Journal of Economics and Administrative Sciences Vol.28 (NO. 134) 2022, pp. 96-109

Journal of Economics
and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq

Using Some Estimation Methods for Mixed-Random Panel Data Regression Models with Serially Correlated Errors with Application

Musaab Amer Rajab
Department of Informatics, Administrative office, and Systems, Ministry of Electricity,
Baghdad, Iraq
Mosab.Aamer1201a@coadec.uobaghdad.edu.iq

Mohammed Sadeq Abdul.Razaq
Department of Statistics, College of Administration and Economics, University of Baghdad,
Baghdad, Iraq
dr_aldouri@coadec.uobaghdad.edu.iq

Received: 12/12/2022 **Accepted:** 29/12/2022 **Published:** December / 2022

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)

Abstract

This research includes the study of dual data models with mixed random parameters, which contain two types of parameters, the first is random and the other is fixed. For the random parameter, it is obtained as a result of differences in the marginal tendencies of the cross sections, and for the fixed parameter, it is obtained as a result of differences in fixed limits, and random errors for each section. Accidental bearing the characteristic of heterogeneity of variance in addition to the presence of serial correlation of the first degree, and the main objective in this research is the use of efficient methods commensurate with the paired data in the case of small samples, and to achieve this goal, the feasible general least squares method (FGLS) and the mean group method (MG) were used, and then the efficiency of the extracted estimators was compared in the case of mixed random parameters and the method that gives us the efficient estimator was chosen. Real data was applied that included the per capita consumption of electric energy (Y) for five countries, which represents the number of cross-sections (N = 5) over nine years (T = 9), so the number of observations is (n = 45) observations, and the explanatory variables are the consumer price index (X1) and the per capita GDP (X2). To evaluate the performance of the estimators of the (FGLS) method and the (MG) method on the general model, the mean absolute percentage error (MAPE) scale was used to compare the efficiency of the estimators. The results showed that the mean group estimation (MG) method is the best method for parameter estimation than the (FGLS) method. Also, the (MG) appeared to be the best and best method for estimating sub-parameters for each cross-section (country).

Keywords: FGLS estimation method, mixed-stochastic parameter regression model, first-order serial correlation, (MG) estimation method.

1. Introduction

To study any of the economic, social, medical or other phenomena that the researcher chooses in his study, he must provide data for that phenomenon from solid and reliable sources, when studying a specific phenomenon during a specific time period, time serious data must be because collected, and this serious may include an autocorrelation problem because it is unstable. In this case, the general least square (GLS) method should be used to estimate the model parameters.

And when studying a certain phenomenon for several sectors of different groups, it is necessary to collect cross-sectional data, which in most cases is a problem of heterogeneity of error variance, so the weighted least square method (WLS) should be used to estimate the modeling parameters. And those random errors in both types of data above are considered the main reason for the occurrence of problems in the data. Instead of analyzing each type of data above separately, in which the researcher may obtain inefficient estimates, it required obtaining another type of data by merging the two types of data above and obtaining what is called Panel Data. Most of the research relied on estimating the parameters and testing them for the panel data on two methods: the generalized least square (GLS) when the variance-covariance matrix is known and the (FGLS) method when the variance-covariance matrix is unknown, and it is one of the methods adopted in this research to estimate the model parameters (Basim Shaliba Muslim 2009, Al-Mafarji 2018).

For example, the phenomenon of the spread of a particular disease in a certain country is classified according to the regions or cities in that country and measured for a specific period, accordingly, the observations of this phenomenon at the level of each city represent the cross-sectional data, while the observations during a period of time for each city and during a certain period of time represent the time series data for example, the phenomenon of the spread of a particular disease in a particular country classified according to the regions or cities in that country and measured for a specific period, accordingly, the observations of this phenomenon at the level of each city represent the cross-sectional data, while the observations during a period of time for each city and during a certain period of time represent the time series data (Kazem & Muslim 2002) .

And the importance of statistical analysis of this type of data is to assess the effects of the explanatory variables on the dependent variable during the specified time period, and the efficient estimation of the model parameters is a major goal in the analysis of the dual data, and that the data collection process in this way leads to obtaining accurate parameters that represent the study population in a way reliable and correct, due to taking into account the time factor and the existence of a correlation between the sample items (Reem 2021).

Among the most important previous studies on the subject, the researchers (Ahmed et al.) in 2009 presented research that included testing panel data models when the regression coefficients are fixed, random, and mixed, where they used simulation to make comparisons between the behavior of several estimation methods, such as random coefficient regression (RCR), classical pooling (CP), and Mean group estimators (MG). In the three cases of regression coefficients, simulation results indicated that (RCR) estimators perform well in the case of small samples if the coefficients are random, while (CP) estimators work well in the case

of the fixed model only, but (MG) estimators work fine if the transactions are random or fixed.

Also, the researcher (Mohammed) presented in 2018 a research that included the study of panel data models when the errors are serially correlated to the first order as well as with the parameters of random regression, and the (GLS) method was used to estimate the parameters when the samples are small, and the researcher suggested an alternative estimator It is the mean group estimator (MG), and the researcher made comparisons of the efficiency of the (GLS) and (MG) estimators. The simulation study conducted by the researcher indicated that the (MG) method is the best and most reliable method than the (GLS) method, especially when the model includes random and fixed estimators. It means a model that contains random- mixed parameters.

