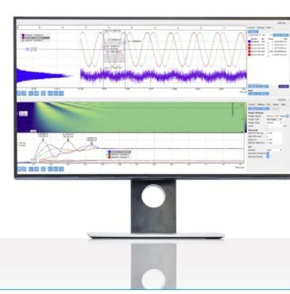
The extended algorithm for quasi maximum likelihood parameter estimation

Cite as: AIP Conference Proceedings **2360**, 020008 (2021); https://doi.org/10.1063/5.0059509
Published Online: 28 September 2021

Alfira Mulya Astuti, Setiawan, Ismaini Zain, et al.



The Extended Algorithm for Quasi Maximum Likelihood Parameter Estimation

Alfira Mulya Astuti[1,2, a)], Setiawan[1)], Ismaini Zain[1)] and Jerry D.T. Purnomo[1)]

[1]*Department of Statistics*, *Institut Teknologi Sepuluh Nopember*, *Jl. Arif Rahman Hakim, Sukolilo, Surabaya 60111, East Java, Indonesia*

[2]*Mathematics Education Department, Universitas Islam Negeri Mataram, Jl. Gajah Mada, Sekarbela, Mataram 83116, West Nusa Tenggara, Indonesia*

[a)] Corresponding author: *alfiramulyastuti@uinmataram.ac.id*

Abstract. This research aims to develop an algorithm of the quasi maximum likelihood estimation method. The model used is the spatial Durbin panel model with dynamic effects for the single equation. The weighted matrix used in this study is equal for each variable This study proposes the generalized estimation equation approach and the Gauss-Newton iteration method for parameter estimation. Modeling of Foreign Direct Investment (FDI) in ASEAN countries from 2010 - 2017 is used for illustration. The best model selection is based on the smallest correlation information criterion (CIC) value.

INTRODUCTION

Regression analysis is an analysis of data that describes the causal relationship between response variables and predictors [1]. Panel data is a type of data used in regression analysis. Regression using panel data is called panel data regression. Regression panel data are widely used in econometric models. Econometric studies continue to develop over time, one of which involves spatial effects. If the model accommodates the links between locations, it is called the spatial econometric model [2]. There are three types of location interaction effects on spatial econometric models, namely: 1) spatial interaction in endogenous variables called spatial autoregressive models (SAR) or spatial lag models, 2) spatial interactions in errors called spatial error models (SEM), and 3) spatial interactions in exogenous variables called spatial Durbin models [3].

The next development of the spatial econometric model of the panel data is the presence of dynamic elements in the model. Models that do not have an offset variable are called static models. The model involving lag variables is called the dynamic model. The static model can only be used for short-term interpretation. While in the dynamic model, interpretation can be done both short and long term. Observations containing dynamic elements and using spatial panel data resulted in autocorrelation. Data containing autocorrelation cannot use regular linear models or Generalized Linear Models (GLM) to model. Liang and Zeger [4] suggest using the Generalized Estimating Equation (GEE) approach to solve the correlated data. This is consistent with the studies conducted by Hardin and Hilbe [5], Anwar [6], and Lu et al [7].

This paper discusses parameter estimation with the extended quasi-maximum likelihood parameter estimation algorithm for the spatial Durbin dynamic panel model in a single equation. The approach used is GEE. The results of the study are presented in the following order. The first session examines the reasons for the development of parameter estimation with the GEE approach. Session 2 presents the spatial dynamic Durbin model on a single equation. Session 3 describes the algorithm development of the QML estimation method. Session 4 illustrates the approach chosen using empirical data of ASEAN countries' FDI for 2010-2017. The last session concludes the study results.

The 6th International Conference on Basic Sciences 2020 (ICBS 2020)
AIP Conf. Proc. 2360, 020008-1–020008-5; https://doi.org/10.1063/5.0059509
Published by AIP Publishing. 978-0-7354-4116-3/$30.00

SPATIAL DURBIN DYNAMIC PANEL MODEL ON SINGLE EQUATIONS.

