International Journal
of Discoveries and Innovations in Applied Sciences
Vol. 4 Issue 3 | pp. 44-54 | e-ISSN: 2792-3983
Available online @ https://openaccessjournals.eu/index.php/ijdias

OPEN ACCESS JOURNALS

Article

Comparing Estimation Methods for Multilevel Regression Parameters

Kareem Khalaf Aazer

Department of Statistics, University of Sumer, Al-Rifai, Iraq
Correspondence: Kareemalataby28@gmail.com

Abstract. Multilevel regression models are widely utilized in various fields, such as health, education, and agriculture, due to their ability to address dependencies between variables at different levels of aggregation. This study compares two estimation methods for parameters in multilevel regression models: the maximum likelihood method (MLE) and the variance constraints method (RCR). Utilizing simulation experiments repeated 1000 times across different sample sizes (15, 40, and 60), the study evaluates these methods using the Akaike Information Criterion (AIC) and mean square error (MSE). Results indicate that the MLE consistently outperforms the RCR in terms of lower AIC and MSE values, especially as sample size increases. Despite the closeness in efficiency between the methods at smaller sample sizes, the MLE's superiority becomes more pronounced with larger samples, suggesting its robustness and reliability for parameter estimation in multilevel regression models. This finding is crucial for researchers seeking accurate and efficient estimation techniques in hierarchical data analysis.

Keywords: Multilevel Regression, Maximum Likelihood, Variance Constraints, Parameter Estimation, Simulation Study

1. Introduction

There is an increasing interest in educational and social research regarding the challenges of defining the relationships between variables that are at different levels of aggregation.In the field of school effectiveness research, one can be interested in studying the impact of the school budget on the academic performance of children.However, the former variable is specified at the school level, whereas the later variable is defined at the student level.This leads to challenges in accurately understanding the interrelationships between these factors.These challenges can be addressed by utilizing multilevel models, as suggested by Bryk and Raudenbush (1992), de Leeuw and Kreft (1986), Goldstein (1995), Longford (1993), and Raudenbush (1988).In the above example, students are organized hierarchically within schools. In a multilevel model, the students would constitute the first level, while the schools would represent the secondary level.The multilevel paradigm is primarily used in regression and analysis of variance models, as demonstrated by Bryk and Raudenbush. However, it can be applied to any statistical modeling of data where there are nested elementary units inside aggregates.Longford provides illustrations of multilevel factor analytical models and generalized linear models [1].

Citation: Kareem Khalaf Aazer, Comparing Estimation Methods for Multilevel Regression Parameters, International Journal of Discoveries and Innovations in Applied Sciences 2024, 4(3), 44–54

Received: 10th Jan 2024
Revised: 10th April 2024
Accepted: 24th April 2024
Published: 14th May 2024

Copyright: © 2024 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

Furthermore, the multilevel paradigm has been applied in several instances within the field of IRT models.Adams, Wilson, and Wu (1997) examine the inclusion of latent proficiency characteristics as dependent variables in a regression analysis.The study demonstrates that a regression model using latent proficiency variables can be understood as a two-level model. The first level is comprised of the item response measurement model, which represents the within-student model [2]. The second level is the model on the student population distribution, which represents the between-students model.Additionally, Adams et al. demonstrate that this method leads to a suitable handling of measurement error in the dependant variable of the regression model.Mislevy and Bock (1989) demonstrated another use of multilevel modeling in the context of IRT models. They developed a hierarchical IRT model that combines group-level and student-level effects.Both applications can be considered as specific instances of the overall method described in this context.This strategy involves using a multilevel regression model to analyze latent proficiency factors. The model allows for predictors at both the student-level and group-level [3][4].

2. The Multilevel Regression Model

The multilevel regression model is known in the statistical literature under a variety of names: hierarchical linear model, random coefficient model, variance component model, and mixed (linear) model [5][6]. Most often it assumes hierarchical data, with one response variable measured at the lowest level and explanatory variables at all existing levels. Conceptually, the model is often viewed as a hierarchical system of regression equations. For example, assume we have data in J groups or contexts and a different number of individuals N_f in each group. On the individual (lowest) level we have the dependent variable Y_{if} and the explanatory variable X_{if}, and on the group level we have the explanatory variable Z_f. Thus, we have a separate regression equation in each group:

$$Y_{if} = \beta_{0f} + \beta_{1j} X_{if} + e_{if} \qquad (1)$$

In Eq. (1) β_0 is the usual regression intercept, β_1, is the regression slope for the explanatory variable X, and e_{if} is the residual term. The regression coefficients β carry a subscript j for the groups, which indicates that the regression coefficients may vary across groups. The variation in the regression coefficients β_j is modeled by explanatory variables and random residual terms at the group level:

$$\beta_{if} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \qquad (2)$$

$$\beta_{1f} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \qquad (3)$$

Substitution of Eqs. (2) and (3) into Eq. (1) produces the single-equation version of the multilevel regression model:

$$Y_{if} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} Z_j X_{if} + u_{1f} X_{if} + u_{0j} + e_{if} \qquad (4)$$