In this paper, the parameters of the panel data model with mixed stochastic parameters will be estimated, and these models include mixed parameters, that is, some of them are random and the other is non-random (fixed), that is (β_{1i}) is a vector for the parameters that are supposed to be random, and (β_2) is a vector for the parameters that are supposed to be non-random (fixed).

The main objective of this research is to identify some notes on how to choose a good estimator for panel data when the sample size is small, and the errors are serially correlated of the first order, as well as with mixed random regression parameters.

2. Materials and Methods

2.1 Mixed-Stochastic Parameter Regression of Panel Data Models (MSPR)

In this type of model, the GLS will be destined to the model when the features are mixed i.e. each other random and the other non-random (fixed) when such a situation occurs, the mixed random model is written as follows:

$$y_{it} = \sum_{k=1}^{K} \beta_{ki} x_{kit} + u_{it}$$

$$y_{it} = x_{it}\beta_i + u_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T$$

$$y_i = X_i\beta_i + u_i$$

$$y_i = X_{1i}\beta_{1i} + X_{2i}\beta_2 + u_i \qquad \dots (1)$$

Where y_i & u_i is:

y_i: A Vector of order (T*1) from the observations of the department variable for section (i).

u_i: A Vector of order (T*1) for the random errors of section (i).

And that:

$X_i = (X_{1i}, X_{2i})$

They are the matrices of the observations on the independent variables

$X_{1i} = T \times K_1$ & $X_{2i} = T \times K_2$, where K_1 & K_2 respectively.

K_1 : Explanatory variables for random parameters (β_{1i})

K_2 : Explanatory variables for fixed parameters (β_2)

And that:

$$\beta_i = \left(\beta'_{1i}, \beta'_2\right)'$$

where (β_{1i}) is the (K_1 X1) vector for parameters that are assumed to be random with an average of ($\overline{\beta}_1$) and the variance-covariance matrix is ($\gamma_{\beta 1}$), but (β_2) is the (K_2 X1) vector for parameters that are assumed to non-random (fixed), where:

$$K = K_1 + K_2$$

Model (1) applies to each cross section under the assumption that:

$$\beta_{1i} = \overline{\beta}_1 + \pi_{\beta 1} \qquad \dots (2)$$

And we can merge (N) from the sectional equations as follows:

$$Y = G\overline{\beta} + \tau \qquad \dots (3)$$

Whereas

$$X = (X_1', \dots, X_N')', \overline{\beta} = \left(\overline{\beta}_1{}', \beta_2'\right)', \tau = (\tau_1', \dots, \tau_N')', \tau_i = X_{1i}\pi_{\beta 1} + u_i$$

Under (Swamy 1970) assumptions, this model has been examined by (Swamy 2012) and (Rosenberg 1973); we examine this model under our assumptions from (1) to (4):

Assumption (1): $E(u_i) = 0; \forall i = 1, 2, \dots, N$

Assumption (2): Explanatory variables are not random (in repeated samples), then we assume independence with other variables in the model and the value of the rank $(X_i'X_i) = K; \quad \forall i = 1, \dots, N, \quad where\ K < T, N$

Assumption (3): The errors have a constant variance for each individual (cross section), but there is a problem of heterogeneity of the variation in the cross sections, in addition to being serially correlated of the first order, meaning that the random error for each period depends linearly on the random error of the previous periods.

$$u_{it} = \phi_i u_{i,t-1} + \varepsilon_{it}; \quad |\phi_i| < 1, \qquad \text{where } \phi_i \text{ for } i = 1, \dots, N$$

$$E(\varepsilon_{it}) = 0,$$

$$E\left(u_{i,t-1}\varepsilon_{jt}\right) = 0; \quad \forall i, j, \text{ and } t. \text{ And}$$

$$E\left(\varepsilon_{it}\varepsilon_{js}\right) = \begin{cases} \sigma^2_{\varepsilon_i} & if\ t = s; i = j \\ 0 & otherwise \end{cases} \qquad i, j = 1, \dots, N; \quad t, s = 1, \dots, T$$

It is assumed that the errors in the initial or primary time period have the same characteristics as in the subsequent periods, so we assume that

$$E\left(u^2_{i0}\right) = \sigma^2_{\varepsilon_i}/1 - \phi_i^2; \quad \forall i.$$

Assumption (4): The regression model feature vector is determined as $\beta_i = \overline{\beta} + \pi_i$ **where** $\overline{\beta} = \left(\overline{\beta}_1, \dots, \overline{\beta}_K\right)'$

$\overline{\beta}$**: It is a vector of non-random (fixed) parameters of order (K * 1) estimated by the method of least squares (OLS).**

$$\pi_i = (\pi_{i1}, \dots, \pi_{iK})'$$

π_i**: The random error vector of the parameters is of order (K*1).**

β_i**: A vector of order (K*1).**

$$\text{And } E\left(\pi_i\pi_j'\right) = \begin{cases} \gamma^* & if\ i = j \\ 0 & if\ i \neq j \end{cases} \quad i, j = 1, \dots, N; \quad k = 1, \dots, K$$

