Adjusting for Missing Information in Pooled Model through the Error Term

Dr. Abdalla Ahmed Alkhalifa Abdalla
Department of Statistics - Faculty of Science - Tabuk University (KSA)

Abstract: Griffiths and Anderson (1982) used nonlinear model to adjust for missing information through the error term, this paper adjust for missing information in linear model through the error term.

Keywords: firm effect; time effect; qualitative variable; error term; linear model

1. Introduction

If time series data and cross section data are combined "Pooled data", it's expected that the explanatory variables will not explain all variations in the dependent variable, because there may be variations due to time (time-series) or due to firm (cross-section) or due to qualitative variables in the model. The linear model for pooled data is characterized by the problem of missing information due to that, the explanatory variables are not account for all variations in the dependent variable.

Griffiths and Anderson (1982), used pooled data to estimate a production function of pastoral zone of eastern Australia, they used nonlinear model and adjusting for missing information through the error term u_it for firm "i" and time"t", i.e. The error term as a function of time effect "g_t" and firm effect "f_i" and qualitative variables "h(x)" in two cases:
1. u_it = (f_i + g_t + e_it) h_i
2. u_it = f_i + g_t + e_it h_i

The objective of this paper is to find out how the linear model estimated when data are pooled, putting in consideration time and firm effects and the variables that are qualitative to adjust for missing information in the linear model using the first case of Griffiths and Anderson.

2. The Model

\[ y_{it} = \alpha + \sum_{k=1}^{K} b_k x_{it} + u_{it} \]

Where:
- \( y_{it} \) is the observations of the dependent variable for firm .i at time t.
- \( X_{it} \) is the observation of the independent variable for firm .i at time t for the k\textsuperscript{th} regressor.
- \( a \) and \( b_k \) are parameters.
- \( u_{it} \) is the error term associated with \( y_{it} \).

To adjust for missing information through the error term we let:
\[ u_{it} = (f_i + g_t + e_{it}) h_i \]

Properties of \( u_{it} \) are:
1. \( E(u_{it}) = 0 \)

2. \( E(u_{it}u_{jt}) = E[(f_i + g_t + e_{it}) h_i][(f_j + g_t + e_{jt}) h_j] = E(f_i f_j h_i h_j) + E(g_t g_t h_i h_j) + E(e_{it} e_{jt} h_i h_j) + E(f_i e_{jt} h_i h_j) + E(g_t e_{it} h_i h_j) + E(e_{it} e_{jt} h_i h_j) = \sigma_{f_i}^2 + \sigma_{g_t}^2 + \sigma_{e_{it}}^2 \)

The variance covariance matrix is:
\[ \mathbf{V} = E(\mathbf{uu}') \]
\[ = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ u_{21} & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & \cdots & u_{NN} \end{pmatrix} \]

Let:
\[ E(u_{it} u_{jt}') = c_{ij} \]
\[ = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1T} \\ u_{21} & u_{22} & \cdots & u_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ u_{T1} & u_{T2} & \cdots & u_{TT} \end{pmatrix} \]
\[ = \begin{pmatrix} \sigma_{f_i}^2 & \sigma_{f_i}^2 & \cdots & \sigma_{f_i}^2 \\ \sigma_{g_t}^2 & \sigma_{g_t}^2 & \cdots & \sigma_{g_t}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{e_{it}}^2 & \sigma_{e_{it}}^2 & \cdots & \sigma_{e_{it}}^2 \end{pmatrix} \]

And:
\[ E(u_{it} u_{jt}') = c_{ij} \]
\[ = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1T} \\ u_{21} & u_{22} & \cdots & u_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ u_{T1} & u_{T2} & \cdots & u_{TT} \end{pmatrix} \]
\[
\begin{pmatrix}
  h_{11} & h_{1j} & 0 & 0 & \ldots & 0 \\
  0 & h_{2j} & h_{2j} & 0 & \ldots & 0 \\
  \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & 0 & \ldots & 0 & h_{jT} \end{pmatrix}
\]

Therefore:
\[
v = \begin{pmatrix}
  c_{11} & c_{12} & \ldots & c_{1N} \\
  c_{21} & c_{22} & \ldots & c_{2N} \\
  \vdots & \vdots & \ddots & \vdots \\
  c_{N1} & c_{N2} & \ldots & c_{NN}
\end{pmatrix}
\]

Which can be written as:
\[
V = H \phi H
\]

Where,
\[
\phi = \sigma_f^2 (I_N \otimes I_T) + \sigma_s^2 (I_N \otimes I_T) + \sigma_e^2 I_{NT}
\]

\[
v^{-1} = H^{-1} \phi^{-1} H^{-1}
\]

Then \(v^{-1}\) can be written as:
\[
v^{-1} = H^{-1} [I - z(z'z + D^{-1})^{-1} z'] H^{-1}
\]

Where,
\[
z = [(I_N \otimes I_T)(I_N \otimes I_T)]
\]

\[
D = \begin{pmatrix}
  \sigma_f^2 & 0 \\
  0 & \sigma_s^2
\end{pmatrix}
\]

And if the variance components are known, the generalized least square estimates are:
\[
\hat{\beta} = (x'v^{-1}x)^{-1}x'v^{-1}y
\]

3. Conclusion

It can be concluded that missing information can be captured through the error term.

4. Recommendation

Other papers can be done in the same topic to cover other methods used by Griffiths and Anderson.

References: