1266	0.5969	0.2612
i	-0.1924	0.7440	0.000807
x	-0.3160	1.4903	0.1994
wsum	-0.1266	0.5969	0.2612
k	0.8076	0.7440	0.000807
y	-0.3160	1.4903	0.1994

Output 31.1.6 *continued*

Reduced Form for Exogenous Variables					
Variable	**wp**	**g**	**t**	**yr**	**Intercept**
c	1.2915	0.6347	-0.1959	0.1640	46.7273
p	0.7682	0.9724	-1.1087	-0.0510	42.7736
w	0.5132	0.6496	-0.0726	0.2156	31.5721
i	-0.0100	-0.0127	0.0145	0.000667	27.6184
x	1.2815	1.6219	-0.1814	0.1647	74.3457
wsum	1.5132	0.6496	-0.0726	0.2156	31.5721
k	-0.0100	-0.0127	0.0145	0.000667	27.6184
y	1.2815	1.6219	-1.1814	0.1647	74.3457

The multiplier matrices (requested by the INTERIM=2 and TOTAL options) are printed next, as shown in Output 31.1.7 and Output 31.1.8.

Output 31.1.7 SIMLIN Procedure Output — Interim Multipliers

Interim Multipliers for Interim 1					
Variable	**wp**	**g**	**t**	**yr**	**Intercept**
c	0.829130	1.049424	-0.865262	-.0054080	43.27442
p	0.609213	0.771077	-0.982167	-.0558215	28.39545
w	0.794488	1.005578	-0.710961	0.0125018	41.45124
i	0.574572	0.727231	-0.827867	-.0379117	26.57227
x	1.403702	1.776655	-1.693129	-.0433197	69.84670
wsum	0.794488	1.005578	-0.710961	0.0125018	41.45124
k	0.564524	0.714514	-0.813366	-.0372452	54.19068
y	1.403702	1.776655	-1.693129	-.0433197	69.84670

Interim Multipliers for Interim 2					
Variable	**wp**	**g**	**t**	**yr**	**Intercept**
c	0.663671	0.840004	-0.968727	-.0456589	28.36428
p	0.350716	0.443899	-0.618929	-.0401446	10.79216
w	0.658769	0.833799	-0.925467	-.0399178	28.33114
i	0.345813	0.437694	-0.575669	-.0344035	10.75901
x	1.009485	1.277698	-1.544396	-.0800624	39.12330
wsum	0.658769	0.833799	-0.925467	-.0399178	28.33114
k	0.910337	1.152208	-1.389035	-.0716486	64.94969
y	1.009485	1.277698	-1.544396	-.0800624	39.12330

Output 31.1.8 SIMLIN Procedure Output — Total Multipliers

Total Multipliers

Variable	wp	g	t	yr	Intercept
c	1.881667	1.381613	-0.685987	0.1789624	41.3045
p	0.786945	0.996031	-1.286891	-.0748290	15.4770
w	1.094722	1.385582	-0.399095	0.2537914	25.8275
i	0.000000	0.000000	-0.000000	0.0000000	0.0000
x	1.881667	2.381613	-0.685987	0.1789624	41.3045
wsum	2.094722	1.385582	-0.399095	0.2537914	25.8275
k	2.999365	3.796275	-4.904859	-.2852032	203.6035
y	1.881667	2.381613	-1.685987	0.1789624	41.3045

The last part of the SIMLIN procedure output is a table of statistics of fit for the simulation, as shown in Output 31.1.9.

