Simulation Models for Business Planning and Economic Forecasting

Donald Erdman, SAS Institute Inc., Cary, NC

ABSTRACT

Simulation models are useful in many diverse fields. This paper illustrates the use of the SAS/ETS® MODEL procedure with examples of a business planning model and a macro econometric forecasting model. In addition, several new features of the MODEL procedure are introduced. Topics discussed include: nonlinear and simultaneous relationships; estimation of unknown model parameters; dynamic processes; goal-seeking solutions; and Monte Carlo simulation to construct confidence bounds.

INTRODUCTION

The most sophisticated approach to forecasting or to predicting the effect of policy changes is to build a full scale simulation model. Simulation models can be developed for any complex system. When business or economic data are involved, the model is often of a company or of a whole economic system. Simulation models of a company are often called financial planning models.

The MODEL procedure allows you to specify, estimate, and solve large nonlinear models. A model can be made up of many model equations, which are specified using programming statements similar to those available in the SAS® DATA step. The system of nonlinear equations can be simultaneous and can contain dynamic or lagged relationships. You can also include time series processes like ARMA errors in model equations.

In contrast to the NLIN procedure in the SAS/STAT® product, the MODEL procedure analytically computes any necessary derivatives required for estimation or simulation and handles systems of equations.

This paper introduces the MODEL procedure through examples of business planning models and macro econometric models and provides an overview of the estimation and simulation of nonlinear systems of equations.

A BUSINESS PLANNING MODEL

Assume that you are the new business planner for the J.D. Burnes Toaster company. What is usually of most interest to a company is the profit margin. The profit for the toaster company would be the income from sales minus the cost of production.

The gross monthly profits in dollars can be written as

$$profits = tsales \times \mathbf{S} - (tsales \times \mathbf{V} + \mathbf{F})$$

where $tsales$ is the quantity sold, S is the unit price, V is the variable cost per unit, and F is the fixed cost in dollars. Examples of fixed costs are wages, insurance, advertising, and interest payments. For the fictitious toaster company,

$$\mathbf{F} = wages + insurance + ad\_cost$$

To complete the *profits* model, a model of toaster sales (*tsales*) must be determined as well as a model of the variable cost function (*V*).

Toaster Sales

The initial task is to develop a model for monthly toaster sales. First plot the data.

```
title 'Monthly Toaster Sales';
proc gplot data=toast;
   axis1 order=('01jan89'd to
               '01apr93'd by year);
   plot tsales*date / haxis=axis1;
   symbol1 i=join ;
run;
```
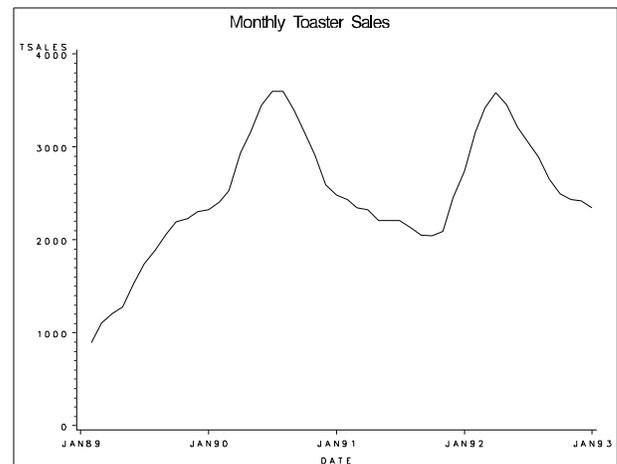

Figure 1. Monthly Toaster Sales

From the plot you determine that toaster sales might be correlated with toaster sales from a previous month. So you write down your first equation as

$$tsales_t = \alpha_1 \, tsales_{t-1} + \alpha_2 + \epsilon_t$$

Using the MODEL procedure, this model is fit to the data by the following statements:

```
proc model data=toast;

   tsales = a1 * lag(tsales) + a2;

   fit tsales;
run;
```

