

Spatial Dependence, Nonlinear Panel Models, and More New Features in SAS/ETS® 14.1

Jan Chvosta, SAS Institute Inc.

ABSTRACT

SAS/ETS® 14.1 delivers a substantial number of new features to researchers who want to examine causality with observational data in addition to forecasting the future. This release adds count data models with spatial effects, new linear and nonlinear models for panel data, the X13 procedure for seasonal adjustment, and many more new features. This paper highlights the many enhancements to SAS/ETS software and demonstrates how these features can help your organization increase revenue and enhance productivity.

INTRODUCTION

Many new features were added to SAS/ETS in 14.1, enhancing its capability for econometric modeling. This paper briefly outlines these features and then discusses several of them in more detail.

The following features were added to SAS/ETS 14.1:

- The COUNTREG procedure adds the TEST statement, three statements that enable you to include spatial effects in a model, and more Bayesian analysis features.
- The HPCOUNTREG procedure adds the TEST statement and support for the Conway-Maxwell distribution.
- The HPPANEL procedure adds support for the between-groups estimator, the between-time-periods estimator, and pooled OLS regression.
- The MODEL procedure adds the %EQAR and %EQMA macros.
- The PANEL procedure adds more general Hausman specification tests, comparison tables for multiple models, and Hausman and Taylor (1981) and Amemiya and MaCurdy (1986) estimators.
- The QLIM procedure adds the RANDOM statement, which enables you to estimate the random-intercept models, and more Bayesian analysis features.
- The SASEFRED interface engine supports Linux X64 (LAX) hosts; real-time periods for data from the Federal Reserve Economic Data (FRED) database; logging of diagnostics in the SAS log; requests for useful information about categories, tags, groups, and releases; and the use of blanks in pathnames in options.
- The SASEQUAN interface engine supports Linux X64 (LAX) hosts, up to nine Quandl codes, logging of diagnostics in the SAS log, and the use of blanks in pathnames in options. It also ensures unique names by appending the variable number to variable names.
- The SASEXFS interface engine supports Linux X64 (LAX) hosts, logging of diagnostics in the SAS log, the UNIVERSE= option in the ExtractFormulaHistory factlet, and the use of blanks in pathnames in options.
- The SSM procedure adds the DEPLAG statement, which simplifies the specification of models that have lagged values of response variables in the observation equation.
- The VARMAX procedure supports vector error correction models in ARMA-GARCH form; linear equality and inequality constraints in vector error correction models; covariance and standard errors of the parameter estimates of the adjustment coefficient matrix; covariance matrix of innovations in vector error correction models; outputs of parameter estimates of the long-run parameters and the error correction trend parameters; Wald tests on any parameters in vector correction models except the long-run parameters and the error correction trend parameters; specification of initial values; and a new estimation method, the conditional maximum likelihood (CML) method.

- The new X13 procedure incorporates the X12 procedure in response to the US Census Bureau's inclusion of the X-12-ARIMA methodology in the X-13 ARIMA-SEATS program. PROC X13 also adds numerous options, displays additional tables, and changes the default value of the MAXITER= option to 1,500.

Count Data Modeling with Spatial Effects

The COUNTREG procedure is a powerful tool that enables you to perform count data regression in cases where the dependent variable takes nonnegative or integer values. Starting with SAS/ETS 14.1, several options are available that enable you to model count data that contain spatial effects by adding spatial lags of regressors (SLX) as right-hand-side variables. These types of models are useful because spatial effects contain information about geographical location. You can use this information to enhance your econometric modeling, because geographical regions are often related and these related regions contain additional information that you can use to enhance your analysis. For example, sales of a product in one region are often related to sales of the same product in neighboring regions; the crime rate in one county is often related to crime rates in neighboring counties; house prices in one district are often related to house prices in neighboring districts. Adding to the model these SLX compound variables that depend on the neighboring structure can increase the efficiency of your estimator and reduce bias.

