3584-2019

A Selective Overview of Panel Data with Applications in SAS®

Oral Capps, Jr., Executive Professor and Regents Professor, Co-Director of the Agribusiness, Food and, Consumer Economics Research Center, Department of Agricultural Economics, Texas A&M University

ABSTRACT

The nature and use of panel data in the economics and business literature are discussed. Various model specifications are presented including variants of the Parks model, the Da Silva model, one-way or two-way random effects models (error component models), one-way or two-way fixed effects models, and seemingly unrelated regression (SUR) models. We illustrate these respective models with a practical example to demonstrate the use of SAS in handling panel data. Additionally, emphasis is placed on the interpretation of the empirical results.

Word count: 81

Key Words: Panel data, SAS, model specifications

Author: Oral Capps, Jr., Executive Professor and Regents Professor, Co-Director of the Agribusiness, Food and, Consumer Economics Research Center, Department of Agricultural Economics, Texas A&M University

Email: ocapps@tamu.edu

Phone: (979) 845-8492

A Selective Overview of Panel Data with Applications in SAS®

Section I. Introduction

In applied econometrics, most researchers use methods of analysis developed either for cross-sectional data or time-series data. At times, practitioners have access to data not only over time but also by cross-section. This type of data set often is referred to as pooled data or panel data, describing each of a number of cross-sectional entities (for example, individuals, households, states, firms, securities, brands of products), across a sequence of time periods.

For example, analysts might want to consider the demand for gasoline on which they have data on a monthly basis as well as by state. The monthly data might run from January 2002, to December 2006. The number of possible observations for analysis is 3,000, the product of 60 time periods of monthly observations for each of the 50 states in the United States. As another illustration, analysts might want to understand the behavior of a sample of households over time in purchasing a given food item. In this case, the cross-sectional units correspond to the number of households in the sample (let's say, for the sake of round numbers, 1,000), and the time period units might correspond to four quarters of a particular year. In this instance, the total number of observations available to the analyst is 4,000.

The extant literature is replete with pooled data sets used in empirical applications. Examples are: (1) annual demand for natural gas in residential and commercial markets by state (Balestra and Nerlove, 1966); (2) demand for gasoline and diesel fuel for agricultural use by extension districts in Virginia (Capps and Havlicek, 1978); (3) monthly rates of return for a sample of securities (Dielman, Nantell, and Wright, 1980); (4) yields of pine trees on various plots (Ferguson and Leech, 1978); (5) hourly demand for electricity by individual households (Granger et al., 1979); (6) monthly consumer purchases of gasoline for different states (Mehta, Narasimham, and Swamy, 1978); (7) per capita consumption of fluid milk over ten regions of the United States (Ward and McDonald, 1986); and monthly coupon promotions for U.S. households (Ward and Davis, 1978)

Empirical applications of micro-econometrics often involve longitudinal or panel data in which cross-sectional entities are observed over time. Three notable examples of panel data are the Panel Study of Income Dynamics (PSID), the National Longitudinal Surveys of Labor Market Experience (NLS), and the Nielsen HomeScan Panel. These panel data sets rest on interviews from thousands of individuals or households over time. Baltagi (2001), Dielman (1983, 1989), and Nerlove (2003) are definitive sources of information about panel data procedures, with extensive references to sources of panel data, examples of applications of panel data, and discussions of the advantages and limitations of panel data.

Section II. Key Question

The key question in dealing with panel data is whether analysts can we live with common pooled estimates of the coefficients of the explanatory variables. If the answer to this question is yes, then single-equation models are the way to proceed with the use of PROC PANEL. If the answer to this question is no, then we can run separate regressions by cross-sectional unit, that is construct a seemingly unrelated regression (SUR) model, provided sufficient observations exist to not only estimate parameters of the model but also to conduct statistical tests of hypotheses. In this case, PROC SYSLIN or PROC MODEL are the appropriate SAS procedures. This decision is consistent with Sims' (1989) principle that a meaningful modeling effort should be, in part, tailored to the analytical purpose at hand. To illustrate, can we live with a common pooled estimate of the own-price elasticity of demand for gasoline in the United States, or do we need to consider individual own-price elasticities by state or by region?

As a starting point, analysts might run different regressions by cross-sectional unit, and test for the equality of the coefficients of common explanatory variables across the cross-sectional entities. This hypothesis test often rests on the use of F-tests. If one cannot reject the null hypothesis of the equality of these coefficients, then on the basis of statistical grounds one might pool the data, use a single-equation specification, and subsequently obtain unique estimates of the coefficients of common explanatory variables, such as a unique estimate of the own-price elasticity of demand. If however one rejects this hypothesis, then statistically speaking one might want to run separate regressions, or at least allow for different parameter estimates of the coefficients associated with the common explanatory variables by cross-sectional unit.

The advantages of pooling the data by cross-section and by time period are: (1) increasing the sample size of observations which from a purely statistical point of view increases the power of various test of hypotheses; (2) alleviating potential collinearity problems through combining variation across micro units with variation over time (Kuh and Meyer, 1957); (3) dealing with heterogeneity in the micro units; and (4) allowing more detailed analysis of dynamic adjustment through the examination of reactions over time by the cross-sectional entities. The limitations of pooling the data often rest on the use of more complex econometric procedures to deal with time-series and cross-sectional variability in the sample of observations. Aside from adding complexities to the econometric model, either in the way of model specification or model estimation, there is virtually little downside in pooling the data.

Section III. Single-Equation Model Specification Associated with the Pooling of Time-Series and Cross-Sectional Data or with the Use of Panel Data

The single-equation model specification associated with the use of panel data or pooled data is as follows:

$$Y_{it} = \beta_0 + \beta_1 X_{it,1} + \beta_2 X_{it,2} + \ldots + \beta_k X_{it,k} + \varepsilon_{it} \tag{1}$$

$$\underbrace{i = 1, 2, \ldots, N;}_{\text{cross-sections}} \quad \underbrace{t = 1, 2, \ldots, T}_{\text{time periods}}$$

Generically, Y_{it} represents the dependent (or endogenous) variable, $X_{it,1},\ldots,X_{it,k}$ correspond to the k explanatory variables for cross-section I in the time period t and ε_{it} represents the error term for cross-section i in time period t. In order to obtain the appropriate parameter estimates of the coefficients of the set of explanatory variables, it is necessary to initially stack the pooled data by cross-section. That is, without particular ordering of cross-sections, arrange the data for cross-section 1 and its associated t observations of the dependent variable and the set of explanatory variables over time, then do the same for cross-section 2 through cross-section k. The SORT procedure (PROC SORT) allows the stacking of the data into the required cross-sectional time-series format. Assuming a balanced design, wherein the same number of time periods exists in each cross-section, the sample size then is the product of the number of cross-sections (N) and the number of time periods in each cross-section (T), or simply NT. Subsequently, the parameter estimates of the coefficients of the set of explanatory variables are given as:

(2)
$$\hat{\beta} = (X^T\Omega^{-1}X)^{-1}X^T\Omega^{-1}Y$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \quad Y = \begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1T} \\ Y_{21} \\ Y_{22} \\ \vdots \\ Y_{NT} \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_{11,1} & \cdots & X_{11,k} \\ 1 & X_{12,1} & \cdots & X_{12,k} \\ \vdots & \vdots & & \vdots \\ 1 & X_{1T,1} & \cdots & X_{1T,k} \\ 1 & X_{21,1} & \cdots & X_{21,k} \\ 1 & X_{22,1} & \cdots & X_{22,k} \\ \vdots & \vdots & & \vdots \\ 1 & X_{NT,1} & \cdots & X_{NT,k} \end{bmatrix}$$

The variances of these estimated parameters are given by:

(3)
$$var(\hat{B}) = (x^T\Omega^{-1}x)^{-1}$$

The estimates of the standard errors of the parameter estimates are simply the square root of the estimates of the variances of these estimated parameters. Essentially, the parameter estimates and the estimates of the variances of these estimated parameters rest on the use of generalized least squares (GLS). The variance-covariance matrix of the error terms is given as (see Kmenta, 1986):

(4)
$$\Omega = \begin{bmatrix} E(\varepsilon_{11}^2) & E(\varepsilon_{11}\varepsilon_{12}) & \dots & E(\varepsilon_{11}\varepsilon_{1T}) & E(\varepsilon_{11}\varepsilon_{21}) & E(\varepsilon_{11}\varepsilon_{22}) & \dots & E(\varepsilon_{11}\varepsilon_{NT}) \\ E(\varepsilon_{12}\varepsilon_{11}) & E(\varepsilon_{12}^2) & \dots & E(\varepsilon_{12}\varepsilon_{1T}) & E(\varepsilon_{12}\varepsilon_{21}) & E(\varepsilon_{12}\varepsilon_{22}) & \dots & E(\varepsilon_{12}\varepsilon NT) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ E(\varepsilon_{1T}\varepsilon_{11}) & E(\varepsilon_{1T}\varepsilon_{12}) & \dots & E(\varepsilon_{1T}^2) & E(\varepsilon_{1T}\varepsilon_{21}) & E(\varepsilon_{1T}\varepsilon_{22}) & \dots & E(\varepsilon_{1T}\varepsilon_{NT}) \\ E(\varepsilon_{21}\varepsilon_{11}) & E(\varepsilon_{21}\varepsilon_{12}) & \dots & E(\varepsilon_{21}\varepsilon_{1T}) & E(\varepsilon_{21}^2) & E(\varepsilon_{21}\varepsilon_{22}) & \dots & E(\varepsilon_{21}\varepsilon_{NT}) \\ E(\varepsilon_{22}\varepsilon_{11}) & E(\varepsilon_{22}\varepsilon_{12}) & \dots & E(\varepsilon_{22}\varepsilon_{1T}) & E(\varepsilon_{22}\varepsilon_{21}) & E(\varepsilon_{22}^2) & \dots & E(\varepsilon_{22}\varepsilon_{NT}) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ E(\varepsilon_{NT}\varepsilon_{11}) & E(\varepsilon_{NT}\varepsilon_{12}) & \dots & E(\varepsilon_{NT}\varepsilon_{1T}) & E(\varepsilon_{NT}\varepsilon_{21}) & E(\varepsilon_{NT}\varepsilon_{22}) & \dots & E(\varepsilon_{NT}^2) \end{bmatrix}_{NTxNT}$$

This specification provides a general framework for the discussion of different models designed to deal with pooled cross-section and time series observations or panel data. In particular, the relationship among the error terms of various cross-sectional entities at some specific time is likely to be different from the relationship among error terms of a specific cross-sectional entity at different periods of time.

Section IV. Typical Assumptions When Dealing with the Pooling of Time Series and Cross-Sectional Data or Panel Data

As previously discussed, the key element to obtain the parameter estimates and their associated variances (and standard errors) rests on the elements of the omega matrix. In this section, we discuss the typical assumptions made regarding the elements of the variance-covariance matrix of the error terms. In this light, we discuss: (1) the pooled model; (2) the Parks model; (3) the error components model; (4) the DaSilva model; and (5) the analysis of covariance model.

Pooled Model

The pooled model is very straightforward. With this model, we simply use ordinary least squares (OLS) to estimate the parameters and the associated standard errors. This procedure is tantamount to the assumption that the omega matrix (the variance-covariance matrix of the error terms) is the identity matrix of order NT x NT. This assumption in most cases is untenable.

Parks Model

The Parks model (Parks, 1969) recognizes heteroscedasticity and mutual correlation across cross-sections and simultaneously recognizes autoregressive schemes within cross-sections. This model is applicable only for balanced designs. With the Parks model, we assume that the error terms of the respective cross-sections are heteroscedastic (unequal variances) and that the error terms within each cross-section follow a first-order autoregressive process.

(5)
$$E(\varepsilon_{it}^2) = \sigma_i^2 \quad \text{(heteroscedasticity)}$$
$$E(\varepsilon_{it}\varepsilon_{jt}) = 0 \ \ (\mathrm{i} \neq \mathrm{j}) \quad \text{(cross-sectional independence)}$$
$$\varepsilon_{\mathrm{it}} = \rho_i \varepsilon_{i,t-1} + u_{it} \quad \text{(autoregression)}$$

The parameter ρ potentially varies from one cross-sectional unit to another. With this set of assumptions, the variance-covariance matrix of error terms is given by (see Kmenta, 1986):

(6a)
$$\Omega = \begin{bmatrix} \sigma_1^2 P_1 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 P_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 P_N \end{bmatrix}$$

(6b) *where*
$$P_i = \begin{bmatrix} 1 & \rho_i & \rho_i^2 & \cdots & \rho_i^{T-1} \\ \rho_i & 1 & \rho_i & \cdots & \rho_i^{T-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 \end{bmatrix}$$

Note that each of the 0's represents a (TxT) matrix of zeros. Importantly, in this version of the Parks model, we assume that the error terms for different cross-sectional units at time t are mutually independent.

In another version of this model, we may relax this assumption of mutual independence.

(7)
$$
\begin{array}{ll}
E(\varepsilon_{it}^2) = \sigma_i & \text{(heteroscedasticity),} \\
\boxed{E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij}} & \boxed{\text{(mutual correlation),}} \\
\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + u_{it} & \text{(autoregression),} \\
u_{it} \sim N(0, \phi_{ii}), & \\
E(\varepsilon_{i,t-1} u_{jt}) = 0, & \\
E(u_{it} u_{jt}) = \phi_{ij}, & \\
E(u_{it} u_{js}) = 0 \qquad (t \neq s), & \\
i, j = 1, 2, \ldots, N \qquad t = 1, 2, \ldots, T. &
\end{array}
$$

In this case, the variance-covariance matrix of the respective error terms looks like (see Kmenta, 1986):

(8a)
$$
\Omega = \begin{bmatrix} \sigma_{11}P_{11} & \sigma_{12}P_{12} & \cdots & \sigma_{1N}P_{1N} \\ \sigma_{21}P_{21} & \sigma_{22}P_{22} & \cdots & \sigma_{2N}P_{2N} \\ \vdots & \vdots & & \vdots \\ \sigma_{N1}P_{N1} & \sigma_{N2}P_{N2} & \cdots & \sigma_{NN}P_{NN} \end{bmatrix}
$$

(8b)
$$
P_{ij} = \begin{bmatrix} 1 & \rho_j & \rho_j^2 & \cdots & \rho_j^{T-1} \\ \rho_i & 1 & \rho_j & \cdots & \rho_j^{T-2} \\ \rho_i^2 & \rho_i & 1 & \cdots & \rho_j^{T-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 \end{bmatrix}
$$

The parameter σ_{ii} and σ_{ij} ($i \neq j$) form the symmetric matrix phi described as:

(8c)
$$
phi = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ & \sigma_{22} & \cdots & \sigma_{2N} \\ & & \cdots & \sigma_{NN} \end{bmatrix} \; NxN
$$

The Parks method estimates a first-order autoregressive model with contemporaneous correlation of the error terms of the cross-sectional units. With the assumption of cross-sectional independence in in the Parks model, it is necessary to obtain N estimates of the variance parameters and N estimates of the rho parameters. However, with the assumption of mutual correlation of the cross-sectional units, it is necessary to obtain N(N+1)/2 +N estimates of the respective parameters.