The results shown in Figure 2 indicate that the parameter estimates A1 and A2 are significant, however there can be other variables that further describe toaster sales.

```
                    MODEL Procedure
                    OLS Estimation

            Nonlinear OLS Summary of Residual Errors

                 DF    DF
Equation Model Error       SSE       MSE  Root MSE R-Square Adj R-Sq

TSALES     2    47    1377000  29297.88   171.17   0.9253   0.9238


              Nonlinear OLS Parameter Estimates

                            Approx.      'T'    Approx.
      Parameter   Estimate  Std Err     Ratio  Prob>|T|

      A1          0.902792  0.03740     24.14   0.0001
      A2        268.300974  96.01155     2.79   0.0075


        Number of Observations      Statistics for System
        Used             49         Objective       28102
        Missing           0         Objective*N   1377000
```

Figure 2. First Model Estimation

From your previous experience as a financial planner, you know that sales are usually related to the price of the item. You assume this relationship is nonlinear. You could expect that toaster sales are affected by the price of bread. If bread is too expensive, no one will want a toaster. Because many people give toasters as wedding presents you also suspect that the toaster sales are dependent on the number of weddings. Also the state of the economy could have an effect on toaster sales.

The new toaster sales model is

$$tsales_t = \alpha_1\,tsales_{t-1} + \alpha_2 tprice^{\alpha_3} + \alpha_4\,nwed_t + \alpha_5\,gnp_t + \alpha_6\,bprice_t + \alpha_7 + \epsilon_t$$

The gross national product (GNP) is used as a measure of the state of the economy. This model is estimated using the following SAS code:

```
proc model data=toast;
    parms  a2 -15 a3 1.5 ;

    tsales = a1 * lag(tsales) +
        a2 * tprice**a3 + a4 * nwed +
        a5 * gnp + a6 * bprice + a7;

    fit tsales /  out=resid outresid;
    id month;
run;
```

The results for the new model are shown in Figure 3.

```
                    MODEL Procedure
                    OLS Estimation

            Nonlinear OLS Summary of Residual Errors

                 DF    DF
Equation Model Error       SSE       MSE  Root MSE R-Square Adj R-Sq

TSALES     7    42    340968   8118.30   90.10159  0.9815   0.9789


              Nonlinear OLS Parameter Estimates

                            Approx.      'T'    Approx.
      Parameter   Estimate  Std Err     Ratio  Prob>|T|

      A2        -59.126320  83.05319    -0.71   0.4805
      A3          1.014765   0.47531     2.13   0.0386
      A1          1.004923   0.02324    43.24   0.0001
      A4         18.598140   3.31132     5.62   0.0001
      A5         14.697573  11.67527     1.26   0.2150
      A6        395.046466   1386.1      0.28   0.7770
      A7       -126.938741  276.62830   -0.46   0.6487


        Number of Observations      Statistics for System
        Used             49         Objective        6959
        Missing           0         Objective*N    340968
```

Figure 3. Toaster Estimation Results

The price of bread (BPRICE) and the intercept term (A7) are dropped from the estimation because these parameters estimates are not significant. The new estimation results are shown in Figure 4. A plot of the residuals is shown in Figure 5.

```
proc gplot data=resid;
    plot tsales*date / vref=0
        haxis=axis1;
    title "Residual";
    symbol1 v=plus;
    run;
```

```
                    MODEL Procedure
                    OLS Estimation

            Nonlinear OLS Summary of Residual Errors

                 DF    DF
Equation Model Error       SSE       MSE  Root MSE R-Square Adj R-Sq

TSALES     5    44    342476   7783.55   88.22442  0.9814   0.9797


              Nonlinear OLS Parameter Estimates

                            Approx.      'T'    Approx.
      Parameter   Estimate  Std Err     Ratio  Prob>|T|

      A2        -93.625585  54.06347    -1.73   0.0903
      A3          0.862519   0.18872     4.57   0.0001
      A1          1.001615   0.02079    48.18   0.0001
      A4         18.327620   2.98386     6.14   0.0001
      A5         14.771541  11.39839     1.30   0.2018


        Number of Observations      Statistics for System
        Used             49         Objective        6989
        Missing           0         Objective*N    342476
```
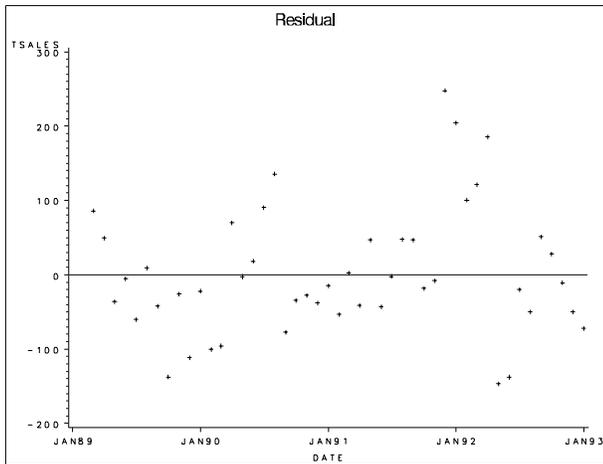
Figure 4. New Toaster Estimation Results