Debarsy et al [8] have studied the dynamic spatial Durbin model on a single equation. This model was later redeveloped by Lee and Yu [9]. The dynamic spatial Durbin model of a single equation is written on Equation (1).

$$\mathbf{y}_{nt} = \lambda \mathbf{W}_n \mathbf{y}_{nt} + \gamma \mathbf{y}_{n,t-1} + \rho \mathbf{W}_n \mathbf{y}_{n,t-1} + \mathbf{X}_{nt}\boldsymbol{\beta} + \mathbf{W}_n \mathbf{X}_{nt}\boldsymbol{\theta} + \mathbf{c}_n + \alpha_t l_n + \mathbf{v}_{nt}, \quad t = 1, 2, \ldots, T \tag{1}$$

where $\mathbf{y}_{nt} = [y_{1t}, y_{2t}, \ldots, y_{nt}]'$ is the $n \times 1$ column vector of the response variable. $\mathbf{X}_t$ is $n \times k$ matrix of predictor variable. $\mathbf{v}_{nt} = [v_{1t}, v_{2t}, \ldots, v_{nt}]'$ is $n \times 1$ column vector of error. It is assumed to be $IIDN(0, \sigma_v^2 \mathbf{I}_N)$. $\mathbf{W}_n$ is $n \times n$ matrix of spatial weighted. $\mathbf{c}_n$ is $n \times 1$ column vector of individual effect. α_t is time effect. l_n is $n \times 1$ column vector with element being one. $\boldsymbol{\delta} = [\lambda, \gamma, \rho, \boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2]$ are parameters in the model. Equation (1) can be constructed in a reduced form which is written in Equation (2).

$$\mathbf{y}_{nt} - \lambda \mathbf{W}_n \mathbf{y}_{nt} = \gamma \mathbf{y}_{n,t-1} + \rho \mathbf{W}_n \mathbf{y}_{n,t-1} + \mathbf{X}_{nt}\boldsymbol{\beta} + \mathbf{W}_n \mathbf{X}_{nt}\boldsymbol{\theta} + \mathbf{c}_n + \alpha_t l_n + \mathbf{v}_{nt}$$

$$(\mathbf{I}_n - \lambda \mathbf{W}_n)\mathbf{y}_{nt} = \gamma \mathbf{y}_{n,t-1} + \rho \mathbf{W}_n \mathbf{y}_{n,t-1} + \mathbf{X}_{nt}\boldsymbol{\beta} + \mathbf{W}_n \mathbf{X}_{nt}\boldsymbol{\theta} + \mathbf{c}_n + \alpha_t l_n + \mathbf{v}_{nt}$$

$$\mathbf{y}_{nt} = (\mathbf{I}_n - \lambda \mathbf{W}_n)^{-1} [\gamma \mathbf{y}_{n,t-1} + \rho \mathbf{W}_n \mathbf{y}_{n,t-1} + \mathbf{X}_{nt}\boldsymbol{\beta} + \mathbf{W}_n \mathbf{X}_{nt}\boldsymbol{\theta} + \mathbf{c}_n + \alpha_t l_n + \mathbf{v}_{nt}] \tag{2}$$

$(\mathbf{I}_n - \lambda \mathbf{W}_n)^{-1}$ is assumed as invertible. The statistical conditions on the spatial and temporal parameters in the dynamic model must qualify the standard requirements, namely: $-1 < \gamma < 1$ for time effect and $\frac{1}{w_{\min}} < \lambda < 1$ for spatial effect, $w_{\min}$ is the eigenvalue of the weighted matrix. The achievement of a stationary condition in the model when the characteristic roots of the $(\mathbf{I}_n - \lambda \mathbf{W}_n)^{-1}(\gamma \mathbf{I}_n + \rho \mathbf{W}_n)$ matrix should be in the unit circle with the following case [8].

$$\gamma + (\lambda + \rho) w_{\max} < 1 \quad \text{if } \lambda + \rho \geq 0 \tag{3}$$

$$\gamma + (\lambda + \rho) w_{\min} < 1 \quad \text{if } \lambda + \rho < 0 \tag{4}$$

$$\gamma - (\lambda - \rho) w_{\max} > -1 \quad \text{if } \lambda - \rho \geq 0 \tag{5}$$

$$\gamma - (\lambda - \rho) w_{\min} > -1 \quad \text{if } \lambda - \rho < 0 \tag{6}$$

Lee and Yu [9] reviewed parameter estimation on spatial Durbin dynamic panel models using 2SLS and ML methods. The parameters in the model are identified by the moment of relation on the 2SLS and the log likelihood function or the quasi-likelihood function. Furthermore, the model being studied was tested with a Monte Carlo simulation. A different case with Debarsy et al [8] that estimated the parameters using the Bayesian Markov Chain Monte Carlo procedure.