Error Components Model

In the error components model, the error term is composed of three independent components – one component associated with time, the second component associated with the cross-sectional units, and the third component varying in both dimensions (Fuller and Battese, 1974; Wallace and Hussain, 1996; Wansbeek and Kapteyn, 1989; and Nerlove, 1971).[1]

Mathematically, as per Kmenta (1986),

(9)
$$\varepsilon_{it} = u_i + v_t + w_{it} \quad (\mathrm{i} = 1, 2, \ldots, \mathrm{N};\ \mathrm{t} = 1, 2, \ldots, \mathrm{T}),$$

where
$$u_i \sim N(0, \sigma_u^2),$$
$$v_t \sim N(0, \sigma_v^2),$$
$$w_{it} \sim N(0, \sigma_w^2),$$

and the components u_t, v_t, and w_{it} satisfy the following conditions:

(10)
$$E(u_i v_t) = E(u_t w_{it}) = E(v_t w_{it}) = 0,$$
$$E(u_i u_j) = 0 \ (\mathrm{i} \neq \mathrm{j}),$$
$$E(u_t v_s) = 0 \ (\mathrm{t} \neq \mathrm{s}),$$
$$E(w_{it} w_{is}) = E(w_{it} w_{jt}) = E(w_{it} w_{js}) = 0 \quad (\mathrm{i} \neq \mathrm{j}; \mathrm{t} \neq \mathrm{s}).$$

Based on this set of assumptions, ε_{it} is homoscedastic with variance given by:

(11)
$$VAR(\varepsilon_{it}) = \sigma^2 = \sigma_u^2 + \sigma_v^2 + \sigma_w^2$$

This expression corresponds to the sum of the variances of each of the three components. Further, the correlation coefficient of ε_{it} and ε_{jt} ($i \neq j$) that is, the correlation of the error terms of given cross-sectional units at a given point of time, is expressed as:

(12)
$$\sigma_v^2 / (\sigma_u^2 + \sigma_v^2 + \sigma_w^2)$$

[1] This specification also is known as a two-way random effects model. At times, analysts may employ a one-way random effects model, wherein the error term is composed of two independent components. To conform to space limitations, we discuss only the two-way random effects model.

The correlation coefficient of ε_{it} and ε_{is} ($t \neq s$) that is, the correlation of the error terms of a given cross-sectional unit at two different points of time is expressed as:

(13) $$\sigma_u^2 / (\sigma_u^2 + \sigma_v^2 + \sigma_w^2)$$

For each cross-sectional unit, the correlation of the error terms over time remains unchanged irrespective of how far apart they are at different time periods. This assumption contrasts sharply with the assumption of first-order autoregression in the Parks model, which implies that the degree of correlation declines geometrically with the time distance involved. Finally, with the error components model, the correlation coefficient between ε_{it} and ε_{js} is 0 for $i \neq j$ and $t \neq s$ (Kmenta, 1986).

With the error components model, the variance-covariance matrix of the error terms is expressed as (Kmenta, 1986):

(14a) $$\Omega = \begin{bmatrix} \sigma_u^2 A_T & \sigma_v^2 I_T & \cdots & \sigma_v^2 I_T \\ \vdots & \vdots & & \vdots \\ \sigma_v^2 I_T & \sigma_u^2 A_T & \cdots & \sigma_v^2 I_T \\ \vdots & \vdots & & \vdots \\ \sigma_v^2 I_T & \sigma_v^2 I_T & \cdots & \sigma_u^2 A_T \end{bmatrix}$$

I_T an identity matrix of order txt.

where A_T is a (TxT) matrix defined as

(14b) $$A_T = \begin{bmatrix} \sigma^2/\sigma_u^2 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & \sigma^2/\sigma_u^2 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & \sigma^2/\sigma_u^2 \end{bmatrix}$$

As such, to estimate the elements of the omega matrix, it is necessary to obtain estimates of only three parameters. We estimate elements of Ω as (Kmenta, 1986):

(15a) $$\hat{\sigma}_w^2 = \frac{1}{(N-1)(T-1)} \sum_{i=1}^{N} \sum_{t=1}^{T} \left[e_{it} - \frac{1}{T} \sum_{t=1}^{T} e_{it} - \frac{1}{N} \sum_{i=1}^{N} e_{it} \right]^2,$$

(15b) $$\hat{\sigma}_u^2 = \frac{1}{T} \left\{ \frac{1}{(N-1)T} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} e_{it} \right]^2 - \hat{\sigma}_w^2 \right\},$$

(15c) $$\hat{\sigma}_v^2 = \frac{1}{N} \left\{ \frac{1}{N(T-1)} \sum_{t=1}^{N} \left[\sum_{i=1}^{T} e_{it} \right]^2 - \hat{\sigma}_w^2 \right\},$$

where e_{it} represents the residuals obtained by applying the OLS method to the pooled data.

Da Silva Model

The Da Silva model (Da Silva, 1975) is also known as the variance-component moving average model. This model is similar to the error components model previously discussed. The model consists of a variance component for cross-sections and a various component for time periods. The Da Silva method estimates a variance-component moving-average error process, that is the error term corresponds to a moving-average time-series of order m < T-1 for each cross-section i. As such,

$$(16)\quad \varepsilon_{it} = \alpha_0 e_{it} + \alpha_1 e_{it-1} + \cdots + \alpha_m e_{it-m} \qquad t = 1, \dots, T;\ i = 1, \dots, N$$

In this specification, $\alpha_0, \alpha_1, \dots, \alpha_m$ are parameters, and e_{it} is a sequence of mutually uncorrelated random variables such that for each $i\ E(e_t) = 0\ and\ E(e_t^2) = \sigma_e^2$.