In general, there will be more than one explanatory variable at the lowest level and also more than one explanatory variable at the highest level. Assume that we have P explanatory variables X at the lowest level, indicated by the subscript p ($p = 1, \dots, P$), and Q explanatory variables Z at the highest level, indicated by the subscript q ($q = 1, \dots, Q$), Then, Eq. (4) becomes the more general equation:

$$Y_{if} = \gamma_{00} + \sum_P \gamma_{p0} X_{pif} + \sum_q \gamma_{0q} Z_{qj} + \sum_q \sum_p \gamma_{pq} Z_{pj} X_{pij} + \sum_p u_{pj} X_{pif} + u_{0j} + e_{if} \qquad (5)$$

In Eq. (5), they are the usual regression coefficients, the $\boldsymbol{u}$ terms are residuals at the group level, and the e term represents the residual at the individual level. The regression coefficients are identified as the fixed part of the model because this part does not change over groups or individuals. The residual error terms are identified as the random or stochastic part of the model.[2][7]

The assumptions of the most commonly used multi-level regression model are that the residuals at the lowest level e_{if} have a normal distribution with a mean of zero and a common variance σ^2 in all groups. The second-level residuals u_{0j} and u_{pj} are assumed to be independent of the lowest level errors e_{if} and to have a multivariate normal distribution with means of zero. Other assumptions, identical to the common assumptions of a multiple regression analysis, are fixed predictors and linear relationships. Most multilevel software assumes by default that the variance of the residual errors e_{if} is the same in all groups. However, certain forms of heteroscedasticity can be explicitly modeled.

3. Estimation methods for a multilevel model

Marginal likelihood and model evidence for the linear model [8]

The marginal likelihood for linear models can often be found in Bayesian textbooks, such as [8], though for clarity, we include the this using the notation above. The marginal likelihood is:

$$p(M,\sigma^2) = \int_{R^4} \quad p(D|M,\beta,\sigma^2)p(M)d\beta$$

$$= \int_{R^4} \quad \frac{1}{(x\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}} exp\; exp\; (-\frac{1}{2}\;(\beta-\mu)^T\Sigma^{-1}(\beta$$

$$-\mu)) \prod_i \quad \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{1}{2\sigma^2}(y_i-\beta^T x_i)^2)d\beta$$

$$= \frac{1}{(2\pi)^{(d+n)/2}|\Sigma|^{\frac{1}{2}}\,\sigma^2} \int_{R^4} \quad exp\; exp\; (-\frac{1}{2}\;(\beta-\mu)^T\Sigma^{-1}(\beta-\mu)$$

$$+\frac{1}{\sigma^2}\sum_i \quad (y_i-\beta^T x_i)^2))d\beta$$

Finally, we get the marginal likelihood for the simple multilevel linear model: [2]

$$p\left(M,\sigma_y^2,\sigma_\eta^2\right) = \frac{|\Sigma|^{1/2}}{(2\pi\sigma_m^2)^{n/2}|\Sigma|^{1/2}} \prod_j \quad \left(\sqrt{\frac{\sigma_y^2}{\sigma_y^2+\eta_j\sigma_\eta}}\right) \times$$

$$exp\; exp\; \left(-\frac{1}{2}\left(\mu^T\Sigma^{-1}\mu + \frac{1}{\sigma_y^2}\sum_{i,j} \quad y_{ij}^2\right.\right.$$

$$\left.\left.-\frac{1}{\sigma_y^2}\left(\sum_j \quad \frac{\sigma_\eta^2}{\sigma_y^2+\eta_j\sigma_\eta^2}\left(\sum_i \quad y_{ij}\right)^2\right) - \mu^{-T}\Sigma^{-1}\acute{\mu}\right)\right)$$

As before, a version of this marginal likelihood derivation can also be found in [8], but, in this case, it is given in simplified matrix algebra form, where the dependence on σ_y^2 and σ_η^2, is left unspecified.