PARAMETER ESTIMATION METHOD

The algorithm used to develop the quasi maximum likelihood estimation method is the generalized estimation equation (GEE) approach. GEE is used to solve correlated data and an extension of quasi score likelihood. The GEE approach models a known function of the marginal expectation of response variables *as* a linear function of one or more predictor variables. GEE describes a random component model for each marginal response with a general link function and variance. The link function of the spatial Durbin dynamic panel model uses the normal distribution link function, i.e: $g[E(\mathbf{y}_{nt})] = \mu$. GEE estimator of $\boldsymbol{\delta}$ is obtained by using quasi score likelihood in Equation (7).

$$\sum_{n=1}^{N} \mathbf{D}_n \mathbf{V}^{-1} (\mathbf{y}_{nt} - E[\mathbf{y}_{nt}]) = \mathbf{0} \tag{7}$$

where $\mathbf{D} = \frac{\partial [E(\mathbf{y}_{nt})]}{\partial \boldsymbol{\delta}}$ is a diagonal matrix of mean response partial derivative. $\mathbf{V} = \phi \mathbf{A}^{\frac{1}{2}} \mathbf{R} \mathbf{A}^{\frac{1}{2}}$. $\mathbf{A}$ is a diagonal matrix with entry being $\mathrm{var}(\mathbf{y}_{nt})$. $\mathbf{R}$ is a working correlation matrix of $\mathbf{y}_{nt}$. ϕ is constant [10]. The general structure of the working correlation matrix used in GEE modeling can be seen in Table 1.

TABLE 1. The Structure of Working Correlation Matrix [11]

Correlation Structure	Corr (Y_{ij}, $Y_{jj'}$)	Sample Matrix	Estimator
Independent	$Corr(Y_{ij}, Y_{ij'}) = \begin{cases} 1 & j = j' \\ 0 & j \neq j' \end{cases}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	
Unstructured	$Corr(Y_{ij}, Y_{ij'}) = \begin{cases} 1 & j = j' \\ \alpha_{jj'} & j \neq j' \end{cases}$	$\begin{pmatrix} 1 & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & 1 & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & 1 \end{pmatrix}$	$\hat{\alpha}_c = \frac{1}{(n-P)\hat{\phi}} \sum_{i=1}^{n} e_{ij} e_{ij'}$
Exchangeable	$Corr(Y_{ij}, Y_{i,j+c}) = \begin{cases} 1 & j = j' \\ \alpha & j \neq j' \end{cases}$	$\begin{pmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{pmatrix}$	$\hat{\alpha} = \frac{1}{(N-P)\hat{\phi}} \sum_{i=1}^{n} \sum_{j>j'} e_{ij} e_{ij'}$ $N = \frac{1}{2} nM(M-1)$
AR (1)	$Corr(Y_{ij}, Y_{i,j+c}) = \alpha^{c}$, $c = 0,1,\ldots,M-j$	$\begin{pmatrix} 1 & \alpha & \alpha^2 \\ \alpha & 1 & \alpha \\ \alpha^2 & \alpha & 1 \end{pmatrix}$	Since $E(e_{ij}, e_{i,j+c}) \approx \alpha^{c}$, α can be estimated by the regression *slope of* $\log(e_{ij}, e_{i,j+c})$ on $\log(c)$

Equation 7 shows that the equation is not closed-form. The closed-form is an estimator that still contains other parameter values. Thus, the estimation is done numerically with the Fisher scoring typed method. This method is an extension of the Gauss-Newton iteration [11]. The algorithm used is contained in Equation (8).

$$\hat{\boldsymbol{\delta}}^{(j+1)} = \hat{\boldsymbol{\delta}}^{(j)} + \left\{ \sum_{i=1}^{N} \mathbf{D}_n^T \left(\hat{\boldsymbol{\delta}}^{(j)}\right) \mathbf{V}_n^{-1} \left(\hat{\boldsymbol{\delta}}^{(j)}\right) \mathbf{D}_n \left(\hat{\boldsymbol{\delta}}^{(j)}\right) \right\}^{-1} \left\{ \sum_{n=1}^{N} \mathbf{D}_n^T \left(\hat{\boldsymbol{\delta}}^{(j)}\right) \mathbf{V}_n^{-1} \left(\hat{\boldsymbol{\delta}}^{(j)}\right) \boldsymbol{\zeta}_n \left(\hat{\boldsymbol{\delta}}^{(j)}\right) \right\}, \tag{8}$$

where $\boldsymbol{\zeta}_n \left(\hat{\boldsymbol{\delta}}^{(j)}\right) = \left(\mathbf{y}_n - E(\mathbf{y}_n)\right)$.