The spatial-effects models in the COUNTREG procedure are currently available for the following models:

- Poisson regression
- Conway-Maxwell-Poisson regression
- negative binomial regression with quadratic (NEGBIN2) and linear (NEGBIN1) variance functions (Cameron and Trivedi 1986)
- zero-inflated Poisson (ZIP) model (Lambert 1992)
- zero-inflated Conway-Maxwell-Poisson (ZICMP) model
- zero-inflated negative binomial (ZINB) model

For example, you can estimate the Conway-Maxwell-Poisson zero-inflated model with dispersion by using the following statements:

```
proc countreg data=carsale;
  model y=x1 x2/dist=poisson;
  zeromodel y~z1 z2/link=normal;
  dispmodel y~g1 g2;
run;
```

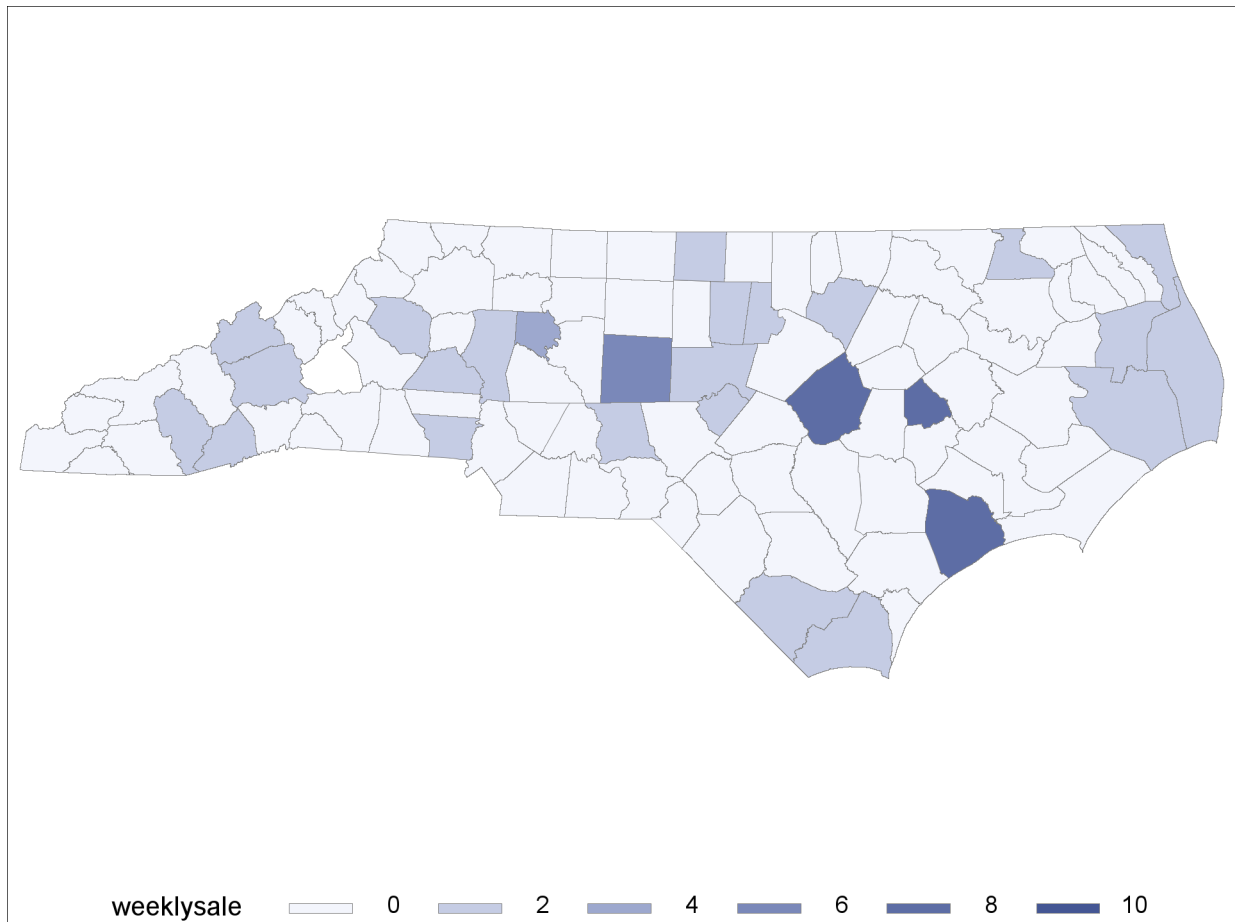
If you want to convert the Conway-Maxwell-Poisson zero-inflated model with dispersion to a model that contains spatial lags, you need to use code such as the following to provide a matrix of spatial neighbors, W , that contains the information about the spatial relationship and use SPATIALEFFECTS, SPATIALZEROEFFECTS, and SPATIALDISPEFFECTS statements to indicate what variables to use to assemble the spatial lag variables. The SPATIALZEROEFFECTS and SPATIALDISPEFFECTS statements are optional and can be used only if ZEROMODEL and DISPMODEL statements are present.

```
proc countreg data=carsale wmat=W;
  model y=x1 x2/dist=poisson;
  zeromodel y~z1 z2/link=normal;
  dispmodel y~g1 g2;
  spatialeffects x3 x4;
  spatialzeroeffects z3 z4;
  spatialdispeffects g3 g4;
  spatialid county;
run;
```