, *For k=1,…,K,* and also assume that $E\left(\pi_i u_{jt}\right) = 0, \forall i \text{ and } j$

$$\gamma^* = diag\{\gamma_k^*\}$$

So the variance-covariance matrix to (τ) is:

$$E(\tau\tau') = V + Z_{\beta_1}\left(I_N \otimes \gamma_{\beta 1}\right) Z'_{\beta_1} \qquad \dots (4)$$

$$E(\tau\tau') = \Pi$$

whereas

$$Z_{\beta_1} = diag\{X_{1i}\}$$

$$Z_{\beta_1} = \begin{pmatrix} X_{11} & 0 & \cdots & 0 \\ 0 & X_{12} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_{1N} \end{pmatrix}$$

And that the estimator (GLS) to ($\overline{\beta}$) is:

$$\widehat{\beta}_{MSPR-SC} = (X'\Pi^{-1}X)^{-1}X'\Pi^{-1}Y$$

$$\widehat{\beta}_{MSPR-SC} = \begin{pmatrix} X_1'\Pi^{-1}X_1 & X_1'\Pi^{-1}X_2 \\ X_2'\Pi^{-1}X_1 & X_2'\Pi^{-1}X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1'\Pi^{-1}Y \\ X_2'\Pi^{-1}Y \end{pmatrix} \quad \dots (5)$$

Whereas

$$X_1 = (X_{11}', \dots, X_{1N}')', X_2 = (X_{21}', \dots, X_{2N}')'$$

It should be noted that the mixed random model is a special case of a random model when the variances in some parameters are assumed to be zero (Reda Abonazel 2018).

2.2 Mixed-Stochastic Parameter Panel Data Regression Model Algorithm:

The feasible estimator ($\widehat{\beta}_{MSPR-SC}$) can be obtained by the following algorithm (Abonazel 2018):

- The first step/ is to calculate ($\hat{\gamma}^*$) in the stochastic parameter model using the coordinate estimator for ($\hat{\sigma}^2_{\varepsilon_i}$ & Ω_{ii}) are calculated as in the following formula:

$$\gamma^* = \left[\frac{1}{N-1}\left(\sum_{i=1}^{N} \beta_i^* \beta_i^{*'} - \frac{1}{N}\sum_{i=1}^{N} \beta_i^* \sum_{i=1}^{N} \beta_i^{*'}\right)\right] - \frac{1}{N}\sum_{i=1}^{N} \sigma^2_{\varepsilon_i}\left(X_i'\Omega_{ii}^{-1}X_i\right)^{-1} \dots (6)$$

$$\beta_i^* = \left(X_i'\Omega_{ii}^{-1}X_i\right)^{-1} X_i'\Omega_{ii}^{-1}y_i \quad \dots (7)$$

And that the coordinated estimates for (ϕ_i & $\sigma^2_{\varepsilon_i}$) are calculated according to the following formulas:

$$\widehat{\phi}_i = \frac{\sum_{t=2}^{T} \hat{u}_{it}\hat{u}_{i,t-1}}{\sum_{t=2}^{T} \hat{u}^2_{i,t-1}} \quad \dots (8)$$

$$\hat{\sigma}^2_{\varepsilon_i} = \frac{\hat{\varepsilon}_i'\hat{\varepsilon}_i}{T-K'} \quad \dots (9)$$

Likewise, each of ($\hat{\varepsilon}_{i1}$) & ($\hat{\varepsilon}_{it}$) is calculated according to the following formulas:

$$\hat{\varepsilon}_{i1} = \hat{u}_{i1}\sqrt{1-\widehat{\phi}_i^2} \quad \dots (10)$$

$$\hat{\varepsilon}_{it} = \hat{u}_{it} - \widehat{\phi}_i\hat{u}_{i,t-1} \qquad \text{for } t = 2, \dots, T \qquad \dots (11)$$

- The second step/ find the estimate of ($\gamma_{\beta 1}$) and let it be ($\hat{\gamma}_{\beta 1}$) by eliminating the rows and columns of the fixed parameter (the one inside the vector) from the matrix ($\hat{\gamma}^*$).
- The third step/ is to find an estimate of (Π) and let it be ($\widehat{\Pi}$), using ($\hat{\gamma}_{\beta 1}$) and the consistent estimator in the equation (8) & (9).
- The fourth step/ Obtaining the feasible estimator ($\widehat{\beta}_{MSPR-SC}$) for the mixed random parameter model using ($\widehat{\Pi}$) in equation (5).

And for how to determine the fixed parameters in the model, the test (Swamy) will be used for the randomness of the parameters, and this test is performed because (π_i) is constant for each (i), as in assumption (4), and as a result, we can test the random variance indirectly by test whether fixed parameter vectors are equal. That is, (H_0) be:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_N = \overline{\beta}$$

H_1: At least the two are not equal
The statistical test is:

$$S = \sum_{i=1}^{N} \left(\hat{\beta}_i - \hat{\beta}_S\right)' \frac{x_i' x_i}{\hat{\sigma}_i^2} \left(\hat{\beta}_i - \hat{\beta}_S\right) \quad \dots (12)$$

Whereas

$$\hat{\hat{\beta}}_S = \left[X'\left(\hat{\Sigma}_H \otimes I_T\right)^{-1} X\right]^{-1} \left[X'\left(\hat{\Sigma}_H \otimes I_T\right)^{-1} Y\right] \quad \dots (13)$$

where($\hat{\Sigma}_H$) is the estimated matrix of (Σ_H) and (Swamy) (Swamy 1970) it is by the null hypothesis H_0, the test statistic in (12) is asymptotic distributed on a chi-square with degrees of freedom K (N-1), such as T, N→∞ it is fixed.