The next step is to determine the sandwich covariance matrix. This matrix is one of the variances that can be used in QMLE and is written in Equation 9.

$$Cov\left(\hat{\boldsymbol{\delta}}\right) = \hat{\mathbf{V}}_{Sand} = \mathbf{B}\left(\hat{\boldsymbol{\delta}}\right)^{-1} \mathbf{M}\left(\hat{\boldsymbol{\delta}}\right) \mathbf{B}\left(\hat{\boldsymbol{\delta}}\right)^{-1} \tag{9}$$

where:

$\mathbf{B}\left(\hat{\boldsymbol{\delta}}\right) = \sum_{n=1}^{N} \mathbf{D}_n^T \left(\hat{\boldsymbol{\delta}}\right) \mathbf{V}_n^{-1} \left(\hat{\boldsymbol{\delta}}\right) \mathbf{D}_n \left(\hat{\boldsymbol{\delta}}\right)$ is "*bread*". $\mathbf{M}\left(\hat{\boldsymbol{\delta}}\right) = \sum_{n=1}^{N} \mathbf{D}_n^T \left(\hat{\boldsymbol{\delta}}\right) \mathbf{V}_n^{-1} \left(\hat{\boldsymbol{\delta}}\right) Cov(\mathbf{y}_n) \mathbf{V}_n^{-1} \left(\hat{\boldsymbol{\delta}}\right) \mathbf{D}_n \left(\hat{\boldsymbol{\delta}}\right)$ is "*meat*".

$Cov(\mathbf{y}_n) = [\mathbf{y}_n - E(\mathbf{y}_n)][\mathbf{y}_n - E(\mathbf{y}_n)]^T$ and $E(\mathbf{y}_n)$ is mean of $\mathbf{y}_n$ [11].

Correlation Information Criteria (CIC) is the model selection criteria to show the best working covariance structure in the spatial Durbin dynamic panel model. Equation 10 is the formula used to determine the CIC value.

$$CIC(R) = tr\left(\boldsymbol{\Omega} \hat{\mathbf{V}}_{Sand}\right) = tr\left(\mathbf{B}(\boldsymbol{\delta}) \hat{\mathbf{V}}_{Sand}\right) \tag{10}$$

THE EMPIRICAL STUDY

The empirical study in this research is to model FDI in ASEAN countries. Data sourced from World Bank Group. The number of years is 8 years, starting from 2010 - 2017. The research unit is ASEAN countries. The variable used in this study is FDI as an endogenous variable. Gross Domestic Product (GDP) and Trade openness (Trade) as exogenous variables. FDI data in the previous year is an endogenous lag variable. The weighting matrix is formed based on the consideration of export-import activities between ASEAN countries. The model used can be seen in Equation 11.

$$\ln FDI_{nt} = \lambda \mathbf{W}_n \ln FDI_{nt} + \gamma \ln FDI_{n,t-1} + \rho \mathbf{W}_n \ln FDI_{n,t-1} + \beta_1 \ln GDP_{nt} + \beta_2 \ln Trade_{nt} + \theta_1 \mathbf{W}_n \ln GDP_{nt} + \theta_2 \mathbf{W}_n \ln Trade_{nt} + \mathbf{c}_{n,FDI} + \alpha_{FDI,t} l_n + \mathbf{v}_{nt} \quad (11)$$

where $-1 < \gamma < 1$ dan $\lambda + \gamma + \rho < 1$.

Estimation of parameters in Equation 11 uses the GEE approach with the Fisher scoring typed method and Gauss-Newton iteration assistance. Working correlation used is independent, exchangeable, and autoregressive. Next, the estimation results were compared with the MLE method. The analysis results are presented in Table 2.

