BAREKENG: Journal of Mathematics and Its Applications
September 2026 Volume 20 Issue 3 Page 1869-1892
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss3pp1869-1892

COMPARISON OF MIXED EFFECT REGRESSION TREE (MERT) AND LINEAR MIXED MODEL (LMM) FOR CLUSTERED DATA ON CASE STUDY HOUSEHOLD POVERTY IN WEST JAVA PROVINCE

Nur Fitriyani Sahamony[1], **Asysta Amalia Pasaribu**[2*], **Bagus Sartono**[3], **Khairil Anwar Notodiputro**[4]

[1]*Faculty of Business and Social Science, Universitas Binawan*
Jln. Dewi Sartika Raya, Jakarta Timur, 13630, Indonesia

[2]*Department of Statistics, School of Computer Science, Universitas Bina Nusantara*
Jln. Raya Kb. Jeruk, Kemanggisan, Jakarta Barat, 11530, Indonesia

[1,2,3,4]*Study Program of Statistics and Data Science, IPB University*
Jln. Raya Dramaga Kampus IPB, Babakan, Kec. Dramaga, Bogor, Jawa Barat, 16680, Indonesia

*Corresponding author's e-mail: * **asysta.amalia@binus.ac.id***

Article Info

Article History:
Received: 2nd July 2025
Revised: 26th January 2026
Accepted: 16th March 2026
Available online: 8th April 2026

***Keywords**:*
Clustered data;
Linear mixed model;
Machine learning;
Mixed effect regression tree;
No poverty.

ABSTRACT

This study compares the performance of the Linear Mixed Model (LMM) and the Mixed Effect Regression Tree (MERT) in analyzing the determinants of household consumption expenditure in West Java Province. The LMM integrates fixed and random effects to account for both individual and regional variations, while MERT extends this approach by incorporating a regression tree framework to capture nonlinear relationships and complex interactions among socio-economic variables. Using data from the 2023 National Socioeconomic Survey (SUSENAS), household consumption expenditure is modeled as an indicator of poverty. The results show that key determinants across both models include the gender and age of the household head, highest educational attainment, household size, land and car ownership, and welfare card ownership. Education and asset ownership consistently emerge as major factors influencing household welfare. The MERT model demonstrates superior predictive performance, with lower RMSE and MAE values compared to the LMM, while offering greater interpretability by identifying specific household profiles. Female-headed households with higher education and no car ownership tend to have higher expenditure in the high-income group, whereas female-headed households with welfare cards remain vulnerable in the low-expenditure group. From a policy perspective, these findings highlight the importance of improving educational access, enhancing asset ownership, and strengthening targeted social protection for vulnerable groups. Overall, while both models contribute valuable insights, the MERT model provides a more flexible and powerful framework for identifying and interpreting the determinants of household welfare in West Java.

How to cite this article:

N. F. Sahamony, A. A. Pasaribu, B. Sartono and K. A. Notodiputro., "COMPARISON OF MIXED EFFECT REGRESSION TREE (MERT) AND LINEAR MIXED MODEL (LMM) FOR CLUSTERED DATA ON CASE STUDY HOUSEHOLD POVERTY IN WEST JAVA PROVINCE", *BAREKENG: J. Math. & App.*, vol. 20, no. 3, pp. 1869-1892, Sep, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Clustered data are data sets in which the response variable is measured once for each subject (the unit of analysis), and the units of analysis are grouped into, or nested within, clusters of units. Clustered data can be arranged hierarchically into levels. All data sets suitable for LMM analysis involve at least two levels. At Level 1, observations represent the most granular data, describing individual units of analysis or subjects. At Level 2, the data reflect the clustering of these units (e.g., households within districts). Clustered data can be modeled using several statistical approaches, including the linear mixed model framework [1]-[3]. A linear mixed model (LMM) is a statistical approach suitable for clustered, longitudinal, or repeated-measures data that captures the relationship between a continuous response variable and a set of predictor variables (2). LMMs incorporate both fixed effects parameters associated with observed covariates that apply to the entire population and random effects that account for unobserved heterogeneity across clusters or subjects [1], [3]. The inclusion of random effects enables LMMs to model hierarchical structures and within-group correlations, making them highly flexible for analyzing multilevel data, particularly in socio-economic and health studies [4]. Linear Mixed Model (LMM) structure is particularly useful for poverty analysis because it explicitly accounts for both within-region and between-region variability in household welfare indicators such as per capita expenditure. In the context of West Java, poverty levels are not uniform across districts and cities area has unique socio-economic characteristics, infrastructure, and demographic compositions.

The term linear mixed models refer to models that are linear in parameters and allow covariates to include both fixed and random effects. Fixed effects are typically associated with continuous variables such as weight, baseline scores, or socioeconomic status or categorical variables like gender or treatment groups [1], [2]. These fixed effects are treated as unknown but constant parameters that apply uniformly across all units in the population. Their estimation is often of primary interest because they represent the systematic influence of predictors on the outcome variable [4], [5]. When categorical factor levels are regarded as randomly sampled from a broader population and not of primary analytical concern, they are suitably modeled as random effects in a Linear Mixed Model (LMM). In contrast, fixed effects correspond to covariates of substantive interest, represented by constant parameters. Random effects are latent variables, commonly assumed to follow a normal distribution, enabling LMMs to capture clustering, repeated measures, and hierarchical structures.