Swamy's test can be applied to the mixed stochastic parameter regression as in the (SPR) model. At first, it is assumed that the mixed stochastic parameter regression form in (1) can be rewritten as:

$$y_i = Q_{1i} b_{1i} + Q_{2i} b_{2i} + X_{2i} \beta_2 + u_i \quad \dots (14)$$

Whereas

$\beta_{1i} = (b_{1i}', b_{2i}')'$

And that

b_{1i} is a vector (h_1X1) for random features to be included in some hypothesis test.
b_{2i} is a vector for random features, but these must be excluded from the test.

$X_{1i} = (Q_{1i}', Q_{2i}')'$

And (Q_{1i}) & (Q_{12}) are (T X h_1) & (T X h_2) are respectively matrices from the observations of the explanatory variables and the rest of the other symbols were defined when discussing the equation (1) is discussed. From the above, the random-mixed model can be rewritten in the following way:

$$Y = Q_1 \bar{b}_1 + Q_2 \bar{b}_2 + X_2 \beta_2 + \tau \quad \dots (15)$$

Whereas

τ & X_2 Are defined in (5) and (3) respectively.
Y: A vector of order (N*1) from the observations of the approved variables for all cross-sections.

$Q_1 = (Q_{1i}', \dots, Q_{1N}')', Q_2 = (Q_{2i}', \dots, Q_{2N}')'$

($\bar{b}_1$) & ($\bar{b}_2$) are the mean of the random features (b_{1i}) & (b_{2i}) consecutively.
And it is possible to conduct a test for the randomness of the parameters in the random-mixed model according to the following hypothesis:

H_0 : b_{11}= … = b_{1N}= $\bar{b}_1$
H_1: At least two are not equal

This test is similar to the indirect test for the randomness of the random parameter model, and here we may have a set of sub-parameters that were initially assumed to be random and these parameters will be tested, and the following test statistics will be used:

$$\sum_{i=1}^{N} \left(\hat{b}_{1i} - \hat{\hat{b}}_1\right)' \frac{Q_{1i}' Q_{1i}}{\hat{\sigma}_i^2} \left(\hat{b}_{1i} - \hat{\hat{b}}_1\right) \quad \dots (16)$$

Whereas

$\hat{\hat{b}}_1$ is the vector estimated for the features as a fixed.
$\hat{b}_{1i}$ for I = (1,…,N) are discrete estimates for the features.
When the null hypothesis is accepted, the parameters are fixed, and when it is rejected, the parameters (b_{1i}) are random.

To calculate the sub-parameters for each cross-section, i.e. for each country, the best unbiased linear BLUE estimator in stochastic parameter regression with serial correlation ($\hat{\hat{\beta}}_{SPR-SC}$) is (Abonazel 2018):

$$\hat{\hat{\beta}}_{SPR-SC} = (X'\Lambda^{*-1}X)^{-1}X'\Lambda^{*-1}Y \quad \dots (17)$$

And the variance - covariance matrix is:

$$var\left(\hat{\hat{\beta}}_{SPR-SC}\right) = (X'\Lambda^{*-1}X)^{-1} \quad \dots (18)$$

Whereas:

$$\Lambda^{*} = V + Z(I_N \otimes \gamma^{*})Z' \quad \dots (19)$$

$$V = \begin{pmatrix} \sigma^2_{\varepsilon_1}\Omega_{11} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\varepsilon_2}\Omega_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2_{\varepsilon_N}\Omega_{NN} \end{pmatrix} \quad \dots (20)$$

$$\Omega_{ii} = \frac{1}{1-\phi_i^2}\begin{pmatrix} 1 & \phi_i & \phi_i^2 & \cdots & \phi_i^{T-1} \\ \phi_i & 1 & \phi_i & \cdots & \phi_i^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_i^{T-1} & \phi_i^{T-2} & \phi_i^{T-3} & \cdots & 1 \end{pmatrix} \quad \dots (21)$$

And that

$$\gamma^{*} = \left[\frac{1}{N-1}\left(\sum_{i=1}^{N} \beta_i^{*}\beta_i^{*'} - \frac{1}{N}\sum_{i=1}^{N} \beta_i^{*}\sum_{i=1}^{N} \beta_i^{*'}\right)\right] - \frac{1}{N}\sum_{i=1}^{N} \sigma^2_{\varepsilon_i}\left(X_i'\Omega_{ii}^{-1}X_i\right)^{-1} \quad \dots (22)$$

$$\beta_i^{*} = \left(X_i'\Omega_{ii}^{-1}X_i\right)^{-1}X_i'\Omega_{ii}^{-1}y_i \quad \dots (23)$$

To make the ($\hat{\hat{\beta}}_{SPR-SC}$) estimator feasible, the following consistent estimates of (ϕ_i) and ($\hat{\sigma}^2_{\varepsilon_i}$) are used:

$$\hat{\phi}_i = \frac{\sum_{t=2}^{T} \hat{u}_{it}\hat{u}_{i,t-1}}{\sum_{t=2}^{T} \hat{u}^2_{i,t-1}} \quad \dots (22)$$

$$\hat{\sigma}^2_{\varepsilon_i} = \frac{\hat{\varepsilon}_i'\hat{\varepsilon}_i}{T-K'} \quad \dots (23)$$