TABLE 2. The Parameter Estimation Results

Variable	QMLE			MLE		
	Coefficient	Standard Error	p-value	Coefficient	Standard Error	p-value
lambda	-0.066	0.108	0.55000	-0.270	0.175	0.20409
ln_FDI_t1	0.571	0.108	1.3e-07 ***	0.719	0.086	< 2e-16***
W_ln_FDI_t1	-0.321	0.145	0.02710 **	0.110	0.208	0.59456
ln_GDP	0.377	0.103	0.00027 ***	0.200	0.079	0.01182**
ln_TRADE	0.261	0.161	0.10374*	0.018	0.092	0.84460
W_ln_GDP	0.508	0.201	0.01151**	0.176	0.157	0.26298
W_ln_TRADE	-1.143	0.530	0.03106**	0.045	0.181	0.80264

Signif.codes: '***' = 1%; '**' = 5%; '*'= 11%.

Table 2 shows that the FDI model resulting from the development of the QML estimation algorithm is better than the MLE method. The best model is determined based on the smallest CIC value from the selected working correlation. Table 3 is the CIC value based on the analysis results.

TABLE 3. CIC Value

Working Covariance	CIC Value
Independent	7,46
Exchangeable	7,56
Autoregressive	6,25

Table 3 determines that the best model is the GEE model with autoregressive working correlation. This is because the autoregressive working correlation produces the smallest CIC value. In addition, autoregressive considers the time element. This is consistent with the characteristics of panel data. Therefore, FDI is modeled by the equation written in Equation 12. $\mathbf{c}_n$ is an individual (countries) effect. $\alpha_{t,FDI}$ is time (years) effect. $\mathbf{c}_n$ and $\alpha_{t,FDI}$ value is shown in Table 4.

$$\ln FDI_t = -0,066\mathbf{W} \ln FDI_t + 0,571 \ln FDI_{t-1} - 0,321\mathbf{W} \ln FDI_{t-1} + 0,377 \ln GDP_t + 0,261 \ln Trade_t + 0,508\mathbf{W} \ln GDP_t - 1,143\mathbf{W} \ln Trade_t + \mathbf{c}_{n,FDI} + \alpha_{t,FDI} l_n \quad (12)$$

TABLE 4. The Country and Time Effect

Country	Effect	Time	Effect
Brunei Darussalam	0.2351	2011	-0.2141
Cambodia	-0.1340	2012	0.0797
Indonesia	0.1204	2013	0.3309
Lao PDR	0.0207	2014	-0.1672
Malaysia	-0.1335	2015	-0.0715
Myanmar	0.0366	2016	0.0110
Philippines	-0.2458	2017	0.0311
Singapore	0.4113		
Thailand	0.0213		
Vietnam	-0.3322		

CONCLUSION

The development of the QML estimation method algorithm produces better model than the MLE method. The best model chosen is a model with autoregressive working correlation. The regression coefficient of the resulting model did not change even though it used a different working correlation. However, it still produces a different conclusion.

REFERENCES

1. D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression* (John Wiley and Son, New York, 2000).
2. A. M. Astuti, Setiawan, I. Zain and J. D. Purnomo, Journal of Physics: Conference Series, **1490(012032)**, 1-13 (2020).
3. J. P. Elhorst, *Spatial Econometrics: From Crossectional to Panels Data* (Spinger Heidelberg, New York, 2014).
4. K. Y. Liang and S. L. Zeger, Biometrika, **73(1)**, 13-22 (1986).
5. J. W. Hardin and J. M. Hilbe, *Generalized Estimating Equations 2nd Edition* (Chapman and Hall/CRC, New York, 2012).
6. N. Anwar, "Pemodelan Tingkat Pengangguran di Lima Negara Anggota Asean dengan Regresi Data Panel dan Generalized Estimating Equation," (Institut Pertanian Bogor, Bogor, 2012).
7. C. Lu, W. Wang and J. M. Wooldridge, "Using Generalized Estimating Equations to Estimate Nonlinear Models with Spatial Data," (Cornell University, New York, 2018).
8. N. Debarsy, C. Ertur, and J.P. LeSage, Statistical Methodology, **9**, 158-171 (2012).
9. Lee LF., Yu J., Journal of Applied Econometrics, (2015).
10. A. J. Dobson, *Introduction to Generalized Linear Models, Second Edition* (CRC Press, New York, 2002).
11. J. D. T. Purnomo, "A Modified Generalized Estimating Equation (GEE) Approach for Latent Class Models with Covariate Effects on Measured and Underlying Variables," (Hsinchu, Taiwan, Republic of China, 2018).