Figure 5. Residuals of Toaster Estimation

Programming Statements

The head of the marketing department reminds you that toaster sales should have something to do with the advertising the company purchases. They point out two time periods in which they ran the "All Toasters are Alike So Why Not Buy Ours" advertisement and the "No More Wimpy Bread" advertisement on television.

One way to test for the significance of advertising is to add a conditional parameter (*advert*) that represents the change in toaster sales due to advertising. Two conditional statements can be used to add *advert* to *tsales* during the two advertising periods. So the model will be

$$tsales_t = \alpha_1\ tsales_{t-1} + \alpha_2 tprice^{\alpha_3} + \alpha_4\ nwed_t + \alpha_5\ gnp_t + \epsilon_t$$

when no advertising has occurred and

$$tsales_t = \alpha_1\ tsales_{t-1} + \alpha_2 tprice^{\alpha_3} + \alpha_4\ nwed_t + \alpha_5\ gnp_t + advert + \epsilon_t$$

during the two advertising periods. Note that IF statements are valid programming statements in the MODEL procedure. The following SAS statements estimate a new toaster sales model:

```
proc model data=toast outmodel=tmodel;
   parms  a1 a2 -15 a3 1.5 a4 a5 advert;

   ads = 0;
   if month >= 15 and month < 20
      then ads = advert;
   if month >= 35 and month < 40
      then ads = advert;

   tsales = a1 * lag(tsales) + a2 * tprice**a
            + a4 * nwed + a5 * gnp + ads;

   fit tsales;
run;
```

The OUTMODEL= option is used to write a copy of the model program and estimated parameter values to a catalog.

The new estimation results shown in Figure 6 indicate that the advertisements paid off in a positive way (advert = 271.7). The added parameter also improved the fit of the model.

```
               MODEL Procedure
               OLS Estimation

        Nonlinear OLS Summary of Residual Errors

          DF    DF
Equation Model Error    SSE      MSE  Root MSE R-Square Adj R-Sq

TSALES     6    43    24123   561.00  23.68535  0.9987   0.9985


            Nonlinear OLS Parameter Estimates

                        Approx.     'T'    Approx.
  Parameter  Estimate   Std Err   Ratio   Prob>|T|

  A1         0.902728  0.0066515  135.72   0.0001
  A2       -11.503566   5.79116    -1.99   0.0534
  A3         1.375796   0.17925     7.68   0.0001
  A4        13.674744   0.67089    20.38   0.0001
  A5         9.489171   3.05635     3.10   0.0034
  ADVERT   271.733096  11.36230    23.92   0.0001


     Number of Observations      Statistics for System
     Used              49        Objective    492.3023
     Missing            0        Objective*N     24123
```

Figure 6. Model with Advertising Effect

N-Ahead Simulation

Suppose you want to regularly forecast 3 months ahead and produce a new forecast each month as more data becomes available. It would be good to be able to check the model's forecasting capabilities for that length of time. One way to do this is to forecast over a historical period so that you can compare your forecast with the actual data. In general, *n*-period-ahead simulation allows you to test how well you would have done over time had you been using your model in the past. *n*-period-ahead simulation is sometimes called historical or *ex post* simulation when data used to estimate the model is used as the actual data. When data excluded from the estimation is used as the actual data, this is sometimes called *ex post* forecasting.