Marginal likelihood for a general multilevel linear model [1][4][10]

In the more general case (Eq. 4), the steps are almost identical:

$$p\left(M,\sigma_y^2,v\right)=\int_{R^4}\int_{R^{mx3}}\frac{1}{(2\pi)^{d/2}|\Sigma|^{\frac{1}{2}}}exp\,exp\left(-\frac{1}{2}(\beta-\mu)^T\Sigma^{-1}(\beta-\mu)\right)$$

$$\times\prod_i\frac{1}{(2\pi)^{rn/2}|\Sigma|^{1/2}}exp\,exp\left(-\frac{1}{2}\eta_j^T\sum_{\eta}^{-1}(v)\eta_j\right)$$

$$\times\prod_{i.j}\frac{1}{\sqrt{2\pi\sigma_y^2}}exp\,exp\left(-\frac{(y_{ij}-\beta^Tx_{ij}-\eta_j^Txx_{ij})^2}{2\sigma_y^2}\right)d\eta d\beta$$

$$=\frac{1}{\sigma_y^2|\Sigma_\eta(v)|^{\frac{m}{2}}(2\pi)\quad(n+d+mJ)/2}\times\int_{R^{d+mJ}}exp\,exp\left(-\frac{1}{2}(\beta-\mu)\right)\times$$

$$exp\,exp\left(-\frac{1}{2}\sum_j\eta_j^T\Sigma_\eta^{-1}(v)\eta_j-\frac{1}{2\sigma_y^2}\Sigma_i\left(y_{ij}-\beta^Tx_{ij}-\eta_j^Tz_{ij}\right)^2\right)d\eta d\beta$$

Then:

$$(\beta-\mu)^T\Sigma^{-1}(\beta-\mu)+\sum_i\eta_j^T\sum_{\eta}^{-1}\eta_j+\frac{1}{\sigma_y^2}\sum_{i,j}(y_{ij}-\beta^Tx_{ij}-\eta_j^Tz_{ij})^2$$

$$=(\beta-\mu)^T\Sigma^{-1}(\beta-\mu)\sum_i\left(\eta_j^T\left(\Sigma_\eta^{-1}+\frac{1}{\sigma_y^2}\sum_i z_{ij}z_{ij}^T\right)\eta_j\right.$$

$$-\eta_j^T\left(\frac{1}{\sigma_y^2}\sum_i z_{ij}\left(y_{ij}-\beta^Tz_{ij}\right)\right)$$

$$\left.-\left(\frac{1}{\sigma_y^2}\sum_i(y_{ij}-\beta^Tx_{ij})z_{ij}\right)\eta_j+\frac{1}{\sigma_y^2}\sum_i(y_{ij}-\beta^Tx_{ij})^2\right)$$

$$=(\beta-\mu)^T\Sigma^{-1}(\beta-\mu)$$

$$+\sum_j\left((\eta_j-\mu_{\eta,j})^T\sum_{n,j}^{-1}(\eta_j-\mu_{\eta,j})+\frac{1}{\sigma_y^2}\sum_i(y_{ij}\right.$$

$$\left.-\beta^Tz_{ij})^2-\mu_{\eta j}^T\sum_{\eta j}^{-1}\mu_{n,j}\right)$$

$$=(\beta-\mu)^T\Sigma^{-1}(\beta-\mu)+\sum_j\left((\eta_j-\mu_{n,j})^T\sum_{h,j}^{-1}(\eta_j-\mu_{\eta,j})\right)$$

$$+\frac{1}{\sigma_y^2}\sum_{i,j}y_{ij}^2$$

$$-\frac{1}{\sigma_y^4}\sum_j\left((\sum_i z_{ij}^T\,y_{ij})\sum_{\eta,j}(\sum_k z_{kj}\,y_{kj})\right)+\mu^T\Sigma^{-1}\mu$$

$$-\mu^T\Sigma^{-1}\mu$$

where we now have additional definitions:

$$\sum_{\eta,j}(\sigma_y^2, v) = \left(\Sigma_\eta^{-1} + \frac{1}{\sigma_y^2}\sum_i z_{ij}\, z_{ij}^T\right)^{-1}, \Sigma_\eta^{-1} = \Sigma_\eta^{-1}(v)$$

$$\mu_{\eta,j}(\sigma_y^2, v) = \Sigma_{\eta,j}\left(\frac{1}{\sigma_y^2}\sum_i z_{ij}\,(y_{ij} - \beta^T x_{ij})\right)$$