Output 31.1.9 SIMLIN Procedure Output — Simulation Statistics

Fit Statistics

Variable	N	Mean Error	Mean Pct Error	Mean Abs Error	Mean Abs Pct Error	RMS Error	RMS Pct Error	Label
c	21	0.1367	-0.3827	3.5011	6.69769	4.3155	8.1701	consumption
p	21	0.1422	-4.0671	2.9355	19.61400	3.4257	26.0265	profits
w	21	0.1282	-0.8939	3.1247	8.92110	4.0930	11.4709	private wage bill
i	21	0.1337	105.8529	2.4983	127.13736	2.9980	252.3497	investment
x	21	0.2704	-0.9553	5.9622	10.40057	7.1881	12.5653	private production
wsum	21	0.1282	-0.6669	3.1247	7.88988	4.0930	10.1724	total wage bill
k	21	-0.1424	-0.1506	3.8879	1.90614	5.0036	2.4209	capital stock
y	21	0.2704	-1.3476	5.9622	11.74177	7.1881	14.2214	national income

The OUTEST= output data set contains all the observations read from the EST= data set, and in addition contains observations for the reduced form and multiplier matrices. The following statements produce a partial listing of the OUTEST= data set, as shown in Output 31.1.10:

```
proc print data=b;
   where _type_ = 'REDUCED' | _type_ = 'IMULT1';
run;
```

Output 31.1.10 Partial Listing of OUTEST= Data Set

Simulation of Klein's Model I using SIMLIN

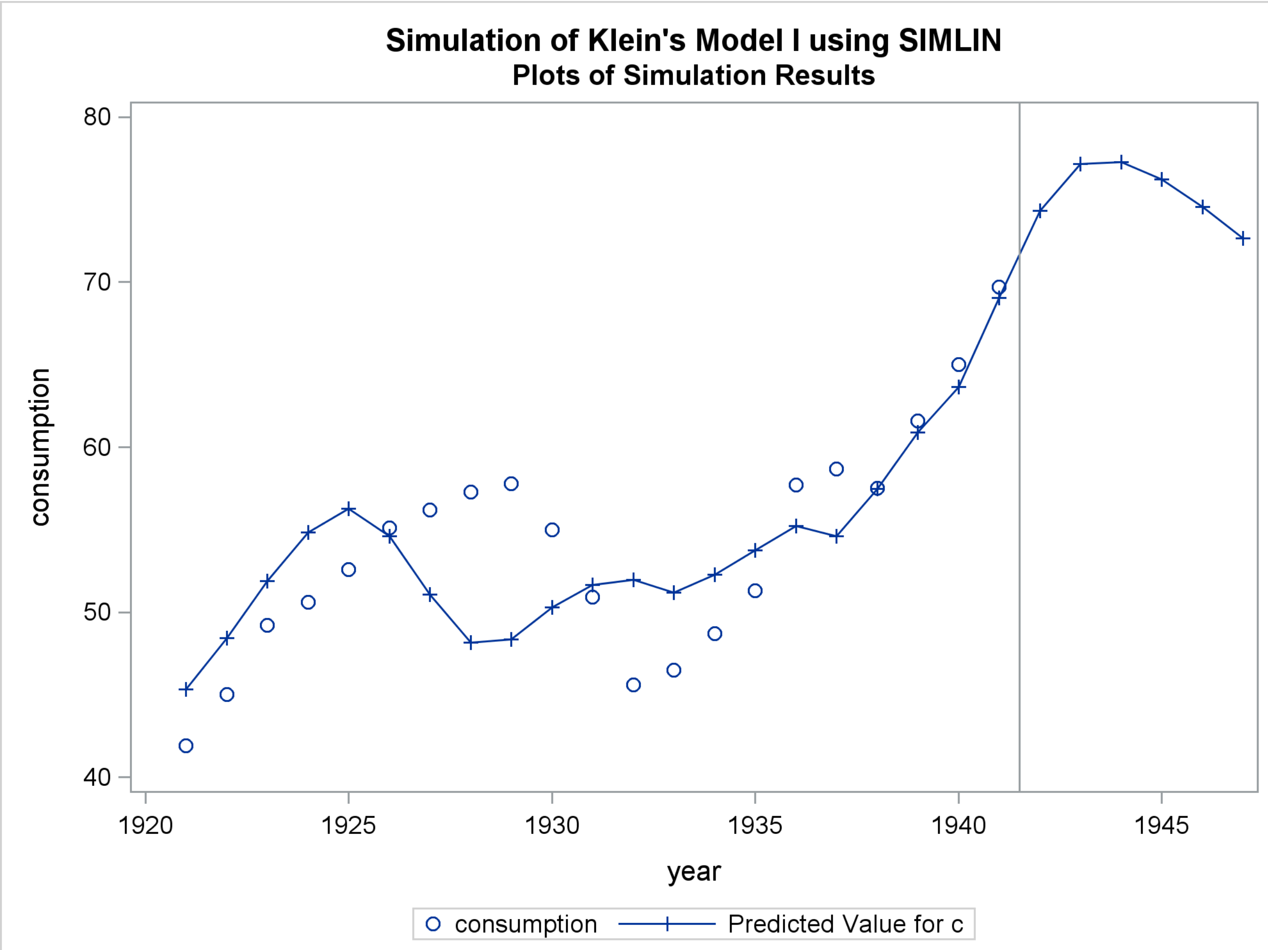
Obs	_TYPE_	_DEPVAR_	_MODEL_	_SIGMA_	c	p	w	i	x	wsum	k
9	REDUCED	c		.	1.63465	0.63465	1.09566	0.63465	0	0.19585	0
10	REDUCED	p		.	0.97236	0.97236	-0.34048	0.97236	0	1.10872	0
11	REDUCED	w		.	0.64957	0.64957	1.44059	0.64957	0	0.07263	0
12	REDUCED	i		.	-0.01272	0.98728	0.00445	-0.01272	0	-0.01450	0
13	REDUCED	x		.	1.62194	1.62194	1.10011	1.62194	0	0.18135	0
14	REDUCED	wsum		.	0.64957	0.64957	1.44059	0.64957	0	0.07263	0
15	REDUCED	k		.	-0.01272	0.98728	0.00445	-0.01272	0	-0.01450	1
16	REDUCED	y		.	1.62194	1.62194	1.10011	0.62194	1	0.18135	0
17	IMULT1	c		.	.	.	.	.	.	.	.
18	IMULT1	p		.	.	.	.	.	.	.	.
19	IMULT1	w		.	.	.	.	.	.	.	.
20	IMULT1	i		.	.	.	.	.	.	.	.
21	IMULT1	x		.	.	.	.	.	.	.	.
22	IMULT1	wsum		.	.	.	.	.	.	.	.
23	IMULT1	k		.	.	.	.	.	.	.	.
24	IMULT1	y		.	.	.	.	.	.	.	.