With this procedure, in addition to the parameters associated with the explanatory variables, the variance components associated with the respective cross-sections and time periods, it is necessary to estimate m+1 parameters in regard to the error term. Importantly, this model rests on the identification of m.[2]

Analysis of Covariance Model

Up to this point, the implicit assumption common to these models discussed thus far using pooled data is that the coefficients of the explanatory variables are equal across time periods and across cross-sectional units. We now discuss the analysis of covariance (ANACOVA) model where this assumption is relaxed. In the ANACOVA model, each cross-sectional unit and each time period is characterized by dummy variables (intercept shifters).

$$(17)\quad \begin{aligned} Y_{it} &= \beta_1 + \beta_2 X_{it,2} + \ldots + \beta_k X_{it,k} + \gamma_2 Z_{2t} + \gamma_3 Z_{3t} + \ldots \\ &+ \gamma_N Z_{Nt} + \delta_2 W_{i2} + \delta_3 W_{i3} + \ldots + \delta_T W_{iT} + \varepsilon_{it} \end{aligned}$$

Z_{it} = 1 for the i^{th} cross-sectional unit
= 0 otherwise

W_{it} =1 for the t^{th} time period
= 0 otherwise (t = 2, 3, …, T)

This model is applicable for unbalanced designs. The ANACOVA model also is known as the Least Squares with Dummy Variables (LSDV) model. Each cross-sectional unit and/or time series unit is characterized by dummy variables. This model works well with unbalanced designs. This model allows analysts to test for equality of intercepts and/or estimated coefficients.

The appropriate estimation technique is OLS. One does not have to include both cross-sectional and time indicator variables in the model. Depending on the magnitude of N, the number of cross-

[2] The choice of m typically rests on the use of model selection critera such as the Akaike Information Criteria (AIC). The Schwart Information Criteria (SIC), or the Hannon-Quinn Information Criteria (HQC).

sectional entities, and T, the number of time periods, the number of parameters to be estimated can greatly increase. The number of parameters to be estimated with N cross-sectional entities is N-1, and the number of parameters to be estimated with T time periods is T-1. Often, interest only lies in the inclusion of the cross-sectional indicator variables.

With the most general form of the ANACOVA model, we can test whether the inclusion of the cross-sectional dummy variables or the inclusion of the time-series dummy variables is necessary. To that end, we perform joint F-tests corresponding to the set of cross-sectional effects or time-series effects. Importantly, the ANACOVA or LSDV model implicitly assumes that the coefficients associated with the set of k explanatory variables, the so-called "slope parameters," are constant across time periods as well as cross-sectional units.

Importantly, violations of the classical assumptions of the error terms might occur with the pooled model or the ANACOVA model. If these error terms exhibit autoregressive or heteroscedastic patterns, then the use of several methods (e.g. the Newey-West procedure (Newey and West, 1987)) that deal with these issues in lieu of OLS is the appropriate estimation method. These methods produce heteroscedasticity-consistent (HCCME) and heteroscedasticity- and autocorrelation-consistent (HAC) covariance matrices. The presence of heteroscedasticity and autocorrelation can result in inefficient and biased estimates of the covariance matrix in the of OLS estimation.

Section V. Fixed and Random Effects

It is worthwhile to note a similarity of the ANACOVA or LSDV model with the error components model. The comparison is especially pertinent with the LSDV model, where we include only intercept shifters corresponding to the cross-sectional units, and with the two-component model, where the decomposition of the error term rests only on the cross-sectional component and a random component.

In the case of the error component model, the specific characteristics of the cross-sections rest on a normally distributed random variable, whereas in the case of the ANACOVA or LSDV model, these specific characteristics rest on parameters associated with dummy variables. Statistically speaking, the error components model often is referred to as a random-effects model, whereas the ANACOVA model is referred to as a fixed-effects model (Mundlak, 1978). The advantage of using the error components model is the reduction in the number of parameters to estimate. Thus, with the error components model, we save on degrees of freedom vis-à-vis the ANACOVA model.

The error components model rests on estimating the variance of the cross-sectional component, whereas the covariance model rests on estimating minimally N-1 additional parameters. Consequently, there is a loss of statistical efficiency associated with the ANACOVA model in comparison to the error components model.

The disadvantage of the random-effects or error components model is the potential correlation of the cross-sectional component with the set of explanatory variables in the model. If this correlation exists, then the parameter estimates of the coefficients associated with the explanatory factors are biased and inconsistent. This issue does not arise with the fixed-effects covariance model. In a nutshell, the crucial factor to consider is the possibility of correlation between the cross-sectional units with the explanatory variables of the model.

It is always good practice to conduct a Hausman test (1978) with random-effects models. Under the null hypothesis of the absence of any correlation between the variance components and the set of explanatory variables, the parameter estimates from the error components model should not be very different from the parameter estimates from the ANACOVA or LSDV model. If $\tilde{\beta}$ denotes the least squares covariance estimator and $\tilde{\beta}$ denotes the error component estimator, then $\hat{q} = \hat{\beta} - \tilde{\beta}$ is the basis for the relevant test statistic. The test statistic is given by:

(18) $$\hat{q}'\left[\mathrm{var}(\hat{q})\right]^{-1}\hat{q} \sim \chi^2_{k-1}, \text{ where } \mathrm{var}(\hat{q}) = \mathrm{var}(\hat{\beta}) - \mathrm{var}(\tilde{\beta}).$$

The test statistic rests on both the magnitude of the parameter estimates and the variance-covariance matrix of the difference between the sets of parameters estimates. Assuming k explanatory variables including a constant term, the test statistic asymptotically follows a chi-squared distribution with k-1 degrees of freedom.

Section VI. Estimation Procedure

The PANEL procedure analyzes a class of linear econometric models that arise when time-series and cross-sectional data are combined. The panel data models can be grouped into several categories depending on the structure of the error term. The PANEL procedure allows the estimation of: (1) the pooled model; (2) one-way and two-way random effects models; (3) one-way and two-way fixed-effects models; (4) cross-sectionally heteroscedastic and time-wise autogressive error terms of the Parks model; and (5) moving average error terms of the Da Silva model.

With the PROC PANEL procedure for unbalanced or balanced designs, it is necessary initially to sort the data by cross-section and then by time period. The PROC SORT procedure works in this capacity. The model statement for PROC PANEL is similar to the model statement for PROC REG or PROC AUTOREG but for the options. For the pooled model, the appropriate options are / pooled HAC Neweywest. Note that HAC Neweywest allows for the heteroscedastic and autocorrelation correction of the error terms with the Newey-West procedure (Newey and West, 1987). For the one-way or two-way fixed-effects model, the options are / fixone (or fixtwo) HAC Neweywest[3]. In the case of the one-way or two-way random-effects model, there are several options given that PROC PANEL allows four different estimation methods of random-effects models.

Fuller-Battese Method	VCOMP = FB RANONE or RANTWO
Wansbeek-Kapteyn Method (default)	VCOMP = WK RANONE or RANTWO
Wallace-Hussain Method	VCOMP = WH RANONE or RANTWO
Nerlove Method	VCOMP = NL RANONE or RANTWO

[3] With the option fixone, PROC PANEL provides estimates of the coefficients of cross-sectional dummy variables. With the option fixonetime, PROC PANEL provides estimates of the coefficients of time-series dummy variables.

Additionally with the option bp in the case of one-way random effects models or the option bp2 in the case of two-way random effects models, it is possible to test the null hypothesis of the absence of random-effects is done. The test developed by Breusch and Pagan (1980) is distributed as a x^2-statistic with one degree-of-freedom for one-way random effects models and with two degrees-of-freedom for two-way random-effects model.