And that

$$\hat{u}_i = (\hat{u}_{i1}, \dots, \hat{u}_{iT})', \hat{u}_i = y_i - X_i\hat{\beta}_i \;, \hat{\beta}_i = (X_i'X_i)^{-1}X_i'y_i \,,$$

While

$$\hat{\varepsilon}_i = (\hat{\varepsilon}_{i1}, \hat{\varepsilon}_{i2}, \dots, \hat{\varepsilon}_{iT})', \hat{\varepsilon}_{i1} = \hat{u}_{i1}\sqrt{1-\hat{\phi}_i^2}\,, \hat{\varepsilon}_{it} = \hat{u}_{it} - \hat{\phi}_i\hat{u}_{i,t-1} \qquad \text{for } t = 2, \dots, T$$

It should be noted that ($\hat{\hat{\beta}}_{SPR-SC}$) can be rewritten as a weighted average estimator (GLS) for each cross section (Abonazel 2018).

$$\hat{\hat{\beta}}_{SPR-SC} = \sum_{i=1}^{N} W_i^{*}\beta_i^{*} \quad \dots (24)$$

And that

$$W_i^{*} = \left\{\sum_{i=1}^{N}\left[\gamma^{*} + \sigma^2_{zi}\left(X_i'\Omega_{ii}^{-1}X_i\right)^{-1}\right]^{-1}\right\}^{-1}\left\{\sum_{i=1}^{N}\left[\gamma^{*} + \sigma^2_{\varepsilon_i}\left(X_i'\Omega_{ii}^{-1}X_i\right)^{-1}\right]^{-1}\right\} \quad \dots (25)$$

It turns out that ($\hat{\hat{\beta}}_{SPR-SC}$) in formula (24) is a weighted average of the (OLS) estimates for a given cross-section. Finally, the formula (24) benefits from the fact that (Poi 2003):

$$(A + BDB')^{-1} = A^{-1} - A^{-1}BEB'^{A^{-1}} + A^{-1}BE(E + D)^{-1}EBA'^{-1} \quad \dots (26)$$

And that

(A) And (D) are non-singular matrices of (m*n) degree, and (B) a matrix of (m*n) degree (Rao, et.al, 1973, p.33).

$E = (B'A^{-1}B)^{-1}$(Rao, et.al, 1973, p.33) … (27)

In addition to the estimation of ($\hat{\hat{\beta}}_{SPR-SC}$), the researcher often wishes to obtain estimations of the (β_i) vectors of cross-sections as well, if the interest is limited to the class of estimators (β_i^*) for which it is (Griffiths, et.al, 541):

$E(\hat{\beta}_i^*/\beta_i) = \beta i$

And an estimator (OLS) for a single cross section (b_i) is appropriate. However, if there is no condition on (β_i), the best unbiased linear estimator is:

$$\hat{\beta}_i = \hat{\beta} + \gamma^* x'_i\,(x_i\gamma^* x'_i + \sigma_{ii} I)^{-1}(y_i - X_i\hat{\beta})$$

$$\hat{\beta}_i = (\gamma^{*-1} + \sigma_{ii}{}^{-1}X_i'X_i)^{-1}\left(\sigma_{ii}{}^{-1}X_i'X_i b_i + \gamma^{*-1}\hat{\beta}\right) \quad \dots (27)$$

To obtain the variance ($\hat{\beta}_i$) , Green (1997,672) suggested the formula (30):

$$\hat{\beta}_i = [A_i \quad (I - Ai)]\begin{bmatrix}\hat{\beta} \\ b_i\end{bmatrix} \quad \dots (28)$$

Whereas:

$$A_i = (\gamma^{*-1} + \sigma_{ii}{}^{-1}X_i'X_i)^{-1}\,\gamma^{*-1} \quad \dots (29)$$

$$\text{Var}(\hat{\beta}_i) = [A_i \quad (I - A_i)]\ \text{Var}\begin{pmatrix}\hat{\beta} \\ b_i\end{pmatrix}\begin{bmatrix}A_i' \\ (I - A_i)'\end{bmatrix} \quad \dots (30)$$

Whereas:

$$\text{Var}\begin{pmatrix}\hat{\beta} \\ b_i\end{pmatrix} = \begin{bmatrix}\text{Var}(\hat{\beta}) & \text{Cov}(\hat{\beta}, b_i) \\ \text{Cov}(\hat{\beta}, b_i) & \text{Var}(b_i)\end{bmatrix} \quad \dots (31)$$

The estimator ($\hat{\beta}$) using the (GLS) method is consistent and effective. According to (Lemma 2.1) in (Hausman 1978):

Asymptotically $\text{Cov}(\hat{\beta}, b_i)$ = Asymptotically $\text{Var}(\hat{\beta})$ – Asymptotically $\text{Cov}(\hat{\beta}, \hat{\beta} - b_i) = \text{Var}(\hat{\beta})$

After doing some mathematical operations, we get:

Asymptotically $\text{Var}(\hat{\beta}_i) = \text{Var}(\hat{\beta}) + (I - A_i)\{\text{Var}(b_i) - \text{Var}(\hat{\beta})\}(I - A_i)'$