Obs	y	klag	plag	xlag	wp	g	t	yr	Intercept
9	1.29151	-0.12366	0.74631	0.19863	1.29151	0.63465	-0.19585	0.16399	46.7273
10	0.76825	-0.18946	0.89347	-0.06173	0.76825	0.97236	-1.10872	-0.05096	42.7736
11	0.51321	-0.12657	0.59687	0.26117	0.51321	0.64957	-0.07263	0.21562	31.5721
12	-0.01005	-0.19237	0.74404	0.00081	-0.01005	-0.01272	0.01450	0.00067	27.6184
13	1.28146	-0.31603	1.49034	0.19944	1.28146	1.62194	-0.18135	0.16466	74.3457
14	1.51321	-0.12657	0.59687	0.26117	1.51321	0.64957	-0.07263	0.21562	31.5721
15	-0.01005	0.80763	0.74404	0.00081	-0.01005	-0.01272	0.01450	0.00067	27.6184
16	1.28146	-0.31603	1.49034	0.19944	1.28146	1.62194	-1.18135	0.16466	74.3457
17	.	.	.	.	0.82913	1.04942	-0.86526	-0.00541	43.2744
18	.	.	.	.	0.60921	0.77108	-0.98217	-0.05582	28.3955
19	.	.	.	.	0.79449	1.00558	-0.71096	0.01250	41.4512
20	.	.	.	.	0.57457	0.72723	-0.82787	-0.03791	26.5723
21	.	.	.	.	1.40370	1.77666	-1.69313	-0.04332	69.8467
22	.	.	.	.	0.79449	1.00558	-0.71096	0.01250	41.4512
23	.	.	.	.	0.56452	0.71451	-0.81337	-0.03725	54.1907
24	.	.	.	.	1.40370	1.77666	-1.69313	-0.04332	69.8467