There are fifty months in the toaster sales data set, so there are 47 periods on which a 3-months-ahead forecast can be performed (remove one observation to fill the lag). Therefore 47 separate 3-months-ahead forecasts can be performed, and the third month in each forecast can be compared to the actual value. This process is done automatically for you with the NAHEAD= option on the SOLVE statement.

The following SAS statements perform a 3-months-ahead simulation for the toaster sales model:

```
proc model data=toast model=tmodel;

   solve / nahead=3 out=tnahead outactual
           outpredict;
   id date;
run;

proc gplot data=tnahead;
   plot tsales*date=_TYPE_ / haxis=axis1;
   title "3-Ahead Simulation";
   symbol1 v=none i=join;
   symbol2 v=triangle i=join;
   run;
```
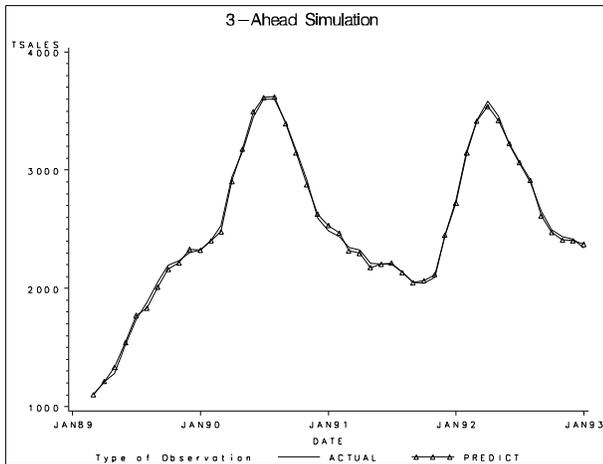
Figure 7. 3-Ahead Simulation vs Actual

The results of the 3-ahead simulation shown in Figure 7 indicate that the toaster sales model is very good at predicting toaster sales 3 months ahead based on a historical comparison.

Goal Seeking

There are several variables in this fictitious company that are under the control of the management. One important variable is the amount of money spent on advertising. If the toaster sales model is rewritten to reflect directly the increase (or decrease) in the number of toasters sold as a function of dollars spent on advertising, an informed decision could be made on how much to spend on advertising.

The variable AD_COST, containing the monthly expenditures on advertising, is added to the data set. The new model is assumed to have the following form:

$$tsales_t = \alpha_1 \, tsales_{t-1} + \alpha_2 tprice^{\alpha_3} + \alpha_4 \, nwed_t$$
$$+ \, \alpha_5 \, gnp_t + \alpha_6 \, ad\_cost_t + \epsilon_t$$

The advertising effect is normally considered nonlinear, but because the toaster company has advertised only twice, it would be difficult to get the functional form of the dependency. It enters the model linearly.

Once the new model is estimated, goal seeking questions such as "How much should be spent on advertising to get a specific profit?" can be asked. Goal seeking is done with the MODEL procedure by using a SOLVE statement. List on the SOLVE statement the variables you want to solve for and provide a data set containing values for the remaining variables.

To determine how to increase the profits of the toaster company to $223,000, use the following SAS statements:

```
data goal;
   tprice = 11.1;
   nwed = 27.45;
   gnp = 4.16;
   ad_cost = .;
   tsales = .;
   profit = 223000;
   month = 51;
output;
run;
data goal; merge toast goal;
   by month;
```

```
run;

proc model data=toast;
   parms   a1 a2 -15 a3 1.5 a4-a6;

   tsales = a1 * lag(tsales) + a2 * tprice**a
      + a4 * nwed + a5 * gnp + a6 * ad_cost;

   insurance = 14000;
   wages = 95500;
   F = wages + insurance + ad_cost;

   profit = tsales * 150.5 -
            (tsales * 23.2 + F);

   fit tsales;

run;
   solve ad_cost tsales
         satisfy=(profit tsales)
         / data=goal solveprint;
range month 51;
run;
```

The result of the goal seeking simulation is shown in Figure 8. Note that a constant cost per toaster has been assumed. You must decide whether or not to believe this model outside the domain in which it was fit. If the model is appropriate, then spending approximately $18,800 on advertising would result in a net profit near $223, 000.

```
Solution Values:
AD_COST:     18844   TSALES:      2760
```

Figure 8. Goal Seeking Output

All of the DATA step financial functions are available in the MODEL procedure. If other financial quantities such as net present value or the internal rate of return are desired over the profit value, they can be obtained easily.

