

Panel Data ITV Class
Brian Gaines

Assignment 1. (12.10.05)

Find an article on a topic of interest to you that makes use of panel data (in the broadest sense, i.e. data having multiple dimensions). Discuss the nature of the data and what models the authors used, along with anything they mentioned about why they chose the particular model (e.g. why fixed or random effects or why they needed to “correct” the standard errors). This is meant to be only a brief summary and need not exceed a page or two. Also, you will not have to do your final paper/note on this same topic or using these same data, though you are, of course, free to do so.

Assignment 2. (19.10.05)

Marc Nerlove’s many influential papers on panel-data econometrics are now collected in a book published by Cambridge, *Essays in Panel Data Econometrics* (2002). The book opens with a long, fairly technical history of statistical models for panel data that is worth reading. An appendix to that chapter that describes decomposition of the sum of squares of a variable observed in a two-dimensional panel is attached. However, it appears that Nerlove changed his mind about notation midway through, and didn’t proof read for typos, so the appendix is not strictly correct. Your assignment is to fix the mistakes and write down the proper matrices to perform the decomposition. To be clear, I’m not looking for formal proofs—just corrected matrices and equations. If you want to illustrate/check your work by hand, you will probably want to use a *small* example, say $N=3$ and $T=2$. Alternatively (optionally), you can write short program using your corrected matrices and then apply it to a real data set.

APPENDIX A

Decomposition of the Sum of Squares of a Variable
Observed in a Two-Dimensional Panel¹

Let x_{it} be a typical observation associated with a panel consisting of N individuals, on each of which we have T observations. Denote

the mean across time for individual i : $\bar{x}_i = \frac{1}{T} \sum_1^T x_{it}$,

the mean across individuals for time period t : $\bar{x}_t = \frac{1}{N} \sum_1^N x_{it}$,

the overall mean across individuals and over time periods t : \bar{x}

$$= \frac{1}{NT} \sum_1^T \sum_1^N x_{it}. \quad (1)$$

The following notation is useful: a single parameter or observation, even if on a random variable, is denoted by a lower case latin or greek letter, with or without subscripts for individuals (groups) and/or time dimension. Matrices are generally denoted by upper-case latin letters. It will generally be true that vectors and scalars can be given in the same notation without ambiguity and great saving in time, although many authors are at pains to distinguish. The identity matrix of order K is denoted by I_K , a vector of K ones by ℓ_K , and a $K \times K$ matrix of ones by $J_K = \ell_K \ell_K'$. I make extensive use of the kronecker product of two matrices.

Using the notation introduced in the preceding paragraph and the vector $x = (x_{11}, \dots, x_{1T}, \dots, x_{N1}, \dots, x_{NT})'$ of "stacked" observations, we can describe the variability of the observations x in terms of the following decomposition of the sum of squares:

¹ This appendix is adapted from Chapter 19 of Nerlove (1999a).

overall SS around the general mean:

$$\sum_1^N \sum_1^T (x_{it} - \bar{x})^2 = x' \left[I_{NT} - \frac{J_{NT}}{NT} \right] x;$$

within-group SS, around the group means:

$$\sum_1^N \sum_1^T (x_{it} - \bar{x}_i)^2 = x' \left[I_{NT} - \left(I_N \otimes \frac{J_T}{T} \right) \right] x;$$

between-group SS, variation of group means around the overall mean:

$$\sum_1^N \sum_1^T (\bar{x}_i - \bar{x})^2 = x' \left[\left(I_N - \otimes \frac{J_T}{T} \right) - \frac{J_{NT}}{NT} \right] x;$$

within-period SS, around period means:

$$\sum_1^N \sum_1^T (x_{it} - \bar{x}_t)^2 = x' \left[I_{NT} - \left(\frac{J_N}{N} \otimes I_T \right) \right] x;$$

between-period SS, variation of period means around the overall mean:

$$\sum_1^N \sum_1^T (\bar{x}_t - \bar{x})^2 = x' \left[\left(\frac{J_N}{N} \otimes I_T \right) - \frac{J_{NT}}{NT} \right] x;$$

residual SS:

$$\sum_1^N \sum_1^T (x_{it} - \bar{x}_i - \bar{x}_t + \bar{x})^2 = x' \left[I_{NT} - \left(I_N \otimes \frac{J_T}{T} \right) - \left(\frac{J_N}{N} \otimes I_T \right) + \frac{J_{NT}}{NT} \right] x. \quad (2)$$

The matrices of the quadratic forms decomposing the total SS of the variable x are, more compactly,

$$V = \left[I_{NT} - \frac{J_{NT}}{NT} \right], \quad \text{rank} = NT - 1;$$

$$B_N = \left[I_{NT} - \left(I_N \otimes \frac{J_T}{T} \right) \right] = \left(I_N - \frac{J_N}{N} \right) \otimes \frac{J_T}{T}, \quad \text{rank} = (T - 1)N;$$

$$B_T = \left[\left(I_N \otimes \frac{J_T}{T} \right) - \frac{J_{NT}}{NT} \right] = \frac{J_N}{N} \otimes \left(I_T - \frac{J_T}{T} \right), \quad \text{rank} = N - 1;$$

$$W_N = \left[I_{NT} - \left(\frac{J_N}{N} \otimes I_T \right) \right] = I_N \otimes \left(I_T - \frac{J_T}{T} \right), \quad \text{rank} = T(N - 1);$$

$$W_T = \left[\left(\frac{J_N}{N} \otimes I_T \right) - \frac{J_{NT}}{NT} \right] = \left(I_N - \frac{J_N}{N} \right) \otimes I_T, \quad \text{rank} = T - 1;$$

$$\begin{aligned}
 W^* &= \left[I_{NT} - \left(I_N \otimes \frac{J_T}{T} \right) - \left(\frac{J_N}{N} \otimes I_T \right) + \frac{J_{NT}}{NT} \right] \\
 &= \left(I_N - \frac{J_N}{N} \right) \otimes \left(I_T - \frac{J_T}{T} \right), \quad \text{rank} = (N-1)(T-1). \quad (3)
 \end{aligned}$$

The properties of the matrices, all of which are symmetric and idempotent, defined in (3), that are useful in obtaining the spectral decomposition of the variance-covariance matrix of the disturbances in a general three-component random-effects model, are

$$\begin{aligned}
 V &= W_N + W_T = B_N + B_T = W^* + W_T + B_T; \\
 W_N B_N &= W_T B_T = W^* B_N = W^* B_T = B_N B_T = 0; \\
 V \frac{J_{NT}}{NT} &= W^* \frac{J_{NT}}{NT} = W_N \frac{J_{NT}}{NT} = W_T \frac{J_{NT}}{NT} = B_N \frac{J_{NT}}{NT} = B_T \frac{J_{NT}}{NT} = 0. \quad (4)
 \end{aligned}$$

From the decomposition (2), we see that the information contained in the sum of squares or cross products among panel variables can be decomposed into six parts as detailed there. These pieces of information are more or less valuable according to the amount of "noise" in the system related to the different components of the variation. What an error-components model does is to weight these bits inversely according to their error variances. What maximum likelihood does is to estimate the weights and the relationship among observed variables simultaneously. What fixed-effects models do is to throw away a part of the observable variation bearing on the relationship to be estimated as if the error associated with that component were very large. All this in the name of avoiding correlation between the error components and the explanatory variables, which is the very source of the information being discarded.