The actual and predicted values for the variable C are plotted in Output 31.1.11.

```
title2 'Plots of Simulation Results';
proc sgplot data=c;
   scatter x=year y=c;
   series x=year y=chat / markers markerattrs=(symbol=plus);
   refline 1941.5 / axis=x;
run;
```

Output 31.1.11 Plot of Actual and Predicted Consumption



Example 31.2: Multipliers for a Third-Order System

This example shows how to fit and simulate a single-equation dynamic model with third-order lags. It then shows how to convert the third-order equation into a three-equation system with only first-order lags, so that the SIMLIN procedure can compute multipliers. (For more information, see the section "Multipliers for Higher-Order Lags" on page 2327.)

The input data set TEST is created from simulated data. A partial listing of the data set TEST produced by PROC PRINT is shown in Output 31.2.1.

Output 31.2.1 Partial Listing of Input Data Set

Simulate Equation with Third-Order Lags
Listing of Simulated Input Data

Obs	y	ylag1	ylag2	ylag3	x	n
1	8.2369	8.5191	6.9491	7.8800	-1.2593	1
2	8.6285	8.2369	8.5191	6.9491	-1.6805	2
3	10.2223	8.6285	8.2369	8.5191	-1.9844	3
4	10.1372	10.2223	8.6285	8.2369	-1.7855	4
5	10.0360	10.1372	10.2223	8.6285	-1.8092	5
6	10.3560	10.0360	10.1372	10.2223	-1.3921	6
7	11.4835	10.3560	10.0360	10.1372	-2.0987	7
8	10.8508	11.4835	10.3560	10.0360	-1.8788	8
9	11.2684	10.8508	11.4835	10.3560	-1.7154	9
10	12.6310	11.2684	10.8508	11.4835	-1.8418	10

The REG procedure processes the input data and writes the parameter estimates to the OUTEST= data set A.

```
title2 'Estimated Parameters';
proc reg data=test outest=a;
   model y=ylag3 x;
run;

title2 'Listing of OUTEST= Data Set';
proc print data=a;
run;
```

Output 31.2.2 shows the printed output produced by the REG procedure, and Output 31.2.3 displays the OUTEST= data set A that is produced.

Output 31.2.2 Estimates and Fit Information from PROC REG

Simulate Equation with Third-Order Lags
Estimated Parameters

The REG Procedure
Model: MODEL1
Dependent Variable: y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	173.98377	86.99189	1691.98	<.0001
Error	27	1.38818	0.05141		
Corrected Total	29	175.37196			

Root MSE	0.22675	R-Square	0.9921
Dependent Mean	13.05234	Adj R-Sq	0.9915
Coeff Var	1.73721		

Output 31.2.2 *continued*

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.14239	0.23657	0.60	0.5523
ylag3	1	0.77121	0.01723	44.77	<.0001
x	1	-1.77668	0.10843	-16.39	<.0001

Output 31.2.3 The OUTEST= Data Set Created by PROC REG

Simulate Equation with Third-Order Lags
Listing of OUTEST= Data Set

Obs	_MODEL_	_TYPE_	_DEPVAR_	_RMSE_	Intercept	ylag3	x	y
1	MODEL1	PARMS	y	0.22675	0.14239	0.77121	-1.77668	-1

The SIMLIN procedure processes the TEST data set using the estimates from PROC REG. The following statements perform the simulation and write the results to the OUT= data set OUT2:

```
title2 'Simulation of Equation';
proc simlin est=a data=test nored;
   endogenous y;
   exogenous  x;
   lagged ylag3 y 3;
   id n;
   output out=out1 predicted=yhat residual=yresid;
run;
```

The printed output from the SIMLIN procedure is shown in Output 31.2.4.