$$\hat{\Sigma}(\sigma_y^2, v) = \left(\Sigma^{-1} + \frac{1}{\sigma_y^2}\sum_{i,j} x_{ij}\, x_{ij}^T - \frac{1}{\sigma_y^4}\sum_j \left((\sum_i x_{ij}\, x_{ij}^T)\Sigma_{\eta,j}\,(\sum_k z_{kj}\, x_{kj}^T)\right)\right)^{-1}$$

$$\hat{\mu}(\sigma_y^2, v) = \hat{\Sigma}\left(\Sigma^{-1}\mu + \frac{1}{\sigma_y^2}\sum_{i,j} x_{ij}\, y_{ij} - \frac{1}{\sigma_y^4}\sum_j \left((\sum_i x_{ij}\, x_{ij}^T)\Sigma_{\eta,j}\,(\sum_k z_{kj}\, x_{kj}^T)\right)\right)$$

Finally, we get the marginal likelihood for the more general multilevel linear model:

$$p(M, \sigma_y^2, v) = \frac{|\hat{\Sigma}|^{1/2}}{(2\pi\sigma_y^2)^{n/2}|\Sigma|^{1/2}|\Sigma_\eta|^{J/2}}\prod_j |\Sigma_{\eta,j}|^{1/2} \times exp\ exp\left(-\frac{1}{2}\left(\mu^T\Sigma^{-1}\mu + \frac{1}{\sigma_y^2}\sum_{i,j} y_{ij}^2 - \frac{1}{\sigma_y^4}\sum_j \left((\sum_i z_{ij}^T\, y_{ij})\Sigma_{\eta,j}(\sum_k z_{kj}\, y_{kj})\right) - \mu^T\Sigma^{-1}\,\hat{\mu}\right)\right)$$

(7)

Rearranging the term in square brackets and integrating out B, we get:

$$p(M, \sigma^2) = \frac{|\underline{\Sigma}|^{1/2}}{(2\pi\sigma^2)^{n/2}|\Sigma|^{1/2}}\, exp\ exp\left(-\frac{1}{2}\left(\mu^T\Sigma^{-1}\mu + \frac{1}{\sigma^2}\sum_i y_i^2 - \mu^{-T}\Sigma^{-1}\underline{\mu}\right)\right)$$

(6)

where we define:

$$\Sigma(\sigma^2) = \left(\Sigma^{-1} + \frac{1}{\sigma^2}\sum_i x_i\, x_i^T\right)^2, \ \ \underline{\mu}(\sigma^2) = \left(\Sigma^{-1}\mu + \frac{1}{\sigma^2}\sum_i x_i\, y_i\right)$$

Marginal likelihood for a simple multilevel linear model[3][5]

For the simple multilevel linear model (Eq 3):

$$p\left(M,\sigma_y^2,\sigma_\eta^2\right)=\int_{R^4}\int_{R^j}\frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}exp\,exp\left(-\frac{1}{2}(\beta-\mu)^T\Sigma^{-1}(\beta-\mu)\right)$$

$$\times\prod_i\frac{1}{\sqrt{2\pi\sigma_\eta^2}}exp\,exp\left(-\frac{\eta_j^2}{2\sigma_\eta^2}\right)\times\times\prod_{i.j}\frac{1}{\sqrt{2\pi\sigma_y^2}}$$

$$exp\,exp\left(-\frac{(y_{ij}-\beta^Tx_{ij}-\eta_j)^2}{2\sigma_y^2}\right)d\eta d\beta$$

$$=\frac{1}{\sigma_y^n|\Sigma|^{\frac{1}{2}}(2\mu)^{(n+d)/2}}\int_{R^d}exp\,exp\left(-\frac{1}{2}(\beta-\mu)^T\,\Sigma^{-1}(\beta-\mu)\right)$$

$$\times\prod_j\left[\frac{1}{\sqrt{2\pi\sigma_\eta^2}}d\eta_j\right]d\beta$$

We first consider the integral in square brackets, completing the square in η_j in the expression:

$$\frac{\eta_j^2}{\sigma_\eta^2}+\frac{1}{\sigma_y^2}\sum_i\left(y_{ij}-\beta^Tx_{ij}-\eta_j\right)^2$$

$$=\frac{\sigma_y^2+\eta_j\,\sigma_\eta^2}{\sigma_y^2\,\sigma_\eta^2}\left(\eta_j-\frac{\sigma_\eta^2}{\sigma_y^2+\eta_j\sigma_\eta^2}\sum_i(y_{ij}-\beta^Tx_{ij})\right)^2+\frac{1}{\sigma_y^2}\sum_i(y_{ij}-\beta^Tx_{ij})^2-\frac{1}{\sigma_y^2}\frac{1}{\sigma_y^2+\eta_i\sigma_\eta^2}\left(\sum_i(y_{ij}-\beta^Tx_{ij})\right)^2$$