For the Parks model, the appropriate options are / Parks Phi rho, and for the Da Silva model, the appropriate options are / Dasilva m=?. With the Da Silva model, it is necessary to specify m, that is replace ? with an integer, the order of the moving-average process. The default is m=1. The options HAC and Neweywest also may be added with the estimation of the respective random-effects models, the Parks model, and the Da Silva model.

Section VII. Sample Problem of the Use of PROC PANEL

The data for this sample problem consists of five retailers (number of cross-sections) across 165 weeks (number of time periods) or 825 observations in total. This example corresponds to a balanced design. The idea is to pool the data in order to obtain a common estimate of the own-price elasticity for a food product. A simplistic version of the model specification for pedagogical purposes is given by:

$$logunits_{it}=a_0+a_1{*}logp_{it}+a_2{*}logdisc_{it}+a_3{*}logdisp_{it}+a_4{*}logad_{it}+e_{it},$$

where *i* refers to cross-sections (retailers) and *t* refers to time periods (weeks). The dependent variable $logunits_{it}$ refers to the amount of the food product sold by retailer i in week t; $logp_{it}$ refers to the price of the food product sold by retailer i in week t; $logdisc_{it}$ corresponds to the amount of the discount associated with this food product sold by retailer i week t; $logdisp_{it}$ refers to the amount of display space provided for this food product sold by retailer i in week; and $logad_{it}$ corresponds to the amount of advertising expenditures associated with this food product sold by retailer i in week t.

Given the mathematical form of the model, the coefficients correspond to elasticities, the percentage change in units sold of this cereal product attributed to a one percent change in price (a_1), the amount of the discount (a_2), the amount of display space provided (a_3), and the amount of advertising expenditures (a_4). The coefficient a_1 is hypothesized to be negative to conform to the law of demand in economics, while the signs of the remaining coefficients are hypothesized to be positive.

The SAS program for this analysis of a balanced panel of retail food sales data is as follows:

```
Proc Sort data=Ford2.retail_food;
       by retailer date;
run;

OLS pooled model with HAC correction

Proc Panel data=Ford2.retail_food;
       id retailer date;
       model log_units = log_p log_disc log_disp log_ad /pooled HAC Neweywest;
run;

One-way fixed effects model cross-sectional effects only

Proc Panel data=Ford2.retail_food printfixed;
       id retailer date;
       model log_units = log_p log_disc log_disp log_ad /fixone HAC Neweywest;
run;

Two-way random effects model(Fuller-Battese method)

Proc Panel data=Ford2.retail_food;
       id retailer date;
       model log_units = log_p log_disc log_disp log_ad /rantwo vcomp=fb HAC Neweywest bp;
run;

Parks model

Proc Panel data=Ford2.retail_food;
       id retailer date;
       model log_units = log_p log_disc log_disp log_ad /Parks Phi rho HAC Neweywest;
run;

Da Silva model

Proc Panel data=Ford2.retail_food;
       id retailer date;
       model log_units = log_p log_disc log_disp log_ad /Dasilva m=4 HAC Neweywest;
run;
```

The empirical results associated with the use of PROC PANEL for these respective models are exhibited in Tables 1-5. For all five models, the estimated coefficients conform the anticipated signs of the coefficients.

In Table 1, the results for the pooled (OLS) model are presented. In Table 2, the results for the one-way fixed-effects models are listed. The CS1, CS2, CS3, and CS4 variables correspond to dummy variables for retailer 1, retailer 2, retailer 3, and retailer 4 respectively (the cross-sectional fixed effects.) The base or reference category is retailer 5. The reference category is arbitrary, but the reference category corresponds to the last cross-section specified in the sorting procedure. With the PROC PANEL procedure, for the one-way fixed effects model, the F-test associated with the null hypothesis of no fixed effects is calculated. In this example, this null hypothesis is rejected.[4]

[4] The one-way fixed-effects model for cross-sections only requires N-1 dummy variables. The one-way fixed-effects model for time-series requires T-1 dummy variables. In this example N=5 and T=165. To cut down on the number of parameters to be estimated, only the one-way fixed effects model for cross-sections is discussed.

In Table 3, the results for the two-way random-effects model are presented. The Fuller-Battese method is chosen to estimate the variance components.[5] Note that the variance component for cross-sections (σ_u^2) dominates that variance component for time series (σ_v^2). The variance of the error term is the sum of these three components, 0.1678. Based on the Hausman test for random effects, the null hypothesis of no correlation of the variance components with the set of explanatory factors is not rejected. Based on the Breusch-Pagan test, the null hypothesis of the absence of random effects is rejected.

In Table 4, the results of the Parks Model are listed. Note that the first-order autoregressive parameter estimates are positive and noticeably larger than zero. As such, a positive serial correlation process of order 1 of the error term exists in each cross-section. On the basis of this result, the pooled model (Table 1) is not likely to be the appropriate specification. Based on the estimated phi matrix exhibited in Table 4, the diagonal elements, which correspond to the variance of the error terms in each cross-section, are different, ranging from 0.0654 to 0.1850. Consequently, evidence exists to support the contention of heteroscedasticity among the respective cross-sections. The off-diagonal elements of the phi matrix, which correspond to the covariances of the error terms of the respective cross-sections, are rather small, ranging from -0.0411 to 0.0335. On this basis, evidence exists to support the contention that the degree of correlation among cross-sections is negligible.

In Table 5, the results of the Da Silva Method are presented. The order of the moving-average (MA) process in the error terms chosen for this example is four. Twelve different orders of the MA process were considered, 1 through 12. The selection of order 4 stems from the minimization of the mean squared error (MSE) of the residuals. Similar to the two-way random-effects model previously discussed, the variance component for the cross-sections dominates the variance component for the weekly time series.

A comparison of the estimated coefficients of the explanatory variables across all models is exhibited in Table 6. Note that the pooled model yields the lowest goodness-of-fit-measure (R^2**)** at 0.5426. In the other models, the R^2 metrics range from 0.8728 to 0.9407. In addition, in the Da Silva model, none of the estimated coefficients is statistically different from zero despite a goodness-of-fit measure of 0.9040. Moreover, in the one-way fixed-effects model, the two-way random-effects model, and the Parks model, all estimated coefficients are statistically different from zero. Further, the one-way fixed effects model and the two-way random effects model yield very similar estimates of the respective coefficients. But notable differences are evident of the estimates of the own-price elasticity and the advertising elasticity in the Parks model (-1.6581 and 0.4682) vis-à-vis the one-way fixed-effects model (-0.7899 and 0.2793) and the two-way random-effects model (-0.8748 and 0.2869). In economic parlance on the basis of the Parks model, the demand for the food product is elastic, whereas on the basis of the one-way fixed-effects model and the two-way random effects the demand for the food product is inelastic. Finally, the impact from advertising is roughly 1.6 times higher on the basis of the Parks model relative to the impact from advertising on the basis of the one-way fixed-effects model and the two-way random-effects model. Bottom line, the selection of the appropriate specification is not only very important but also should not be made purely on statistical grounds.

[5] The empirical results for the Wallace-Hussein method, the Wansbeek-Kapteyn method, and the Nerlove method were very similar to those obtained using the Fuller-Battese Method.