The following example illustrates the use of the SLX model by using a sample of simulated data for car sales in 100 North Carolina counties. In the **CarSale** data set, **WeeklySale** refers to the number of car sales per week in each county. The variables **x1**, **x2**, and **x3** are assumed to be predictors that affect car sales, such as the median household income, population, and size of each county. [Figure 1](#) shows a map of simulated car sales. The following statements plot the variable **WeeklySale** for each county:

```
proc gmap data=SGF2016.carsale map=SGF2016.nccounty_map;
  choro weeklysale/ coutline=gray levels=5 midpoints=(0 to 10 by 2);
  id County;
run;
```

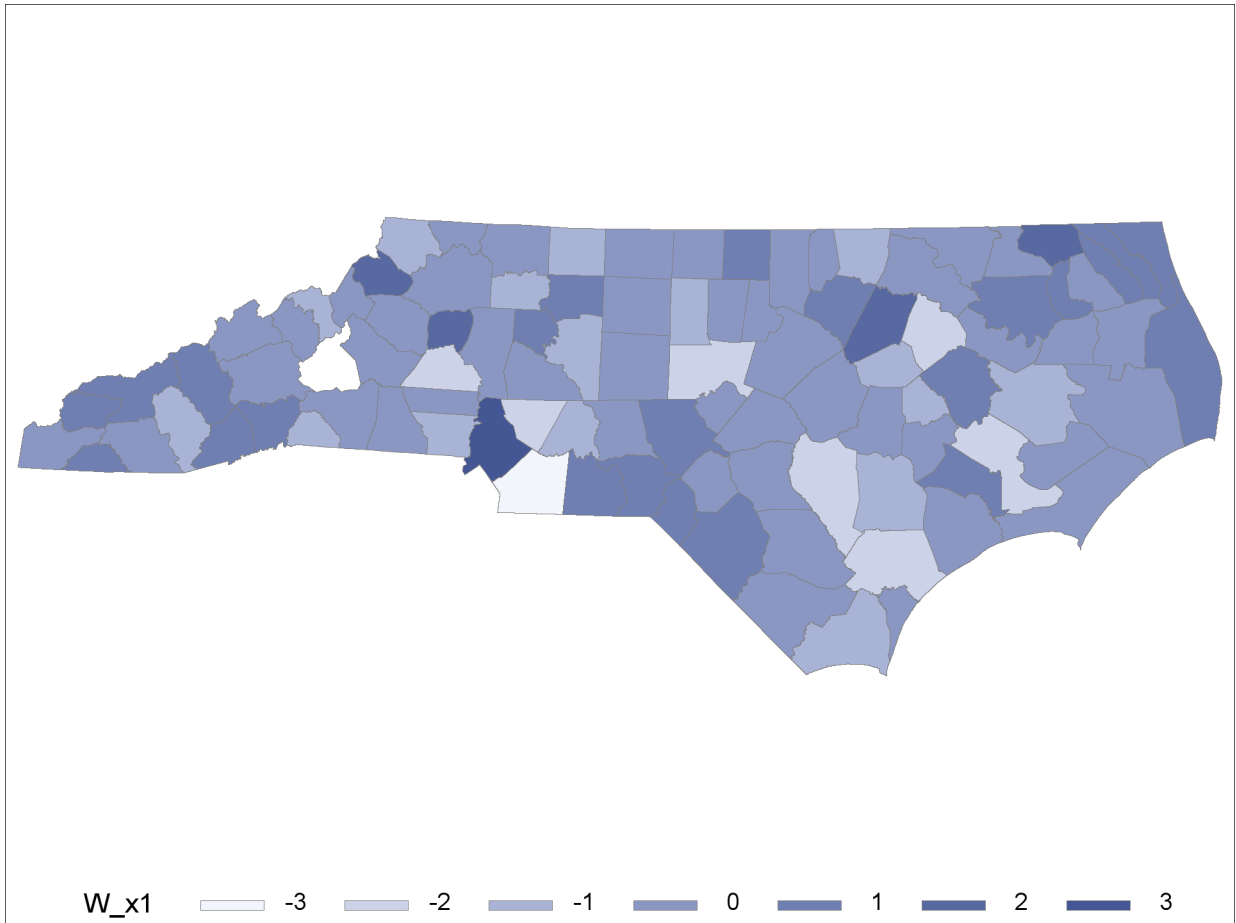
Figure 1 Simulated North Carolina Weekly Car Sales