This gives:

$$\frac{1}{\sqrt{2\pi\sigma_\eta^2}}\int_{-\infty}^{\infty}exp\,exp\left(-\frac{\sigma_j^2}{2\sigma_\eta^2}-\frac{1}{2\sigma_y^2}\sum_i(y_{ij}-\beta^Tx_{ij}-\eta_j)^2\right)d\eta j$$

$$=\sqrt{\frac{\sigma_y^2}{\sigma_y^2+\eta_j\sigma_\eta^2}}\left(exp\,exp\left(-\frac{1}{2\sigma_y^2}\sum_i(y_{ij}-\beta^Tx_{ij})^2-\frac{\sigma_\eta^2}{\sigma_y^2+\eta_j\sigma_\eta^2}\left(\sum_i(y_{ij}-\beta^Tx_{ij})\right)^2\right)\right)$$

Then, rearranging for with β as in the linear model case:

$$(\beta-\mu)^T\Sigma^{-1}(\beta-\mu)+\frac{1}{\sigma_y^2}\sum_{i,j}(y_{i,j}-\beta^Tx_{i,j})\left(y_{i,j}-x_{ij}^T\beta\right)$$

$$-\frac{1}{\sigma_y^2}\sum_j\left(\frac{\sigma_\eta^2}{\sigma_y^4+\eta_j\sigma_\eta^2}\left(\sum_i(y_{ij}-\beta^Tx_{i,j})\right)^2\right)$$

$$(\beta-\hat{\mu})^T\Sigma^{-1}(\beta-\hat{\mu})+\mu^T\Sigma^{-1}\mu+\frac{1}{\sigma_y^2}\sum_{i,j} y_{ij}^2$$
$$-\frac{1}{\sigma_y^2}\sum_j\left(\frac{\sigma_\eta^2}{\sigma_y^2+\eta_j\sigma_\eta^2}(\sum_i (y_{ij})^2\right)-\hat{\mu}^T\hat{\Sigma}^{-1}\underline{\mu}$$

where we define:

$$\hat{\Sigma}(\sigma_y^2,\sigma_\eta^2)=\left(\Sigma^{-1}+\frac{1}{\sigma_y^2}\sum_{i,j} x_{ij}\,x_{ij}^T\right.$$
$$\left.-\frac{1}{\sigma_y^2}\sum_j\left(\frac{\sigma_\eta^2}{\sigma_y^2+\eta_j\sigma_\eta^2}(\sum_i x_{ij})(\sum_k x_{kj}^T)\right)\right)^{-1}$$

$$\hat{\mu}(\sigma_y^2,\sigma_\eta^2)=\hat{\Sigma}\left(\Sigma^{-1}\mu+\frac{1}{\sigma_y^2}\sum_{i,j} x_{ij}\,xy_{ij}\right.$$
$$\left.-\frac{1}{\sigma_y^2}\sum_j\left(\frac{\sigma_\eta^2}{\sigma_y^2+\eta_j\sigma_\eta^2}(\sum_i y_{ij})(\sum_k x_{kj})\right)\right)$$

Employing divergent constraints to estimate the random parameters of the random parameters model

This method relies on employing prior information derived from outside the sample, which is in a mixed electronic form, and combining this data with the limited sample data for these restrictions [11][8][7], as follows:

$$a_i<\beta<C_i \qquad i=1\dots\dots g$$

a_i:Minimum enrollment
C_i:Maxmum enrollment

$$r\geq R\beta \qquad (7)$$

r:(gx1)

R :[g * (k+1)]
B : ((k+1)*1)

$$r\ =\ R\,\beta\ +\ V \qquad (8)$$

$$E\ (V)\ =\ 0, E'(VV')\ =\ G$$
$$G\ =\ \left[s_{11}\ \dots\ 0\,.\,s_{22}\,.\,0\ \dots\ s_{gg}\ \right]$$

$$[y\,r\,]=[x\,R\,]\beta+[u\,v\,] \qquad (9)$$

Since:
y∗: represents a vector of order ((nt+g) * 1) of observations of the dependent variable, which includes the school sample data plus the retained information.
χ∗: a matrix of rank [(nt+g) * (k+1)] of observations of the explanatory variables, which includes the data of the studied sample plus prior information.
B: represents a vector of order ((k+1) * 1) of features to be estimated.
u∗: represents a vector of order [(nt+g) * 1)] of random errors, which includes the random errors of the studied sample data and prior information.
The random errors in Model (68-2) are subject to the following assumptions:

$$E\ (v^*)\ =\ E\,[u\,v\,]\ =\ [0\,0\,]\ =\ 0$$

$$E\,[uu'] = E\,[u\;v]\,[u\;v] = E\,[uu'\;uv'\;vu'\;vv'] = [\Omega_T\;0\;0\;G\,]$$