Multiple Equation Systems

If a model has more than one endogenous variable, you must be careful in choosing an estimation method. If the model has endogenous regressors, then nonlinear ordinary least-squares estimation of these equations will produce biased and inconsistent parameter estimates. This is called *simultaneous equation bias*.

For example, assume that the toaster company has two products, a two-slice toaster and a four-slice toaster. Because the sales of one type of toaster affects the sales of the other, the sales model for both types of toaster should be coupled together. The new model is assumed to have the following form:

$$tsales_t = \alpha_1 \, tsales_{t-1} + \alpha_2 \, tprice^{\alpha_3} + \alpha_4 \, nwed_t + \alpha_5 \, gnp_t + \alpha_6 \, ad\_cost_t + \alpha_7 \, fs\_sales_t + \epsilon_t$$

$$fs\_sales_t = \beta_1 \, fs\_sales_{t-1} - fs\_price^{\beta_2} + \beta_3 \, gnp_t + \beta_4 \, tsales_t + \tilde{\epsilon}_t$$

The two equations must be estimated simultaneously because of the cross dependency of endogenous variables *tsales* and *fs_sales*.

Simultaneous equation bias can be avoided by using an instrumental variables method such as 2SLS and 3SLS or by using full information maximum likelihood estimation (FIML).

FIML is computationally more expensive than an instrumental variables method and assumes that the errors are normally distributed.

On the other hand, FIML does not require the specification of instruments. FIML is selected with the FIML option on the FIT statement.

```
proc model data=toast ;
   parms  a1 a2 -15 a3 1.5 a4-a7;
   parms  b1 b2 0.5 b3 b4 ;

   tsales = a1 *lag(tsales) + a2 * tprice**a3
      + a4 * nwed + a5 * gnp + a6 * ad_cost
      + a7 * fs_sales;

   fs_sales= b1*lag(fs_sales) - fs_price**b2
      + b3 * gnp + b4 * tsales;

   fit tsales fs_sales/ fiml hessian=fda;
run;
```

The estimation results are shown in Figure 9. Note that increasing sales of two-slice toasters positively (0.0968) affects four-slice toaster sales, but increasing sales of four slice toasters negatively (-0.0707) affects the two-slice toaster sales.

```
                    MODEL Procedure
                   FIML Estimation

        Nonlinear FIML Summary of Residual Errors

             DF    DF
Equation  Model  Error      SSE      MSE   Root MSE  R-Square  Adj R-Sq

TSALES       7     42     35611    847.89  29.11853   0.9981    0.9978
FS_SALES     4     45    129788   2884.17  53.70448   0.9772    0.9757


          Nonlinear FIML Parameter Estimates

                         Approx.     'T'    Approx.
   Parameter   Estimate  Std Err    Ratio   Prob>|T|

      A1       0.951236   0.01587    59.95   0.0001
      A2     -18.642667   7.98447    -2.33   0.0244
      A3       1.226497   0.14736     8.32   0.0001
      A4      13.307976   0.83458    15.95   0.0001
      A5      10.479793   3.65095     2.87   0.0064
      A6       0.036417  0.0050365    7.23   0.0001
      A7      -0.070744   0.02758    -2.57   0.0140
      B1       0.813210   0.03601    22.58   0.0001
      B2       1.245051   0.09653    12.90   0.0001
      B3      20.545335   6.66550     3.08   0.0035
      B4       0.096838   0.02410     4.02   0.0002
```

Figure 9. FIML Estimation of Toaster Sales

FIML estimation is a new feature of the MODEL procedure for SAS release 608.