Table 1. The PANEL Procedure Pooled (OLS) Estimates

Model Description	
Estimation Method	Pooled
Number of Cross Sections	5
Time Series Length	165
HAC Kernel	Bartlett
HAC Bandwidth	Newey and West

Fit Statistics			
SSE	495.9952	**DFE**	820
MSE	0.6049	**Root MSE**	0.7777
R-Square	0.5426		

Parameter Estimates				
Variable	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	11.1471	0.3443	32.37	<.0001
logp_1	-2.0666	0.3151	-6.56	<.0001
logdisc	0.0466	0.4250	0.11	0.9128
logdisp	3.6824	0.2599	14.17	<.0001
logad	0.4137	0.1646	2.51	0.0121

Table 2. The PANEL Procedure Fixed One Way Estimates

Model Description	
Estimation Method	FixOne
Number of Cross Sections	5
Time Series Length	165
HAC Kernel	Bartlett
HAC Bandwidth	Newey and West

Fit Statistics			
SSE	64.2781	**DFE**	816
MSE	0.0788	**Root MSE**	0.2807
R-Square	0.9407		

F Test for No Fixed Effects			
Num DF	**Den DF**	**F Value**	**Pr > F**
4	816	1370.14	<.0001

Parameter Estimates				
Variable	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
CS1	0.9071	0.0485	18.70	<.0001
CS2	0.9373	0.0752	12.46	<.0001
CS3	0.9664	0.0358	26.99	<.0001
CS4	-1.1998	0.0647	-18.54	<.0001
Intercept	9.2257	0.4780	19.30	<.0001
logp_1	-0.7899	0.4278	-1.85	0.0652
logdisc	3.6074	0.1822	19.80	<.0001
logdisp	1.8955	0.1219	15.55	<.0001
logad	0.2793	0.0783	3.57	0.0004

Table 3. The PANEL Procedure Fuller and Battese Variance Components (RanTwo)

Model Description	
Estimation Method	RanTwo
Number of Cross Sections	5
Time Series Length	165
HAC Kernel	Bartlett
HAC Bandwidth	Newey and West

Fit Statistics			
SSE	60.2128	**DFE**	820
MSE	0.0734	**Root MSE**	0.2710
R-Square	0.8728		

Variance Component Estimates	
Variance Component for Cross Sections	0.8906
Variance Component for Time Series	0.0054
Variance Component for Error	0.0734

Hausman Test for Random Effects		
DF	**m Value**	**Pr > m**
4	4.60	0.3304

Breusch Pagan Test for Random Effects (Two-Way)		
DF	**m Value**	**Pr > m**
2	35242.4	<.0001

Parameter Estimates				
Variable	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	9.6509	0.7364	13.11	<.0001
logp_1	-0.8748	0.4110	-2.13	0.0336
logdisc	3.5611	0.1793	19.86	<.0001
logdisp	1.9149	0.1223	15.65	<.0001
logad	0.2869	0.0778	3.69	0.0002

Table 4. The PANEL Procedure Parks Method Estimation

Model Description	
Estimation Method	Parks
Number of Cross Sections	5
Time Series Length	165
HAC Kernel	Bartlett
HAC Bandwidth	Newey and West

Fit Statistics					
SSE	783.4189	**DFE**	820	**R-Square**	0.9142
MSE	0.9554	**Root MSE**	0.9774		

Parameter Estimates				
Variable	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	10.7135	0.6425	16.68	<.0001
logp	-1.6581	0.6211	-2.67	0.0077
logdisc	3.2539	0.0605	53.78	<.0001
logdisp	1.7622	0.0373	47.25	<.0001
logad	0.4682	0.0152	30.90	<.0001

First Order Autoregressive Parameter Estimates	
RETAILER	**Rho**
1	0.6392
2	0.9259
3	0.6616
4	0.9231
5	0.5199

Estimated Phi Matrix					
	1	**2**	**3**	**4**	**5**
1	0.0735	0.0030	0.0335	-.0060	-.0044
2	0.0030	0.0654	0.0213	-.0107	-.0292
3	0.0335	0.0213	0.0945	-.0268	-.0411
4	-.0060	-.0107	-.0268	0.0824	0.0124
5	-.0044	-.0292	-.0411	0.0124	0.1850

Table 5. The PANEL Procedure Da Silva Method Estimation

Model Description	
Estimation Method	DaSilva
Number of Cross Sections	5
Time Series Length	165
Order of MA Error Process	4
HAC Kernel	Bartlett
HAC Bandwidth	Newey and West

Fit Statistics			
SSE	738.2562	**DFE**	820
MSE	0.9003	**Root MSE**	0.9488
R-Square	0.9040		

Variance Component Estimates	
Variance Component for Cross Sections	0.8583
Variance Component for Time Series	0.0197

Estimates of Autocovariances	
Lag	**Gamma**
0	0.1019
1	0.0535
2	0.0233
3	0.0167
4	0.0105

Parameter Estimates				
Variable	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	9.1526	234.1	0.04	0.9688
logp_1	-0.4691	118.9	-0.00	0.9969
logdisc	3.4655	3.9757	0.87	0.3836
logdisp	1.9344	1.5614	1.24	0.2157
logad	0.3958	0.9797	0.40	0.6863

Table 6. A Comparison of the Estimated Coefficients in the Explanatory Variables in the Example with the Use of PROC PANEL

Variable	Pooled	One-Way Fixed Effects	Two-Way Random Effects	Parks	Da Silva (m=4)
Intercept	11.1472*	9.2257*	9.6509*	10.7135*	9.1526
log	-2.0666*	-0.7899*	-0.8748*	-1.6581*	-0.4691
logdisc	0.0466	3.6073*	3.5611*	3.2539*	3.4655
logdisp	3.6824*	1.8955*	1.9149*	1.7622*	1.9344
logad	0.4137*	0.2793*	0.2869*	0.4682*	0.3958
R^2	0.5426	0.9407	0.8728	0.9412	0.9040

*Significant at the 0.05 level.
Source: Compiled by the author.

Section VIII. Seemingly Unrelated Regression Model

We now focus on the consideration of separate regressions for each cross-section and/or time series unit characterized by separate equations. Known as the seemingly unrelated regression (Zellner, 1962), this specification (SUR) works only with balanced data. With the use of the SUR procedure, each equation is initially estimated by OLS. The correlation of the error terms then is taken into account in the second step of this procedure. As such, the SUR procedure is a joint generalized least squares estimation technique. In the ensuing discussion, we assume the estimation of N separate regressions, one for each cross-sectional unit. With the estimation of the N separate regressions, we obtain parameter estimates that vary across common explanatory variables, unlike the previously discussed single-equation models associated with the pooling of time-series and cross-sectional data.

With the SUR model, we allow for the presence of contemporaneous correlation across the N error terms. In addition, we may also allow for the presence of autoregressive processes in each equation (Kmenta and Gilbert, 1968). Importantly, this technique allows the coefficients associated with common explanatory variables to vary by cross-sectional unit. As well, joint F-tests can be used in conjunction with each of the k explanatory variables to examine whether the impacts of changes in any explanatory variable vary significantly by cross-sectional entity.