In the SLX model, the spatial lag variables that depend on the weighting matrix, **W**, are calculated internally and added to the model. [Figure 2](#) shows a map that depicts a spatial lag variable, **W_x1**, that is based on the product of the weighting matrix, **W**, and the explanatory variable **x1**. If you carefully examine the map in [Figure 2](#), you can see a tendency for regions with high and low values of **W_x1** to be clustered together. The following statements plot the variable **W_x1** for each county:

```
proc gmap data=SGF2016.carsale map=SGF2016.nccounty_map;
  choro W_x1/ coutline=gray levels=5 midpoints=(-3 to 3 by 1);
  id County;
run;
```

Figure 2 Spatial Lag Variable W_x1



The following set of statements fits the Poisson model with spatial lag effects:

```
proc countreg data=SGF2016.carsale wmat=SGF2016.W;
  model weeklysale=x1 x2 x3/dist=poisson;
  spataleffects x1 x2 x3;
  spatialid county;
  SLX_Test: test W_x1=0,W_x2=0,W_x3=0/all;
run;
```

In **Figure 3**, W_x1, W_x2, and W_x3 denote the spatial lag of x1, x2, and x3, respectively. All six coefficients but one are significant at the 5% level.

Figure 3 Parameter Estimates for Poisson Regression with SLX Effect

The COUNTREG Procedure

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	-1.945692	0.295041	-6.59	<.0001
x1	1	1.225100	0.150402	8.15	<.0001
x2	1	-0.498799	0.171867	-2.90	0.0037
x3	1	0.465696	0.169787	2.74	0.0061
W_x1	1	1.169532	0.387681	3.02	0.0026
W_x2	1	0.810996	0.327123	2.48	0.0132
W_x3	1	-0.797009	0.424839	-1.88	0.0607

You can use the TEST statement to perform a test of joint significance of spatial lags. All three tests shown in Figure 4 indicate that the coefficients are significantly different from zero at the 1% level.

Figure 4 Test Results for $H_0: W_x1=W_x2=W_x3=0$ in Poisson Regression with SLX Effect

Test Results				
Test	Type	Statistic	Pr >	ChiSq Label
SLX_TEST	Wald	16.57183	0.0009	W_x1 = 0, W_x2 = 0, W_x3 = 0
SLX_TEST	L.R.	19.02948	0.0003	W_x1 = 0, W_x2 = 0, W_x3 = 0
SLX_TEST	L.M.	17.27478	0.0006	W_x1 = 0, W_x2 = 0, W_x3 = 0

For more information about the simulated North Carolina car sales data set and a more complete discussion of spatial econometrics modeling, see Wu and Chvosta (2016).

Hausman-Taylor and Amemiya-MaCurdy Estimators for Panel Data

Panel data are obtained when a group of subjects is observed over time and the data are recorded with the subject and time stamps. Depending on the study, subjects can be represented by many different items, including people, goods, and services, on which the data are collected at regular or irregular time intervals. The regression model for panel data can be written as

$$y_{it} = \beta_0 + \beta_x \mathbf{X}_{it} + \beta_z \mathbf{Z}_i + v_{it}$$

and

$$v_{it} = v_i + \epsilon_{it}$$

where i denotes the individual and t is any one of T time points. The regression model has two sets of explanatory variables: a set of \mathbf{X} variables that vary over time (time-varying) and a set of \mathbf{Z} variables that do not vary over time (time-invariant, or TI). The v_i are known as individual (or cross-sectional) effects; the ϵ_{it} are the observation-level regression errors; and v_{it} represents the compound error, which consists of individual and observation-level errors.