And to obtain the feasible estimations of the above formulas, each (σ_{ii}) it can be offset by an OLS estimate (Poi, 2003):

$$\widehat{\sigma_{ii}} = \frac{(y_i - X_i b_i)'(y_i - X_i b_i)}{T_i - K} \quad \dots (32)$$

2.3 Feasible Generalized Least Square (FGLS)

The estimators of (MSPR-SC) need to estimate the elements of matrices (variance-covariance) because they are unknown and to make these estimators feasible, it is suggested to use the following consistent estimators: (ϕ_i) and ($\sigma_{\varepsilon_i}^2$) (Reda Abonazel, 2018):

$$\hat{\phi}_i = \frac{\sum_{t=2}^{T}\hat{u}_{it}\hat{u}_{i,t-1}}{\sum_{t=2}^{T}\hat{u}_{i,t-1}^2} \quad \dots (33)$$

$$\hat{\sigma}_{\varepsilon_i}^2 = \frac{\hat{\varepsilon}_i'\hat{\varepsilon}_i}{T - K'} \quad \dots (34)$$

where

$\hat{u}_i = (\hat{u}_{i1}, \dots, \hat{u}_{iT})' = y_i - X_i\hat{\beta}_i$, $\hat{\beta}_i = (X_i'X_i)^{-1}X_i'y_i$,

While $\hat{\varepsilon}_i = (\hat{\varepsilon}_{i1}, \hat{\varepsilon}_{i2}, \dots, \hat{\varepsilon}_{iT})'$, $\hat{\varepsilon}_{i1} = \hat{u}_{i1}\sqrt{1 - \hat{\phi}_i^2}$, $\hat{\varepsilon}_{it} = \hat{u}_{it} - \hat{\phi}_i\hat{u}_{i,t-1}$ for $t = 2, \dots, T$

For the estimator of (MSPR-SC), the consistent estimator for (Π) say ($\widehat{\Pi}$) was proposed by Abonazel (Abonazel 2018) to obtain the feasible estimator for it.

2.4 Mean Group Estimator (MG)

(Abo Nazel)(Abonazel 2019, Abonazel 2018) suggest using an estimator (MG) as an alternative estimator for the general random regression model is defined as follows:

$$\bar{\beta}_{SMG} = \frac{1}{N}\sum_{i=1}^{N} \hat{\beta}_i \qquad \dots (35)$$

You notice that this estimate is average of the ordinary least square (OLS) estimates, which is ($\hat{\beta}_i$).

And for easy verification the (MG) estimator is constant with ($\bar{\beta}$) when both are (N, T → ∞), (Abonazel 2018) showed the statistical properties of the (MG) estimator.

3. Applications and Discussion of Results

3.1 Description of Data

The data studied in this paper represents the per capita share of electric energy consumption as a dependent variable and the mean explanatory variables affecting it, which are the per capita share of the gross domestic product and the consumer price index; it is panel data that includes five countries, namely Iraq, and its comparison with neighboring countries (N=5), which represent cross-sections measured over nine years (T=9) which in turn represent the time series.

Global ESCWA committee of the United Nations and ESCWA is among the committees that work under the supervision of the economic and social council, and the ESCWA committee was established by the economic commission for western Asia to stimulate the economic activity of several of countries.

3.2 Analysis of Data

In this section, the results of real data analysis will be presented, which were represented by the per capita share of electric energy consumption (Y), the consumer price index (X_1), and the per capita gross domestic product (X_2) for five countries; the estimations of the parameters of the model were extracted and compared between them through the measure the mean absolute percentage error (MAPE).

After describing the data, the mixed model will be tested according to formula (1), in which it is assumed that (β_{1i}) is a vector of random parameters and (β_2) is a vector of parameters that are assumed to be non-random (fixed), where the (Swamy) [3] test was applied to the mixed model according to the following hypothesis for the randomness of the parameters:

$H_0 : b_{11} = \dots = b_{1N} = \bar{b}_1$

H_1**: At least are two not equal**

According to the above hypothesis, the parameter vector (β_{1i}) will be tested for randomness according to the test statistic in the formula (16). The results showed that the value of the test statistic is equal to (0.00011) and the value (P-value = 0) is less than the level of significance (5%). This leads to the rejection of the null hypothesis, which states that the parameters are equal to the vector (β_{1i}) and as a result, the parameter vector (β_{1i}) is random, and assuming that the vector (β_2) is fixed, so the model is random-mixed.

❖ **First stage:**

At this stage, the parameters estimators of the mixed random parameter dual data model were extracted in two Methods (FGLS) and (MG) and compared between them through the (MAPE) measure.

As for the best method in the mixed stochastic parameters estimation model and through the (MAPE) scale, we note that the (MG) is the best method in estimating the parameters of the model than the (FGLS) method, and the results are shown in Table (1).

Table (1)
Shows the preference for estimation methods in the case of the mixed Stochastic parameter model using (MAPE) scale

method measure	FGLS MSPR_SC	MG
MAPE	0.374039	0.360112

As for the values of the feature estimations according to the mixed stochastic parameters estimation model and according to the estimation methods, the estimator (MSPR_SC) using the (FGLS) method was calculated according to formula No. (5), and the estimator using the (MG) were calculated according to formula No. (35), and the results are shown in Table (2).