We use all NT observations in this process, so we preserve the pooled sample. Also, this technique improves the statistical efficiency of the parameter estimates. On an equation-by-equation basis, the standard errors of the estimated coefficients in the SUR model are lower (or minimally the same) in comparison to those obtained through the use of OLS. We may use either PROC SYSLIN or PROC MODEL to estimate SUR models. As well, we may use joint F-tests to determine whether or not coefficients of the explanatory variables vary significantly across cross-sectional units. The only visible downside to the estimation of SUR models is the size of N, the number of equations or cross-sections in the system. Even with a sizeable number of cross-sectional units, this potential limitation is not much of a constraint.

Section IX. Sample Problem of the Use of Seemingly Unrelated Regression

The SAS program for this example of the use of seemingly unrelated regression (SUR) is as follows:

```
data all; merge ret1 ret2 ret3 ret4 ret5; by week;
proc model data=all;
ret1_logunits=a0+a1*ret1_logprice+a2*ret1_logdisc+a4*ret1_logdisp+a5*ret1_logad;
ret2_logunits=b0+b1*ret2_logprice+b2*ret2_logdisc+b4*ret2_logdisp+b5*ret2_logad;
ret3_logunits=c0+c1*ret3_logprice+c2*ret3_logdisc+c4*ret3_logdisp+c5*ret3_logad;
ret4_logunits=d0+d1*ret4_logprice+d2*ret4_logdisc+d4*ret4_logdisp+d5*ret4_logad;
ret5_logunits=e0+e1*ret5_logprice+e2*ret5_logdisc+e4*ret5_logdisp+e5*ret5_logad;
%ar(ret1_logunits,1);
%ar(ret2_logunits,1);
%ar(ret3_logunits,1);
%ar(ret4_logunits,1);
%ar(ret5_logunits,1);
fit ret1_logunits ret2_logunits ret3_logunits ret4_logunits ret5_logunits / sur dw dwprob
out=retailersur;
parms a0 a1 a2 a4 a5 b0 b1 b2 b4 b5 c0 c1 c2 c4 c5 d0 d1 d2 d4 d5 e0 e1 e2 e4 e5;
test a1-b1=0; test a1-c1=0; test a1-d1=0; test a1-e1=0; test b1-c1=0; test b1-d1=0;
test b1-e1=0; test c1-d1=0; test c1-e1=0; test d1-e1=0;
run;
```

Each of the data sets for respective cross-sections (ret1 through ret5) consist of 165 weekly time periods. These five data sets subsequently are merged by week. Five separate equations embody the SUR framework. The macro %***ar*** allows for a first-order autoregressive process of the error terms in each of the respective equations. With the fit command, we estimate the coefficients of the explanatory variables in the five-equation system with the SUR procedure, and we obtain the Durbin-Watson (dw) test statistics along with their associated p-values (dwprob). In this example, we also demonstrate the use of statistical tests associated with the own-price elasticities of the food product for the five different retailers. The null hypothesis is that the respective own-price elasticities are the same across retailers. In this example, with five retailers, there are ten tests of hypotheses with respect to the own-price elasticity of the food product.

The summary of the goodness-of-fit statistics and the statistics associated with autocorrelation in the error terms are exhibited in Table 7. The R^2 and adjusted R^2 metrics range suggest that the model explains a notable amount of variability in the volume of units sold for this food product for each of the respective retailers. These statistics range from roughly 74% to 96%. Based on the estimated coefficients of the %***ar*** macro, a first-order autoregressive process is evident in the error terms of each of the respective five equations. Each of these estimated coefficients is positive (indicative of positive serial correlation of the residuals) and significantly different from zero at the 0.05 level. The Durbin-Watson statistics indicate that the correction for the first-order autoregressive process leads to white noise (the absence of any systematic pattern) in the error terms in each of the five equations.

Table 7. Summary of Goodness-of-Fit Statistics and Statistics Associated with Autocorrelation from the PROC MODEL Procedure (SUR)

Equation	R^2	Adjusted R^2	Coefficient Associated with First-Order Autoregression Process	Durbin-Watson
Retailer 1	0.7429	0.7346	0.5771*	1.9790
Retailer 2	0.9490	0.9474	0.3995*	2.0402
Retailer 3	0.9613	0.9603	0.3637*	1.9291
Retailer 4	0.8957	0.8923	0.3466*	2.0288
Retailer 5	0.9305	0.9283	0.1664*	1.9286

*Significantly different from zero at the 0.05 level.

The estimated coefficients, standard errors, and p-values are summarized in Table 8. The own-price elasticities of the food product are negative as hypothesized and significantly different from zero at the 0.05 level for retailers 2 and 4. The own-price elasticities of the food product are negative for retailers 3 and 5, but these coefficients are not statistically different from zero. The own-price elasticity of the food product for retailer 1 however is positive but this coefficient is not statistically different from zero. The elasticities associated with discounts are positive across all retailers as hypothesized, and these coefficients are significantly different from zero for retailers 2, 3, 4, and 5. The elasticities associated with displays are positive and significantly different from zero across all five retailers as expected. Finally, the elasticities associated with advertising are all positive for the set of retailers as expected but significantly different from zero only for retailers 2 and 3.

The summary of the tests of hypotheses in regard to the equality of the own-price elasticities is exhibited in Table 9. The test statistic (a Wald test) follows a chi-squared distribution with one degree-of-freedom. In three of the ten pairwise tests, the equality of the own-price elasticities is rejected the 0.05 level of significance. This situation occurs for retailer 1 versus retailer 2, retailer 1 versus retailer 3, and retailer 4 versus retailer 5.

Table 8. Parameter Estimates, Standard Errors and p-Values Associated with the Coefficients in the SUR Model

	Retailer 1	Retailer 2	Retailer 3	Retailer 4	Retailer 5
Intercept	9.3516*[a]	12.0315*	13.6615*	11.2548*	8.7373*
	-0.5295[b]	-1.1472	(3.9003)	(1.0789)	(0.7060)
	(0.0000)[c]	(0.0000)	(0.0006)	(0.0000)	(0.0000)
logp	0.0538	-2.0553*	-3.6749	-3.1453*	-0.4014
	(0.5457)	(0.8128)	(3.3099)	(0.7916)	(0.6335)
	(0.9217)	(0.0125)	(0.2682)	(0.0001)	(0.5272)
logdisc	0.7491	2.8040*	4.8289*	4.1428*	3.0511*
	(0.7733)	(0.2153)	(0.3406)	(0.1901)	(0.3774)
	(0.3342)	(0.0000)	(0.0000)	(0.000)	(0.0000)
logdisp	2.1758*	0.8750*	1.0891*	1.0681*	3.2028*
	(0.1636)	(0.2119)	(0.1310)	(0.1463)	(0.2259)
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
logad	0.1231	1.2261*	0.2394*	0.1366	0.1266
	(0.0783)	(0.0971)	(0.1047)	(0.0831)	(0.1056)
	(0.1182)	(0.0000)	(0.0236)	(0.1023)	(0.2327)

[a] Parameter Estimate
[b] Estimates of Standard Error
[c] p-Value
* Significant at the 0.05 level.