The PANEL procedure provides a variety of modeling options that you can use to fit panel data. Depending on the assumptions and theory behind your model, you usually have to choose between fitting a random-effects model and fitting a fixed-effects model. If there is not a strong theoretical foundation for fitting a fixed-effects model, you should fit the random-effects model first and use the Hausman test to see whether the regressors and errors are correlated. Under the null hypothesis, the regressors and errors are not correlated, and failing to reject the null means that the random-effects model is appropriate. However, rejecting the null signals that the fixed-effects model is more appropriate than the random-effects model.

However, it can happen that the random-effects model is rejected but you might not be interested in estimating a fixed-effects model either, because under the fixed-effects model assumptions, coefficients for the set of time-invariant variables, \mathbf{Z} , cannot be estimated. One solution to this problem is to use the Hausman-Taylor estimator. Hausman and Taylor (1981) describe an instrumental-variables approach to dealing with endogeneity due to correlated individual effects. You specify which variables you think are correlated with the individual effects, and the estimator derives a set of instruments based on the uncorrelated variables, their individual-level averages, and the deviations of the variables from these averages. The PANEL procedure fits the model by two-stage least squares (2SLS). During the first stage, the instruments are used to “predict” the correlated variables. At the second stage, estimation proceeds with a modified GLS strategy (for more information, see Baltagi 2008, sec. 7.4).

In the following example, the Hausman-Taylor estimator is fitted to grocery data on meat expenditures by 330 households that shopped regularly at a grocery chain in Raleigh, North Carolina. The data were recorded monthly for the year 2011. The household and the time period are identified by **HouseID** and **Month**. The dependent variable **Meat** records the average amount per visit spent on butcher meats. The variable **Govt** has a value of 1 if government assistance (such as food stamps or WIC) was used during that month; **Hsize** is the household size; **Rural** records whether a rural store location was visited; and **Alcohol** has a value of 1 if at least 10% of the household’s expenditures for the month were for alcoholic beverages. The variable **MealsOut** records the number of meals per month outside the household, as provided in a survey the household filled out when it applied for its customer loyalty card.

The time-invariant variables, **Hsize** and **MealsOut**, are correlated with individual cross-sectional regression errors, and the set of remaining variables in the model serve as their instruments. The following statements fit the Hausman-Taylor model:

```

proc panel data = Grocery;
  id HouseID Month;
  instruments correlated = (Govt MealsOut);
  model Meat = Govt Hsize Rural Alcohol MealsOut / htaylor;
run;

```

Figure 5 provides the results of the Hausman-Taylor estimation. The “Variance Component Estimates” table provides variance estimates for both the household effects (cross sections) and the overall errors, similar to what you get with random effects.

The “Hausman Test against Fixed Effects” table provides the results of a Hausman test that is similar to the test you would get with random effects. The Hausman test compares the Hausman-Taylor estimator to the within estimator. Think of this test as a referendum on your choice of correlated variables. The null hypothesis is that you made an adequate choice, and the results seem to indicate that the hypothesis held. The “Parameter Estimates” table is similar to the one for random-effect models, with added columns to mark the variables that are assumed to be correlated (C) and the variables that are time-invariant (TI). By stipulating the correlated variables and using the Hausman-Taylor model, you also obtain coefficients for the time-invariant variables; this was not possible in fixed-effects estimation. The estimates of coefficients for the time-varying variables (**Govt**, **Rural**, and **Alcohol**) are consistently estimated and similar to the ones that you would obtain from a fixed-effects model.