Table (2)
It shows the parameter values in the case of the mixed stochastic Parameter model according to the estimation methods

Methods Coefficient	FGLS (MSPR_SC)	MG
$_0\beta$	-696.035	-294.797
$_1\beta$	7.802907	9.433515
$_2\beta$	0.431089	0.280955

For the (FGLS) method, the model is as follows:

$$y_i = -696.035 + 7.802907X_1 + 0.431089X_2$$

As (b1 = 7.802907) represents the marginal propensity for per capita consumption of electrical energy, which increases by (7.803) if the consumer price index increases by one unit with the rest of the factors remaining constant, i.e. (X_1 directly affects the per capita consumption share of electrical energy).

As for (b2 = 0.431089), it represents the marginal propensity for per capita consumption of electrical energy, according to which it increases by (0.431) with an increase in the per capita share of GDP by one unit with the rest of the factors constant, meaning that (X_2 directly affects the per capita share of electrical energy consumption).

For the (MG) method, the model is as follows:

$$y_i = -294.797 + 9.433515X_1 + 0.280955X_2$$

As (b1 = 9.433515) represents the marginal propensity for per capita consumption of electric energy, which increases by (9.434) when increasing the consumer price index by one unit, with the rest of the factors remaining constant, meaning that (X_1 directly affects the per capita consumption share of electrical energy).

As for (b2 = 0.280955), it represents the marginal propensity for the per capita consumption of electrical energy, according to which it increases by (0.280955) when the per capita share of the GDP increases by one unit, with the rest of the factors remaining constant, meaning that (X2 directly affects the per capita share of electrical energy consumption).

As for the fixed limit (β_o) for the per capita consumption function of electrical energy, which means that the per capita consumption function of electrical energy is constant and equal to (-696.035) by the (FGLS) method and equal to (-294.797) by the (MG) method when the model is not significant.

- **The second stage**

At this stage, the parameters of the mixed parameters dual data model were estimated for each cross section (country) separately in terms of studying the per capita consumption of electric energy for Iraq and some of its neighboring countries. The results for each country will be mentioned below.

For All Countries and through the (MAPE) scale, and in the case of the panel data model for mixed- stochastic parameters, we note that the method of estimating the Mean Group (MG) is the best compared to the estimator (MSPR_SC), using the (FGLS) method, as shown in Table (3).

Table (3)

Shows the preference of estimation methods for all Countries using the (MAPE) scale

Method Measure	Iraq		Egypt		Jordon		Morocco		Tunisia	
	FGLS (MSPR_SC)	MG	FGLS (MSPR_SC)	MG	FGLS (MSPR_SC)	MG	FGLS (MSPR_SC)	MG	FGLS (MSPR_SC)	MG
MAPE	1.9097	0.0857	0.4797	0.0486	2.5916	0.0191	1.6976	0.0629	0.4495	0.0397

As for the values of the parameters estimates for Iraq and the rest of the countries, according to the mixed stochastic panel data model using the (FGLS) method, they were calculated according to the formula (28), while the estimator of the group mean (MG) estimation method was calculated according to the formula (35) as shown in Table No. (4), which represents estimates of features for all countries.

Table (4)
Shows parameter values for all countries and according to estimation methods

Method Coefficient	Iraq		Egypt		Jordon		Morocco		Tunisia	
	FGLS MSPR_SC	MG	FGLS MSPR_SC	MG	FGLS MSPR_SC	MG	FGLS MSPR_SC	MG	FGLS MSPR_SC	MG
$_0\beta$	-2068.11	-381.82	-597.87	377.79	2711.29	122.41	-329.56	-1026.36	-1108.49	-564.42
$_1\beta$	33.5701	6.5718	-0.65433	6.57180	1.46878	6.5718	-1.90018	6.5718	0.37461	6.5718
$_2\beta$	0.1328	0.35212	0.87035	0.35212	-0.27014	0.35212	0.39951	0.3521	0.62805	0.3521

4. Conclusion

In this paper, after examining the estimators (MSPR-SC) of the two estimation methods (MG) and (FGLS) of the panel data models for the mixed random parameters when the errors are serially correlated of the first order, and after applying the real data, the results indicate that the estimator of the estimation method (MG) has the smallest values for the efficiency measure (MAPE) from (FGLS) method, we conclude that the (MG) method is the best and best method for estimating model parameters

Also, the results of applying the real data for each cross-section (country) indicated that the estimations of the (MG) method are more efficient than the estimations of (FGLS) and for all countries.

We also note that the per capita consumption is directly affected, starting from the consumer price index and the per capita share of the gross domestic product, as the increase of these variables leads to an increase in the per capita consumption of electric energy, and this is consistent with the economic theory.

The most important recommendation for future study is to study the subject of unbalanced panel data, which includes the quality of missing values in the panel data, as well as the use of the MAPE scale in determining the efficient and best methods in the process of estimating parameters because of its flexibility and helps the researcher in getting more accurate and objective results.

References

1.Abonazel, M.R. 2018. Efficiency Comparisons of Different Estimators for Panel Data Models with Serially Correlated Errors: A Stochastic Parameter Regression Approach. International Journal of Systems Science and Applied Mathematics, 3, 2, pp. 37.

2.Abonazel, M.R. 2019. Generalized estimators of stationary random-coefficients panel data models: Asymptotic and small sample properties. REVSTAT-Statistical Journal, 17, 4, pp. 493–521.