$: \Omega$

$$\Omega_{TRC} = \left[\chi_1\psi_1\chi_1 + \delta_{i1}^2 I_T \cdots 0 \;\vdots \ddots \vdots\; 0 \cdots \acute{\chi}_n\psi_n\chi_n + \delta_{i1}^2 I_T\right]$$

Because Model No. (9) suffers from a heterogeneity problem, the general least squares (GIS) method can be applied using the following formula:

$$\hat{\beta}_{QR} = (\chi^{*'}\,Q^{-1}\chi^*)^{-1}\,\chi^{*'}Q^{-1}y^* \qquad (10)$$

$$\left[\left[\chi'\;\;R'\right]\left(\Omega^{-1}\;0\;0\;G^{-1}\right)[\Omega^{-1}\;R\,]^{-1}\right]\left[\left[\chi'\;\;R'\right]\left(\Omega^{-1}\;0\;0\;G^{-1}\right)[y\;r\,]\right]$$

$$\hat{\beta}_{QR} = [(\chi'\;\;\Omega^{-1}\,\chi)^{-1} + R'G^{-1}R]^{-1}[\chi'\;\;\Omega^{-1}\,Y)^{-1} + R'G^{-1}r] \quad (11)$$

The estimation formula in equation (10) is unbiased for the parameters (β).
The variance and covariance matrix of the estimators can be defined as follows

$$Var\,\hat{\beta}_{QR} = [\chi'\Omega^{-1}\chi + R'G^{-1}R] \qquad (12)$$

The formula (11) depends on the value of б_i^2, which is often unknown in applied reality, so it is estimated based on the sample data for each section and using the ordinary least squares method. Since:

$$E(S_{ei}^2) = б_i^2 \qquad i = 1 \ldots\ldots n$$

Therefore, the estimation formula used to estimate the random parameters (6) Swamy's model in the presence of varying constraints is as follows:

$$\hat{\beta}_{QR} = \left[\chi'\hat{\Omega}^{-1}\chi + R'^{G^{-1}}R\right]^{-1}\left[\chi'\hat{\Omega}^{-1}\chi + R'^{G^{-1}}r\right] \quad (13)$$

$$Var\,\hat{\beta}_{QR} = (\chi'\Omega^{-1}\chi + R'^{G^{-1}}R)^{-1} \qquad (14)$$

It should be noted that the combined homogeneity between prior information and sample data must be tested before relying on it according to the following hypothesis:
$H0$: The existence of homogeneity between prior information and sample data.
$H1$: Lack of homogeneity between prior information and sample data.
This hypothesis is chosen in the following form:

$$\chi^2 = (r - R\hat{\beta}_{QR})'[G + S_e^2\,R\,(\chi'\hat{\Omega}^{-1}\chi)^{-1}R']^{-1}(r - R\hat{\beta}_{QR}) \quad (15)$$

1-Employing heterogeneous constraints for the random parameters regression model
The process of employing varying constraints in estimating random parameters in the general model does not differ in its steps from what was presented in the section (employing constraints in the Swamy model), with attention to the change of the matrix Ω^*

$$\hat{\beta}_{GQR} = (\chi'\,Q^{*-1}\chi)^{-1}(\chi'Q^{*-1}\,Y) \qquad (16)$$

$$\hat{\beta}_{GQR} = [(\chi'\Omega_{GRCR}^{*-1}\,\chi)^{-1} + R'G^{-1}R]^{-1}[\chi'\Omega_{GRCR}^{*-1} + R'G^{-1}r] \qquad (17)$$

The variance and covariance matrix are as follows:

$$Var\,\hat{\beta}_{GQR} = (\chi'\,\Omega^{*-1}\chi + R'G^{-1}R)$$

$$\hat{\beta}_{GQR} = [(\chi'\hat{\Omega}_{GRCR}^{*-1}\,\chi)^{-1} + R'G^{-1}R]^{-1}[\chi'\hat{\Omega}_{GRCR}^{*-1} + R'G^{-1}r] \qquad (18)$$