Figure 5 Hausman-Taylor Estimation

**The PANEL Procedure
Hausman and Taylor Model for Correlated Individual Effects (HTaylor)**

Dependent Variable: Meat (Meat purchases per store visit)

Model Description	
Estimation Method	HTaylor
Number of Cross Sections	330
Time Series Length	12

Fit Statistics			
SSE	89144.7983	DFE	3567
MSE	24.9915	Root MSE	4.9992
R-Square	0.1547		

Variance Component Estimates	
Variance Component for Cross Sections	97.29627
Variance Component for Error	24.97519

Hausman Test against Fixed Effects			
Coefficients	DF	m Value	Pr > m
3	1	0.76	0.3824

Parameter Estimates							
Variable	Type	DF	Estimate	Standard Error	t Value	Pr > t	Label
Intercept		1	19.12589	2.4038	7.96	<.0001	Intercept
Govt	C	1	3.583391	0.6649	5.39	<.0001	1 if used government assistance that month
Hsize	TI	1	5.17389	0.3523	14.68	<.0001	Household size
Rural		1	-1.43991	0.3573	-4.03	<.0001	1 if rural location visited at least once
Alcohol		1	2.974996	0.2004	14.85	<.0001	1 if at least 10% spent on alcohol
MealsOut	C TI	1	-1.92242	0.8090	-2.38	0.0175	Meals per week outside of household (survey)

**C: correlated with the individual effects
TI: constant (time-invariant) within cross sections**

Another new feature added to SAS/ETS in 14.1 is the Amemiya-MaCurdy estimator. In principle, it is similar to the Hausman-Taylor estimator, but it uses additional instruments to gain efficiency. You can access the Amemiya-MaCurdy estimator by using the following statements:

```
proc panel data = Grocery;
  id HouseID Month;
  instruments correlated = (Govt MealsOut);
  model Meat = Govt Hsize Rural Alcohol MealsOut / amacurdy;
run;
```

Comparison tables for multiple models were also added to the PANEL procedure in SAS/ETS 14.1. The following statements fit the one-way random-effects, one-way fixed-effects, and Hausman-Taylor estimators:

```
proc panel data = Grocery;
  id HouseID Month;
  model Meat = Govt Hsize Rural Alcohol MealsOut / ranone;
  model Meat = Govt Hsize Rural Alcohol MealsOut / fixone;
  instruments correlated = (Govt MealsOut);
  model Meat = Govt Hsize Rural Alcohol MealsOut / htaylor;
compare;
run;
```

Figure 6 shows the comparison table of parameter estimates and their standard errors for all three fitted models.

Figure 6 Parameter Estimate and Standard Error Comparison

**The PANEL Procedure
Model Comparison**

Dependent Variable: Meat (Meat purchases per store visit)

Comparison of Model Parameter Estimates				
Variable		Model 1	Model 2	Model 3
		RanOne	FixOne	HTaylor
Intercept	Estimate	20.506060	53.894415	19.125895
	Std Err	2.332669	1.649992	2.403772
Govt	Estimate	5.050562	3.591205	3.583391
	Std Err	0.598942	0.665025	0.664876
Hsize	Estimate	5.145648	0	5.173890
	Std Err	0.477447	.	0.352344
Rural	Estimate	-1.410680	-1.454439	-1.439905
	Std Err	0.344892	0.357766	0.357340
Alcohol	Estimate	2.982397	2.992035	2.974996
	Std Err	0.196014	0.201343	0.200391
MealsOut	Estimate	-2.827608	0	-1.922421
	Std Err	0.384842	.	0.808967

For more information about the grocery data set and a more complete discussion of panel data modeling, including the Hausman-Taylor estimator, see Gutierrez and Sanford (2015).

SASEFRED and SASEQUAN Interface Engines

You can access Federal Reserve Economic Data (FRED) and Quandl data by using the SAS/ETS access engine utilities. These data, which contain thousands of economic series, are a great source of information that you can use to enhance your economic analyses. The following features have been added to the SASEFRED and SASEQUAN interface engines in SAS/ETS 14.1:

- Linux X64 (LAX) host support is available.

- The PROXY= option specifies a proxy server and port number to use if the connection times out without a proxy server.
- The URL= option requests information about categories, tags, groups, and releases in the SASEFRED engine.
- The RTSTART= and RTEND= options support the real-time periods for FRED data.
- Enhancements make truncated long variable names unique in the SASEQUAN engine.
- Up to nine Quandl codes are allowed in the IDLIST= option.

Qualitative and Limited Dependent Variable Models for Panel Data

The RANDOM statement in the QLIM procedure enables you to estimate the random-effects panel model in linear, censored response, truncated response, discrete choice, and stochastic frontier models. This section demonstrates the use of the random-effects panel model in an example of a Tobit model where the dependent variable is censored at zero from below.