Output 31.2.4 Output Produced by PROC SIMLIN

Simulate Equation with Third-Order Lags
Simulation of Equation

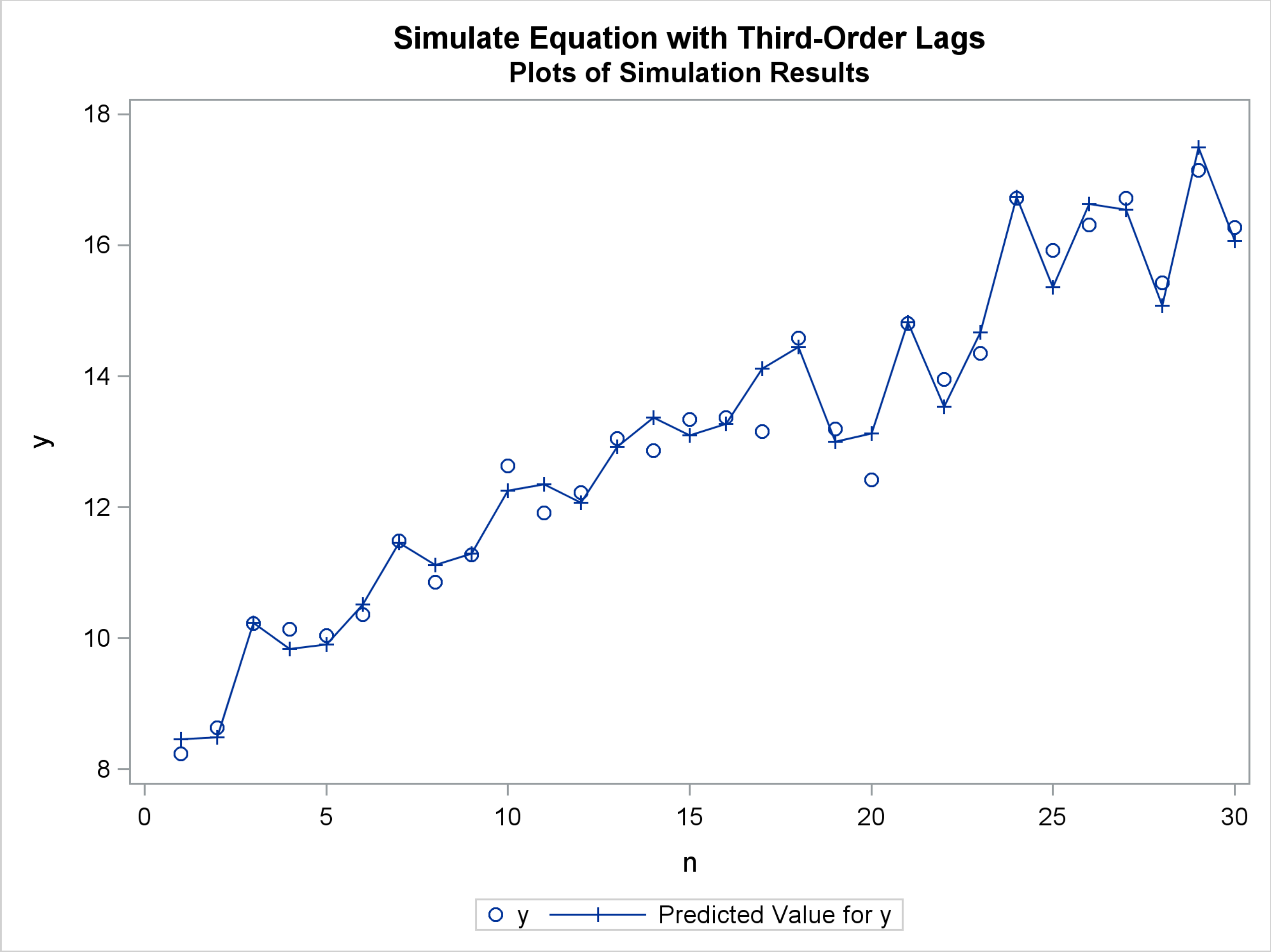
The SIMLIN Procedure

Fit Statistics							
Variable	N	Mean Error	Mean Pct Error	Mean Abs Error	Mean Abs Pct Error	RMS Error	RMS Pct Error
y	30	-0.0233	-0.2268	0.2662	2.05684	0.3408	2.6159