Table 9. Summary of the Tests of Hypotheses in Regard to the Equality of the Own-Price Elasticities of the Food Product across the Respective Retailers in the SUR Model

Null Hypothesis (Ho*)	Test Statistic [a]	p-value [b]	Decision
Elasticity for Retailer 1 = Elasticity for Retailer 2	4.64	0.0312	**Reject Ho**
Elasticity for Retailer 1 = Elasticity for Retailer 3	1.29	0.2568	Fail to reject Ho
Elasticity for Retailer 1 = Elasticity for Retailer 4	11.29	0.0008	**Reject Ho**
Elasticity for Retailer 1 = Elasticity for Retailer 5	0.26	0.6102	Fail to reject Ho
Elasticity for Retailer 2 = Elasticity for Retailer 3	0.22	0.6359	Fail to reject Ho
Elasticity for Retailer 2 = Elasticity for Retailer 4	0.93	0.3339	Fail to reject Ho
Elasticity for Retailer 2 = Elasticity for Retailer 5	2.60	0.1071	Fail to reject Ho
Elasticity for Retailer 3 = Elasticity for Retailer 4	0.02	0.8764	Fail to reject Ho
Elasticity for Retailer 3 = Elasticity for Retailer 5	0.91	0.3404	Fail to reject Ho
Elasticity for Retailer 4 = Elasticity for Retailer 5	7.42	0.0064	**Reject Ho**

Bold indicates significance at the 0.05 level.
[a] Test statistic is a Wald statistic.
[b] Test statistic follows an x^2 distribution with degree of freedom.

Section X. Concluding Remarks

- In applied econometrics, practitioners have access to data not only over time but also by cross-section. This type of data set often is referred to as pooled data.
- The key question in dealing with panel data is whether analysts can we live with common pooled estimates of the coefficients of the explanatory variables. If the answer to this question is yes, then single-equation models are the way to proceed with the use of PROC PANEL. If the answer to this question is no, then we can run separate regressions by cross-sectional unit, a seemingly unrelated regression (SUR) model, provided sufficient observations exist to not only estimate parameters of the model but also to conduct statistical tests of hypotheses. In this case, PROC SYSLIN or PROC MODEL are the appropriate SAS procedures.
- PROC PANEL is the appropriate SAS procedure to handle the estimation of the Parks model, the error components model (or random effects model), the Da Silva, and the ANACOVA or LSDV model (or fixed effects model).
- When dealing with pooled data sets with balanced designs, analysts may use the Parks model and the Da Silva model.
- When dealing with pooled data sets with unbalanced designs, analysts may use the error components model as well as the analysis of covariance (ANACOVA) or least squares with dummy variables (LSDV) models.
- Analysts should use the Hausman test in random-effects models to determine if the error components are correlated with the set of explanatory variables.
- Analysts should use the Breusch-Pagan test in random-effects models to determine if random effects are present.
- In the estimation of random-effects models, options include the Fuller-Battese method, the Wansbeck-Kapteyn method, the Wallace-Hussein method, and the Nerlove method. No obvious advantages are evident in conjunction with these estimation methods.
- For balanced designs, the SUR model enables the estimation of separate regressions for each cross-section. Joint F-tests can be used in conjunction with each of the explanatory variables to examine whether the impacts of changes in any explanatory variable vary significantly by cross-sectional entity.

References

Balestra, P., and M. Nerlove. 1966. "Pooling Cross-Section and Time-Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas." *Econometrica* 34: 585-612.

Baltagi, B.H. 2001. *Econometric Analysis of Panel Data*. Second Edition. New York: John Wiley and Sons.

Breusch, T.S. and A.R. Pagan. 1980. "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics." *Review of Economic Studies* 47: 239-253.

Capps, O., Jr., and J. Havlicek, Jr. 1978. "The Demand for Gasoline and Diesel Fuel in Agricultural Use in Virginia." *Southern Journal of Agricultural Economics* 10(1): 59-65.

Da Silva, J.G.C. 1975. *The Analysis of Cross-Sectional Time Series Data*. Ph.D. dissertation, North Carolina State University, Department of Statistics.

Dielman, T.E., T.J. Nantell, and R.L. Wright. 1980. "Price Effects of Stock Repurchasing: A Random Coefficient Regression Approach." *Journal of Financial and Quantitative Analysis* 15: 175-189.

Dielman, T.E. 1989. *Pooled Cross-Sectional and Time Series Data Analysis*. New York: Marcel-Dekker.

Dielman, T.E. 1983. "Pooled Cross-Sectional and Time Series Data: A Survey of Current Statistical Methodology." *American Statistician* 37(2): 111-122.

Ferguson, I.S. and J.W. Leech. 1978. "Generalized Least Squares Estimation of Yield Functions." *Forest Science* 24: 27-42.

Granger, C.W.J., R. Engle, R. Ramanathan, and A. Anderson. 1979. "Residential Load Curves and Time-of-Day Pricing: An Econometric Analysis." *Journal of Econometrics* 9:13-32.

Fuller, W.A., and G.E. Battese. 1974. "Estimation of Linear Models with Crossed-Error Structure." *Journal of Econometrics* 2: 67-78.

Hausman, J. A. 1978. "Specification Tests in Econometrics." *Econometrica* 46: 1251–1271.

Kmenta, J. 1986. *Elements of Econometric*, Second Edition, New York: MacMillan.

Kmenta, J., and R. Gilbert. 1968. "Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions." *Journal of the American Statistical Association* 63: 1180-1200.

Kuh, E., and J.R. Meyer. 1957. "How Extraneous are Extraneous Estimates?" *The Review of Economics and Statistics* 39: 380-393.

Mehta, J.S., G.V.L. Narasimham, and P.A.V.B. Swamy. 1978. "Estimation of a Dynamic Demand Function for Gasoline with Different Schemes of Parameter Variation." *Journal of Econometrics* 7(3): 263-279.

Mundlak, Y. 1978. "On the Pooling of Time Series and Cross-Sectional Data." *Econometrica* 56: 69-86.

Nerlove, M. 1971. "Further Evidence on the Estimation of Dynamic Relations from a Time Series of Cross Sections." *Econometrica*, 39, 359–382.

Nerlove, M. 2003. *Essays in Panel Data Econometrics*. Cambridge: Cambridge University Press.

Newey, W. K. and K.D. West. 1987. “A Simple, Positive Semi-definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica* 55(3): 703-708.

Parks, R.W. 1967. “Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated.” *Journal of the American Statistical Association* 62: 500-509.

Sims, C. 1989. “Models and Their Uses.” *American Journal of Agricultural Economics* 489-494.

Wallace, T., and A. Hussain. 1969. “The Use of Error Components in Combining Cross Section with Time Series Data.” *Econometrica* 37: 55-72.

Wansbeek, T., and Kapteyn, A. 1989. “Estimation of the Error-Components Model with Incomplete Panels.” *Journal of Econometrics* 41: 341-361.

Ward, R.W., and J.W. Davis. 1978. “A Pooled Cross-Section Time-Series Model of Coupon Promotions.” *American Journal of Agricultural Economics* 60(3): 393-401.

Ward, R.W., and W.F. McDonald. 1986. “Effectiveness of Generic Milk Advertising: A Ten Region Study.” *Agribusiness: An International Journal* 2(1): 77-89.

Zellner, A. 1962. “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias.” *Journal of the American Statistical Association* 57, 348–368.