3.Al-Mafarji, R.K.H. 2018. Estimation of random parameters in regression models for panel data with practical application. Baghdad.

4.Basim Shaliba Muslim 2009. Bayesian analysis of regression models for panel data. Baghdad.

5.Griffiths, W.E., Judge, G.G., Hill, R.C., Lütkepohl, H. & Lee, T.-C. 1985. The theory and practice of econometrics. Wiley.

6.Greene, W. H. 1997. Econometric Analysis. 3d ed. Upper Saddle River, NJ: Prentice– Hall.

7.Hausman, J.A. 1978. Specification tests in econometrics. Econometrica: Journal of the econometric society, pp. 1251–1271.

8.Kazem, A.H. & Muslim, B.H. 2002. Advanced Economic Measurement Theory and Practice. Baghdad: Dunia al-Amal Library.

9.Pio, B.P. 2003. From the help desk: Swamy's random-coefficients model. The Stata Journal, 3, 3, pp. 302–308.

10. Rao, C.R., Rao, C.R., Statistiker, M., Rao, C.R. & Rao, C.R. 1973. Linear statistical inference and its applications. Wiley New York.

11. Reda Abonazel, M. 2018. Different estimators for stochastic parameter panel data models with serially correlated errors. Journal of Statistics Applications & Probability, 7, 3, pp. 423–434.

12. Reem Talal Kamel Al-Adly 2021. Using some methods of estimating longitudinal data models with a practical application. Baghdad.

13. Rosenberg, B. 1973. A survey of stochastic parameter regression. NBER.

14. Swamy, P.A.V.B. 1970. Efficient inference in a random coefficient regression model. Econometrica: Journal of the Econometric Society, pp. 311–323.

15. Swamy, P.A.V.B. 2012. Statistical inference in random coefficient regression models. Springer Science & Business Media.

استخدام بعض طرائق التقدير لنماذج انحدار البيانات المزدوجة العشوائية المختلطة ذات الأخطاء المتسلسلة مع التطبيق

الباحث/ مصعب عامر رجب
قسم المعلوماتية والنظم، الدائرة الادارية، وزارة الكهرباء،
بغداد، العراق
Mosab.Aamer1201a@coadec.uobaghdad.edu.iq

أ.د. محمد صادق عبد الرزاق
قسم الأحصاء، كلية الأدارة والأقتصاد،
جامعة بغداد، بغداد، العراق
dr_aldouri@coadec.uobaghdad.edu.iq

Received:12/12/2022 Accepted: 29/12/2022 Published: December / 2022

هذا العمل مرخص تحت اتفاقية المشاع الابداعي نَسب المُصنَّف - غير تجاري - الترخيص العمومي الدولي 4.0
Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)

مستخلص البحث

يتضمن هذا البحث دراسة نماذج البيانات المزدوجة ذات المعلمات العشوائية المختلطة، وهي التي تحتوي على نوعين من المعلمات الاولى عشوائية والاخرى ثابتة، بالنسبة للمعلمة العشوائية تحصل نتيجة الاختلافات في الميول الحدية للمقاطع العرضية، وبالنسبة للمعلمة الثابتة تحصل نتيجة الاختلافات في الحدود الثابتة، والاخطاء العشوائية لكل مقطع عرضي تحمل صفة عدم تجانس التباين بالاضافة الى وجود ارتباط تسلسلي من الدرجة الاولى، والهدف الرئيسي في هذا البحث هو استعمال طرائق كفوءة تتناسب مع البيانات المزدوجة في حالة العينات الصغيرة، ولتحقيق هذا الهدف تم أستخدام طريقة المربعات الصغرى العامة المجدية (FGLS)، وطريقة متوسط المجموعة (MG) ومن ثم مقارنة كفاءة المقدرات المستخرجة في حالة المعلمات العشوائية المختلطة وتحديد الطريقة التي تعطينا المقدر الكفؤ. وتم تطبيق بيانات حقيقية تضمنت حصة استهلاك الفرد من الطاقة الكهربائية (Y) لخمسة دول التي تمثل عدد المقاطع العرضية (N=5) خلال فترة زمنية مدتها تسعة سنوات (T=9) فيكون عدد المشاهدات هي (n=45) مشاهدة والمتغيرات التوضيحية هي الرقم القياسي لاسعار المستهلك (X1) وحصة الفرد من الناتج المحلي الاجمالي (X2) ولتقييم أداء مقدرات كل من طريقة (FGLS) وطريقة (MG)على الأنموذج العام تم أستعمال مقياس (MAPE) لمقارنة كفاءة المقدرات. وأظهرت النتائج أن طريقة تقدير متوسط المجموعة (MG) هي الطريقة الأفضل في تقدير المعلمات من طريقة المربعات الصغرى العامة المُجدية (FGLS). وكذلك ظهرت طريقة تقدير (MG) هي الطريقة المثلى والافضل لمقدرات المعلمات الفرعية لكل مقطع عرضي (دولة).

نوع البحث: ورقة بحثية
المصطلحات الرئيسة للبحث: مقدر المربعات الصغرى العامة المَجدي، الارتباط التسلسلي من الدرجة الأولى ، أنموذج انحدار المعلمة العشوائية المختلطة، تقدير متوسط المجموعة.

*البحث مستل من رسالة ماجستير