4. Stages of building a simulation experiment

This part of the research includes conducting a simulation experiment using the simulation method in the process of comparing estimation methods and determining the optimal methods based on the two comparison scales, the Akaike criterion (AIC) and the mean square error (MSE). Using the

Matlab program, the program was created and a Monte Carlo simulation was carried out, where the simulation experiments relied on generating data. Multilevel regression model

$Y = X\beta + u$ (Level 1)

$\beta = g\gamma + e$ (Level-2) (19)

$t = 1,2,3, \dots (T = 15, T = 40, T = 60), i = 1, 2, 3, j = 0, 1, 2, 3, N = 3T$

The dependent and explanatory variables, random errors, and parameters are as follows:

$Y = [Y_1\ Y_2\ Y_3\ ]_{(3T\times1)}$, $X = [X_1\ 0\ 0\ 0\ X_2\ 0\ 0\ 0\ X_3\ ]_{(3T\times3P)}$, $\beta = \left[\beta_0\ \beta_1\ \beta_2\ \right]_{(3P\times1)}$,

$u = [u_1\ u_2\ u_3\ ]_{(3T\times1)}$, $g = [g_1\ 0\ 0\ 0\ g_1\ 0\ 0\ 0\ g_3\ ]_{(3\times3)}$, $\gamma = \left[\gamma_1\ \gamma_2\ \gamma_3\ \right]_{(6\times1)}$, $e = [e_1\ e_2\ e_3\ ]_{(3\times1)}$

To conduct the simulation under the assumptions of a multilevel regression model (with two levels) for the model where the data was generated using the Bootstrap Sampling method; The data will be generated according to the following steps:

Generating random errors error term according to a normal distribution with a mean and variance defined as follows according to the assumed sample sizes (15, 40, 60), as the generation was done:

$$u_1: normal\ (0.04\ ,0.5)$$
$$u_2: normal\ (0.3,0.5)$$
$$u_3: normal\ (0.6\ ,0.5)$$

Initial values were assumed for the parameters as shown in the following table 1.

Table 1. Initial Value

parameters	β_0	β_1	β_2
Initial value	1	1.5	3

Calculate the values of the variables Y (first iteration) for all three cross-sections based on the values of the random errors and the default values of the parameters described in steps 1 and 2, respectively, and the real data as explanatory variables.

Table 2. The Estimators of The Maximum Likelihood Method (MLE) and the Identical Constraints Method (RCR)

Methods		MLE	RCR
$n = 3T$	j		
15	0	103.044	103.067
	1	1.48885	1.48883
	2	2.99997	2.99992
40	0	56.88664	56.89465
	1	1.48554	1.48551
	2	2.99996	2.99991
60	0	115.88	1.453
	1	1.48712	1.4871
	2	2.99998	2.99993

It is noted from Table (2) that the estimators of the maximum likelihood method (MLE) and the identical constraints method (RCR) are very close in value for all Level-1 parameters of the multi-level model (with two levels) assumed in the simulation experiment. On the contrary, we find that the estimator of the maximum possibility method (MLE) has given estimators It differs from the

RCR method, but the difference is not large. This also applies to the estimators of the Level-2 parameters for the same model and for the same estimation methods, as shown in the following table:

Table 3. Level 2 parameter estimates from the simulation model when It is a pooled model

الطريقة / N=3T	MLE				RCR			
15	γ_{00}	61.312	γ_{01}	0.014	γ_{00}	56.955	γ_{01}	0.01471
	γ_{10}	1.463	γ_{11}	0.312	γ_{10}	1.46215	γ_{11}	0.3111
	γ_{20}	3.043	γ_{21}	0.048	γ_{20}	2.99829	γ_{21}	0.0056
40	γ_{00}	482.9	γ_{01}	0.005	γ_{00}	480.08	γ_{01}	0.0463
	γ_{10}	1.49	γ_{11}	0.002	γ_{10}	1.49	γ_{11}	0.0021
	γ_{20}	3	γ_{21}	0.007	γ_{20}	2.9983	γ_{21}	0.009
60	γ_{00}	427.0891	γ_{01}	0.008	γ_{00}	427.025	γ_{01}	0.008
	γ_{10}	1.47067	γ_{11}	0.008	γ_{10}	1.47027	γ_{11}	0.008
	γ_{20}	2.99994	γ_{21}	0.00023	γ_{20}	2.99913	γ_{21}	0.00023