The following statements plot the actual and predicted values, as shown in Output 31.2.5:

```
title2 'Plots of Simulation Results';
proc sgplot data=out1;
   scatter x=n y=y;
   series x=n y=yhat / markers markerattrs=(symbol=plus);
run;
```

Output 31.2.5 Plot of Predicted and Actual Values



Next, the input data set TEST is modified by creating two new variables, YLAG1X and YLAG2X, that are equal to YLAG1 and YLAG2. These variables are used in the SYSLIN procedure. (The estimates produced by PROC SYSLIN are the same as before and are not shown.) A listing of the OUTEST= data set B created by PROC SYSLIN is shown in Output 31.2.6.

```
data test2;
   set test;
   ylag1x=ylag1;
   ylag2x=ylag2;
run;

title2 'Estimation of parameters and definition of identities';
proc syslin data=test2 outest=b;
   endogenous y ylag1x ylag2x;
   model y=ylag3 x;
   identity ylag1x=ylag1;
   identity ylag2x=ylag2;
run;

title2 'Listing of OUTEST= data set from PROC SYSLIN';
```

```
proc print data=b;
run;
```

Output 31.2.6 Listing of OUTEST= Data Set Created from PROC SYSLIN

Simulate Equation with Third-Order Lags
Listing of OUTEST= data set from PROC SYSLIN

Obs	_TYPE_	_STATUS_	_MODEL_	_DEPVAR_	_SIGMA_	Intercept	ylag3	x	ylag1	ylag2	y	ylag1x	ylag2x
1	OLS	0 Converged	y	y	0.22675	0.14239	0.77121	-1.77668	.	.	-1	.	.
2	IDENTITY	0 Converged		ylag1x	.	0.00000	.	.	1	.	.	-1	.
3	IDENTITY	0 Converged		ylag2x	.	0.00000	.	.	.	1	.	.	-1

The SIMLIN procedure is used to compute the reduced form and multipliers. The OUTEST= data set B from PROC SYSLIN is used as the EST= data set for the SIMLIN procedure. The following statements perform the multiplier analysis:

```
title2 'Simulation of transformed first-order equation system';

proc simlin est=b data=test2 total interim=2;
   endogenous y ylag1x ylag2x;
   exogenous  x;
   lagged  ylag1 y 1  ylag2 ylag1x 1  ylag3 ylag2x 1;
   id n;
   output out=out2 predicted=yhat residual=yresid;
run;
```

Output 31.2.7 shows the interim 2 and total multipliers printed by the SIMLIN procedure.

Output 31.2.7 Interim 2 and Total Multipliers

Simulate Equation with Third-Order Lags
Simulation of transformed first-order equation system

The SIMLIN Procedure

Interim Multipliers for Interim 2

Variable	x	Intercept
y	0.000000	0.0000000
ylag1x	0.000000	0.0000000
ylag2x	-1.776682	0.1423865

Total Multipliers

Variable	x	Intercept
y	-7.765556	0.6223455
ylag1x	-7.765556	0.6223455
ylag2x	-7.765556	0.6223455

References

Maddala, G. S. (1977). *Econometrics*. New York: McGraw-Hill.

Pindyck, R. S., and Rubinfeld, D. L. (1991). *Econometric Models and Economic Forecasts*. 3rd ed. New York: McGraw-Hill.

Theil, H. (1971). *Principles of Econometrics*. New York: John Wiley & Sons.

Subject Index

Syntax Index