Monte Carlo Simulation

How accurate is a forecast? For a linear model, the confidence bounds of the forecast can be easily computed using existing formulas. However, nonlinear models require numerically generated confidence bounds since the distribution of a forecast is unknown. The MODEL procedure provides Monte Carlo techniques, which, when used with the covariance of the parameter estimates and error covariance matrix, can produce approximate error bounds on the forecasts.

For example, confidence intervals can be created for the two product toaster sales model introduced previously. For a forecast to be accurate, accurate forecasts of the exogenous variables must be obtained. Many forecasts can be purchased, obtained for free

(for example, call several churches to determine the number of weddings planned), or produced using another procedure such as PROC FORECAST. It is better to add the equations used to generate the forecast of the exogenous variables to the model so that the effect of the errors of the forecasts of those variables are captured in the Monte Carlo forecast.

For this example, the forecasts for the exogenous variables are assumed to be obtained externally and from other departments in the company. The forecasts are added to the end of the toaster sales data set using the following DATA steps:

```
data forecast;
   input tsales fs_sales fs_price
   gnp tprice month nwed ad_cost;
cards;
.  .  52.9 5.09315 12.7948 51 25.6870  0
.  .  47.1 4.13807 11.1620 52 27.3529  0
.  .  55.6 5.03208  9.0044 53 27.0519  1000
.  .  54.0 4.48161  6.3539 54 29.1117  1000
.  .  49.4 5.29205  7.5116 55 34.3796  1000
.  .  43.0 4.18634  7.3252 56 33.0864  1000
.  .  41.8 4.71916  6.9959 57 31.7647  1000
;

run;
data forecast; merge toast forecast;
   by month;run;
run;
```

Note that the forecast includes a plan to spend $1,000 on advertising for several months.

To perform Monte Carlo simulation, you need the covariance matrix of the parameter estimates and the covariance matrix of equation errors. The following modified FIT statement saves the two covariance matrices for use in the SOLVE statement:

```
proc model data=toast;
   parms  a1 a2 -15 a3 1.5 a4-a7 ;
   parms  b1-b4 ;

   tsales = a1 * lag(tsales) + a2 * tprice**a
      + a4 * nwed + a5 * gnp + a6 * ad_cost
      + a7 * fs_sales;
   fs_sales= b1*lag(fs_sales) - fs_price**b2
           + b3 * gnp + b4 * tsales;

   fit tsales fs_sales/ fiml hessian=fda
           outest=t_est outs=s outcov;
run;
```

A SOLVE statement is added to request a Monte Carlo simulation. The RANDOM=150 option specifies that 150 forecasts be run with different perturbations of the parameters and the equation errors.

```
   solve tsales fs_sales / data=forecast
         estdata=t_est sdata=s random=150
         seed=123 out=monte;
   id month;
   range month 51;
run;
```

To generate a confidence interval plot for the forecast, use PROC UNIVARIATE to generate percentile bounds and PROC GPLOT to plot the graph. The following SAS statements produce the 5%-95% confidence bounds plot for TSALES shown in Figure 10.

```
proc sort data=monte;
   by month;
run;

proc univariate data=monte noprint;
   by month;
   var tsales;
   output out=bounds mean=mean p5=p5 p95=p95;
run;

title
  "Monte Carlo Generated Confidence Bounds";
proc gplot data=bounds;
   axis2 label=('TSales');
   plot mean*month=1 p5*month=2 p95*month=2
      / overlay vaxis=axis2 ;
   symbol1 i=join value=triangle;
   symbol2 i=join value=square l=4;
run;
```
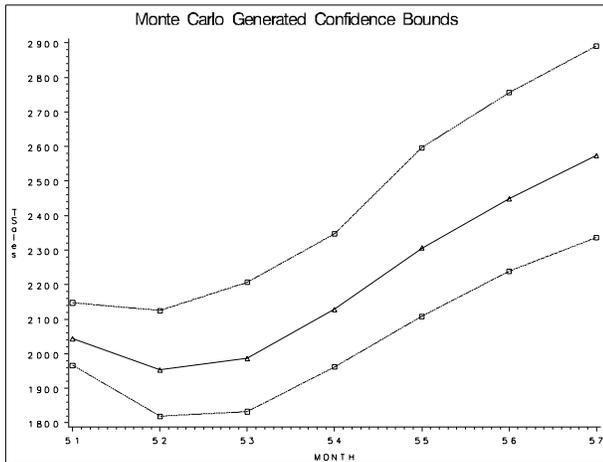