Table 4. Values of the AIC and MSE parameters of the PMRM-2 simulation model when Pooled Model

method		MLE	RCR
N=3T	Standard		
15	AIC	349.1872	349.1869
	MSE	357.5944	357.5941
40	AIC	586.8427	586.8427
	MSE	598.4336	598.4336
60	AIC	2544.097	2544.094
	MSE	2564.593	2564.591

It is clear from Table (4) that the values of the AIC and MSE coefficients for the two maximum likelihood methods (MLE) and the identical constraints method (RCR) are very close at N = 15, and equal for the two sample sizes N = 40 and N = 60, which indicates that their efficiency in the estimation process is close or equal, and when compared Together, we find that the MLE method has smaller values for the comparison coefficients (AIC) and MSE than the identical constraints method, and for all sample sizes adopted in the research (15, 40, 60). It is also noted that the values of the coefficients AIC and MSE increase as the sample size increases.

5. Conclusion

Through the beginners Akaike as well as the mean square error it is recognized that the maximum likelihood method is the best in estimation. We note that the larger the eye size, the better the estimate, and also that the mean square error begins to decline. It is clear from Table (4) that the values of the AIC and mse coefficients for the maximum possibility method and the identical constraints method are very close at N = 15, and equal for the two sample sizes N = 40 and N = 60, which indicates the closeness or equality of their efficiency in the estimation process, according to the coefficients (Standard comparison of AIC and MSE in the process of estimating the parameters of a multilevel model (with two levels) for PMRM-2 panel data when the model is pooled. It is also noted that the values of the coefficients AIC and BIC increase as the sample size increases.

REFERENCES

1. W. J. Browne and D. Draper, "A Comparison of Bayesian and Likelihood-Based Methods for Fitting Multilevel Models," Journal of the Royal Statistical Society: Series A (Statistics in Society), vol. 169, no. 3, pp. 625-641, 2006.
2. L. Beckmann, M. C. Bergmann, F. Fischer, and T. Mößle, "Risk and Protective Factors of Child-to-Parent Violence: A Comparison Between Physical and Verbal Aggression," Journal of Interpersonal Violence, vol. 36, no. 3-4, pp. NP1309-NP1334, 2021.
3. M. S. Kovačević and S. N. Rai, "A Pseudo Maximum Likelihood Approach to Multilevel Modelling of Survey Data," Communications in Statistics-Theory and Methods, vol. 32, no. 1, pp. 103-121, 2003.
4. M. Moeyaert, D. Rindskopf, P. Onghena, and W. Van den Noortgate, "Multilevel Modeling of Single-Case Data: A Comparison of Maximum Likelihood and Bayesian Estimation," Psychological Methods, vol. 22, no. 4, pp. 760-778, 2017.
5. S. D. Pimentel, L. C. Page, M. Lenard, and L. Keele, "Optimal Multilevel Matching Using Network Flows: An Application to a Summer Reading Intervention," The Annals of Applied Statistics, vol. 12, no. 3, pp. 1479-1505, 2018.
6. J. Thurnherr, A. Berchtold, P. A. Michaud, C. Akre, and J. C. Suris, "Violent Adolescents and Their Educational Environment: A Multilevel Analysis," Journal of Developmental & Behavioral Pediatrics, vol. 29, no. 5, pp. 351-359, 2008.
7. S. Rabe-Hesketh and A. Skrondal, "Multilevel Modelling of Complex Survey Data," Journal of the Royal Statistical Society: Series A (Statistics in Society), vol. 169, no. 4, pp. 805-827, 2006.
8. J. Thurnherr, A. Berchtold, P. A. Michaud, C. Akre, and J. C. Suris, "Violent Adolescents and Their Educational Environment: A Multilevel Analysis," Journal of Developmental & Behavioral Pediatrics, vol. 29, no. 5, pp. 351-359, 2008.
9. M. L. Wong, R. K. W. Chan, D. Koh, H. H. Tan, F. S. Lim, S. Emmanuel, and G. Bishop, "Premarital Sexual Intercourse Among Adolescents in an Asian Country: Multilevel Ecological Factors," Pediatrics, vol. 124, no. 1, pp. e44-e52, 2009.
10. R. Triebel, P. Pfaff, and W. Burgard, "Multi-Level Surface Maps for Outdoor Terrain Mapping and Loop Closing," in 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2276-2282, Oct. 2006.
11. J. R. Zubizarreta and L. Keele, "Optimal Multilevel Matching in Clustered Observational Studies: A Case Study of the Effectiveness of Private Schools Under a Large-Scale Voucher System," Journal of the American Statistical Association, vol. 112, no. 518, pp. 547-560, 2017.