In the case of a random-effects censored response model (Tobit), y_{it} is specified as

$$y_{it}^* = \mathbf{x}'_{it} \boldsymbol{\beta} + \epsilon_{it}, \quad t = 1, \dots, T_i, \quad i = 1, \dots, N$$

$$y_{it} = \begin{cases} y_{it}^* & \text{if } y_{it}^* > 0 \\ 0 & \text{if } y_{it}^* \leq 0 \end{cases}$$

where

$$\epsilon_{it} = \mu_i + v_{it}$$

$$v_{it} | \mathbf{x}_i, \mu_i \sim N(0, \sigma^2)$$

$$\mu_i | \mathbf{x}_i \sim N(0, \sigma_\mu^2)$$

where \mathbf{x}_i contains \mathbf{x}_{it} for all t and μ_i is the unobserved error component related to the cross section i .

Because the error component, μ_i , stays constant within each cross section i , it is easy to see that the overall error, ϵ_{it} , violates the independence of errors assumption necessary for ordinary least squares (OLS) regression to produce valid estimates. If you use OLS to estimate the model, ignoring the violation, the dependence of errors will lead to inefficient parameter estimates. Depending on the magnitude of the two error components, these problems can be quite severe and can lead to invalid conclusions. In the following example, the data were simulated in such a way that the overall error, ϵ_{it} , is dominated by the cross-sectional error, μ_i . You can estimate the random-effects panel model by using the following statements:

```
proc qlim data=testdata;
  model y11 = x1 x2 x3 x4 / censored(lb=0);
  random int / subject=id method=halt (hdraw=100 );
run;
```

You estimate the random-intercept Tobit model by using a simulation technique based on 100 Halton draws. The SUBJECT= option indicates the cross-sectional ID variable. [Figure 7](#) shows the results of the estimation.

Figure 7 Random-Effects Panel Data Model
The QLIM Procedure

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	0.863507	0.008039	107.41	<.0001
x1	1	0.975936	0.041799	23.35	<.0001
x2	1	1.031182	0.041676	24.74	<.0001
x3	1	0.966133	0.042181	22.90	<.0001
x4	1	1.001329	0.041286	24.25	<.0001
_Sigma	1	0.107424	0.001455	73.84	<.0001
_Sigma_mu	1	4.591846	0.010013	458.57	<.0001

All parameters are significantly different from zero at the 1% level. In order to demonstrate the violation of the independence of errors assumption, the Tobit model is also fitted by OLS, as follows:

```
proc qlim data=testdata;
  model y11 = x1 x2 x3 x4/ censored(lb=0);
run;
```

Figure 8 shows the results of the OLS estimation.

Figure 8 Ordinary Least Squares Regression
The QLIM Procedure

Parameter Estimates					
Parameter	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	1.062441	0.045026	23.60	<.0001
x1	1	1.353322	0.877047	1.54	0.1228
x2	1	1.478952	0.858995	1.72	0.0851
x3	1	1.188161	0.872618	1.36	0.1733
x4	1	0.952468	0.857986	1.11	0.2669
_Sigma	1	2.891829	0.039344	73.50	<.0001

The parameter estimates for **Intercept**, **x1**, **x2**, **x3**, and **x4** are not considerably different from those shown in Figure 7, but their standard errors are much larger, meaning that some parameter estimates are not significant even at the 10% level.

The VARMAX Procedure

The VARMAX procedure in SAS/ETS 14.1 was enhanced to include the vector error correction models (VECM) in ARMA-GARCH form. You can use the COINTEG statement together with the Q= option in the MODEL statement and the GARCH statement to model the cointegration relationship between multiple time series that have GARCH-type innovations. For example, the following statements fit the VECM(1)-BEKK-GARCH(1,1):

```
proc varmax data=one;
  model y1 y2 / p=1;
  cointeg rank=1;
  garch form=bekk q=1 p=1;
run;
```

The VARMAX procedure also includes the following enhancements:

- The linear equality and inequality constraints for any parameters to be estimated in vector error correction models are supported. You can use the BOUND and RESTRICT statements to study the restricted cointegrated systems.
- The covariance and standard errors of the parameter estimates of the adjustment coefficient matrix and the covariance matrix of innovations in vector error correction models are supported. The outputs of parameter estimates of the long-run parameters and the error correction trend parameters are also supported.
- You can apply the Wald tests, by using the TEST statement, to any parameters in vector correction models except the long-run parameters and the error correction trend parameters.
- You can specify initial values, by using the INITIAL statement, for any parameters to be estimated in vector error correction models. If you do so, you must specify the full-rank initial matrices for both the adjustment coefficient matrix and the long-run parameters.
- A new estimation method, the conditional maximum likelihood (CML) method, is supported. This method is especially suitable for estimating VARMAX models of large samples.
- The log likelihoods for all types of models are output. These outputs are especially useful if you need to perform likelihood ratio (LR) tests.
- Definitions have been revised for the following information criteria: Akaike's information criterion (AIC), the corrected Akaike's information criterion (AICC), the Hannan-Quinn criterion (HQC), and the Schwarz Bayesian criterion (SBC, also referred to as BIC). You can compare more types of models, including all forms of multivariate GARCH models.
- The ECTREND option is supported in the COINTEG statement. All suboptions of the ECM= option in the MODEL statement are now supported in the COINTEG statement. Starting with SAS/ETS 14.1, it is recommended that you use the COINTEG statement instead of the ECM= option to fit vector error correction models.
- A new option, the NLC option, is supported in the COINTEG statement. By using this option, you can explicitly require that the adjustment coefficient matrix and the long-run parameters both be full rank when you numerically maximize the likelihood of a vector error correction model.

The X13 Procedure

The new X13 procedure is an adaptation of the US Census Bureau X13ARIMA-SEATS methodology. It was developed by incorporating the SEATS methodology into the X-12-ARIMA seasonal adjustment program. The functionality previously available in the X12 procedure is included in the new X13 procedure. The following features are available in the X13 procedure:

- The DIFFID= and DIFFIDITER= options have been added to the AUTOMDL statement. The DIFFID= option specifies the estimation method of identifying difference orders, whereas the DIFFIDITER= option specifies the maximum iterations for exact likelihood for DIFFID=EXACTFIRST.
- The EXACT= option has been added to the ESTIMATE statement. The EXACT= option specifies the likelihood function to be used to estimate AR and MA parameters.
- The EASTERMEANS= option has been added to the REGRESSION statement. The EASTERMEANS= option specifies the method to be used to calculate the means for the Easter regression variable.
- The following tables have been added to the displayed output for automatic modeling: "Check of the Residual Ljung-Box Q Statistic," "Initial Automatic Model Selection," and "Final Checks for Identified Model."
- The following tables have been added to those available in the X11 statement: B7, B13, B17, B20, C1, D8B, and E18.

CONCLUSION

Many new features were added to SAS/ETS 14.1. In this paper, most of the features are only briefly discussed. For more information about SAS/ETS and new features in 14.1, visit the SAS Technical Support website at <https://support.sas.com/documentation>.

REFERENCES

- Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*. 4th ed. Chichester, UK: John Wiley & Sons.
- Cameron, A. C., and Trivedi, P. K. (1986). "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests." *Journal of Applied Econometrics* 1:29–53.
- Gutierrez, R. G., and Sanford, K. (2015). "Working with Panel Data: Extracting Value from Multiple Customer Observations." In *Proceedings of the SAS Global Forum 2015 Conference*. Cary, NC: SAS Institute Inc. <http://support.sas.com/resources/papers/proceedings15/SAS1755-2015.pdf>.
- Hausman, J. A., and Taylor, W. E. (1981). "Panel Data and Unobservable Individual Effects." *Econometrica* 49:1377–1398.
- Lambert, D. (1992). "Zero-Inflated Poisson Regression with an Application to Defects in Manufacturing." *Technometrics* 34:1–14.
- Wu, G., and Chvosta, J. (2016). "How Do My Neighbors Affect Me? SAS/ETS Methods for Spatial Econometric Modeling." In *Proceedings of the SAS Global Forum 2016 Conference*. Cary, NC: SAS Institute Inc. <http://support.sas.com/resources/papers/proceedings16/4440-2016.pdf>.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author:

Jan Chvosta
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
jan.chvosta@sas.com

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.