Figure 10. Monte Carlo Confidence Interval Plot

MACRO ECONOMETRIC MODELS

In the forecast of the toaster sales model it was necessary to obtain a forecast of the gross national product. Other models will require other economic indicators to be forecast. Usually, business planning is closely related with the economic systems. This section will demonstrate further features of the MODEL procedure using various macro econometric models from the literature.

Heteroscedasticity

If the data are time series, so that t indexes time, it is possible that ϵ_t depends on ϵ_{t-i} or, more generally, the ϵ_t are not identically and independently distributed. If the errors of a model system are autocorrelated, the standard error of the estimates of the parameters of the system will be biased.

Sometimes the ϵ_t's are not identically distributed because the variance of ϵ_t is not constant. This is known as *heteroscedasticity*. Heteroscedasticity can also make the standard error of the estimates of the parameters biased if the model is estimated ignoring this heteroscedasticity. Using a weighted estimation can sometimes eliminate this problem. If the proper weighting scheme is difficult to determine, generalized methods of moments (GMM) estimation can be used. GMM is a new feature in the MODEL procedure for SAS release 608.

Consider the consumption-based asset pricing model of Hansen and Singleton (Ferson 1992). The model implies that "in equilibrium the price of an asset is the expected discounted value of its future payoff, weighted by a marginal utility of consumption." The model is:

$$ \mathbf{E} \left\{ \beta \left[\frac{\mathbf{C}_{t+1}}{\mathbf{C}_t} \right]^{-\alpha} \mathbf{R}_{t+1} \mid \mathbf{Z}_t \right\} = 1 $$

where $\mathbf{Z}_t$ is a vector of instruments, $\mathbf{C}_t$ is consumption of an asset, and $\mathbf{R}_{t+1}$ is one plus a real rate of return. $\alpha > 0$ is a concavity parameter and β is a time preference parameter. The instruments for this model are a constant, four lags of the real consumption growth, and four lags of the real Treasury bill return.

The assets for this model are returns on government (GB) and corporate (CB) bonds and the smallest (D1) and the largest (D10) equity decile returns. In the following SAS code this consumption-based asset pricing model is estimated using iterated generalized method of moments (ITGMM) estimation in the MODEL procedure:

```
proc model data=harvey;

   endogenous conrat;
   exogenous gb cb d1 d10 ;
   parms beta 1.0 alpha 1.0;

      /* set up lags */
   lc1 = lag(cinst);
   lc2 = lag2(cinst);
   lc3 = lag3(cinst);
   lc4 = lag4(cinst);
   ltb1 = lag(rinst);
   ltb2 = lag2(rinst);
   ltb3 = lag3(rinst);
   ltb4 = lag4(rinst);

   eq.h1 = beta * (1+conrat)**(-alpha)
         * (1+gb) - 1 ;
   eq.h2 = beta * (1+conrat)**(-alpha)
         * (1+cb) - 1 ;
   eq.h3 = beta * (1+conrat)**(-alpha)
         * (1+d1) - 1 ;
   eq.h4 = beta * (1+conrat)**(-alpha)
         * (1+d10) - 1 ;

   fit h1-h4 / itgmm ;
   instruments lc1-lc4 ltb1-ltb4 ;

run;
```

The results of the estimation of this model are shown in Figure 11. A Parzen kernel with bandwidth parameter 2.756 was used in the estimation.

```
                    MODEL Procedure
                    ITGMM Estimation

           Nonlinear ITGMM Summary of Residual Errors

                DF    DF
Equation   Model Error     SSE       MSE      Root MSE

H1          0.5 158.5   0.43960   0.0027735    0.05266
H2          0.5 158.5   0.42362   0.0026727    0.05170
H3          0.5 158.5   3.11923   0.01968      0.14028
H4          0.5 158.5   0.99259   0.0062624    0.07914


                    MODEL Procedure
                    ITGMM Estimation

            Nonlinear ITGMM Parameter Estimates

                              Approx.     'T'    Approx.
          Parameter  Estimate  Std Err   Ratio  Prob>|T|

          BETA       1.004713 0.0036251  277.15  0.0001
          ALPHA      1.066761 0.84551      1.26  0.2089


          Number of Observations      Statistics for System
          Used             159     Objective      0.1991
          Missing            0     Objective*N   31.6519
```

Figure 11. ITGMM Estimation

General Form Equations

Many equation systems cannot be expressed in normalized form. The alternate is called *general* form, for example, the two-equation econometric model used by Bodkin and Klein to fit U.S production data for the years 1909-1949 (Bard 1974, p.133-138).

$$g_1 = c_1 \, 10^{c_2 z_4}(c_5 \, z^{-c_4} + (1 - c_5)z^{-c_4})^{-\frac{c_3}{c_4}} - z_3 = \epsilon_t$$

$$g_2 = [c_5 / (1 - c_5)](z_1 / z_2)^{(-1-c_4)} - z_5 = \tilde{\epsilon}_t$$

where z_1 is capital input, z_2 is labor input, z_3 is real output, z_4 is time in years with 1929 as year zero, and z_5 is the ratio of price of capital services to wage scale. The c_i's are the unknown parameters. z_1 and z_2 are considered endogenous variables. The model is based on the constant elasticity of substitution (CES) theory of production. This equation is written in general form.

The two equations must be estimated simultaneously because of the cross dependency of endogenous variables z_1 and z_2 and the shared parameters c_4 and c_4. The following demonstrates the use of FIML to estimate the Bodkin and Klein econometric model:

```
proc model data=bodkin;
   parms c1-c5;
   endogenous z1 z2;
   exogenous z3 z4 z5;

   eq.g1 = c1*10**(c2 * z4)*(c5*z1**(-c4) +
           (1-c5)*z2**(-c4))**(-c3/c4) - z3;
   eq.g2 = (c5/(1-c5))*(z1/z2)**-1-c4) -z5;

   fit g1 g2 / fiml ;
run;
```

The FIML estimates are shown in Figure 12.

```
                      FIML Estimation

           Nonlinear FIML Summary of Residual Errors

                DF    DF
Equation   Model Error     SSE       MSE      Root MSE

G1          4     37    0.05291  0.0014299    0.03781
G2          1     40    0.01729  0.0004322    0.02079


            Nonlinear FIML Parameter Estimates

                              Approx.     'T'    Approx.
          Parameter  Estimate  Std Err   Ratio  Prob>|T|

          C1        0.583889   0.02186   26.71   0.0001
          C2        0.00588198 0.0006716  8.76   0.0001
          C3        1.362873   0.11469   11.88   0.0001
          C4        0.474992   0.27047    1.76   0.0871
          C5        0.447049   0.05974    7.48   0.0001


          Number of Observations      Statistics for System
          Used              41     Log Likelihood  -110.7774
          Missing            3
```

Figure 12. FIML Estimation of Bodkin and Klein Econometric Model

SUMMARY

The major features of the MODEL procedure have been introduced through a mixture of generated and actual models including

- estimation of unknown model parameters
- forecasting
- Monte Carlo simulation
- N-period ahead simulation

The use of the MODEL procedure to estimate and simulate nonlinear system simulation models has been demonstrated, as well as two new features for SAS release 608, GMM and FIML estimation.

REFERENCES

Bard, Yonathan (1974), *Nonlinear Parameter Estimation*, New York: Academic Press, Inc.

Ferson, Wayne E. and Harvey, Campbell R. (1992), "Seasonality and Consumption-Based Asset Pricing," *The Journal of Finance*,47(2), 511–552.

SAS/ETS User's Guide: Release 608.