This modeling framework has proven effective across disciplines such as epidemiology [1], [3], public health [6], education [7], and cognitive aging [8], offering analytical flexibility when accounting for both population-level and subject-specific variability. Recent methodological developments have further extended LMMs into generalized forms (GLMMs) and joint modeling approaches to address complex interdependencies in large-scale and longitudinal health surveys [2], [3]. Linear Mixed Models (LMM) have been widely applied across numerous scientific domains. In economics and finance, LMMs have been employed to assess the impact of foreign direct investment on labor indicators in transition economies [9], and to explore poverty indicators using multilevel dynamic modeling frameworks [1], [10], [11]. Within health and biomedical sciences, LMMs are utilized to investigate risk factors for malaria and anemia [3], assess psychopharmacological treatment responses [12], and analyze longitudinal progression in Alzheimer's disease among individuals with Down syndrome [13]. In the education sector, they are applied to evaluate reading intervention effectiveness [7] and to model hierarchical performance in national examinations [14]. In the field of psychology, LMMs are adopted to analyze gait dynamics [15] and mental health outcomes among adolescents [16]. Moreover, in agriculture and animal science, LMMs support meta-analyses of agricultural data [17], assess resilience in traditional farming systems [18], and examine genetic traits in poultry breeding programs [19].

Complementing the LMM framework, the Mixed Effects Regression Tree (MERT) model has emerged as a flexible extension capable of handling clustered data structures. This model integrates decision-tree splitting rules to accommodate fixed effects, while employing a node-invariant linear structure to model random effects. It effectively accounts for intra-cluster correlation, such as that between administrative regions, by embedding standard regression tree procedures within the Expectation-Maximization (EM) algorithm framework [4], [6]. Consequently, MERT offers a promising methodological alternative for modeling complex multilevel socioeconomic variables. MERT is considered suitable and complementary to the LMM in poverty analysis because it combines the interpretability of tree-based models with the hierarchical structure of mixed-effects modeling. In the context of West Java, poverty determinants often exhibit nonlinear relationships and interaction effects among socio-economic variables such as education, household size, and asset ownership. Traditional LMMs assume linearity and may not capture these complex patterns effectively. In this study, the LMM model is applied in the socio-economic domain, including household poverty in West Java Province. According to the March 2023 survey by Badan Pusat Statistik (BPS), the poverty rate in West Java was 7.62%. Poverty in West Java is measured by households' ability to meet their basic needs both food and non-food based on per-capita expenditure. Although numerous studies have applied statistical and econometric models such as Ordinary Least Squares (OLS) or Linear Mixed Models (LMM) to analyze poverty and household consumption expenditure, most have focused on linear relationships and homogeneous population structures without considering the complex, nonlinear interactions among socio-economic variables. Moreover, spatial and regional heterogeneity such as differences in living standards and access to resources across districts as often been treated as noise rather than as a meaningful source of variation. Research that integrates machine-learning approaches, particularly Mixed Effect Regression Tree (MERT) models, remains scarce in poverty analysis. Existing studies rarely explore how combining hierarchical random effects with nonlinear tree-based structures can reveal hidden patterns of vulnerability and context-specific poverty determinants. Therefore, this study fills an important gap by applying both LMM and MERT to model per capita expenditure in West Java, offering a more flexible and data-driven framework to understand multidimensional poverty across regions. Poverty can be prevented by focusing on influential factors. The phenomenon of poverty can be addressed by understanding its root causes factors that go beyond individual behavior to include differences in regional conditions where households reside. Therefore, both LMM and MERT models can be used in practice to analyze case studies of per-capita expenditure.

This study aims to evaluate the predictive performance of the Mixed Effects Regression Tree (MERT) model in comparison with the Linear Mixed Model (LMM), using household consumption expenditure per capita data from regencies and cities in West Java Province, Indonesia, for the year 2023. West Java was selected as the study area due to its substantial regional heterogeneity. Some regions consist of large, densely populated cities located in close proximity to the DKI Jakarta metropolitan area, while others are geographically remote and predominantly rural. According to Statistics Indonesia (BPS, 2023), West Java has a population of approximately 49.6 million people, making it the most populous province in Indonesia. Population density varies widely across districts, from 15,051 people per km² in Bandung City and 12,907 people per km² in Bekasi City both highly urbanized areas near Jakarta to only 396 people per km² in Pangandaran Regency, which is largely rural and located in the southern coastal region. In addition, GRDP per capita also shows sharp disparities, ranging from over IDR 170 million in Bekasi City to below IDR 40 million in Garut Regency (BPS West Java, 2023). These contrasts in population density, economic structure, and geographic accessibility illustrate the strong inter-regional heterogeneity that makes West Java an ideal case for mixed-effects modeling of poverty and household consumption expenditure. This geographic and infrastructural diversity results in considerable variation in household characteristics, including socioeconomic status, educational attainment, household size, asset ownership (e.g., cars and land), and possession of government-issued welfare cards [1], [10], [11], [14]. Such variations are known to significantly influence poverty indicators and expenditure levels and therefore must be accounted for in model estimation and comparison [4], [6].

Comparing LMM and MERT is important because both models represent distinct yet complementary approaches to understanding poverty dynamics across heterogeneous regions. The Linear Mixed Model (LMM) offers a statistically rigorous framework for analyzing hierarchical data structures, effectively capturing variation between districts through random effects while assuming linear relationships among predictors. However, socio economic phenomena such as household consumption expenditure often exhibit nonlinear patterns and complex interactions that linear models cannot fully capture. The Mixed Effects Regression Tree (MERT), as an extension of the LMM framework, overcomes this limitation by integrating decision-tree partitioning to identify nonlinear relationships and threshold effects while still accounting for intra-regional correlations through random effects. Therefore, a comparative evaluation of LMM and MERT provides methodological and practical insights into which model more effectively explains regional disparities in poverty and reveals hidden vulnerability patterns. This comparison contributes to the literature by demonstrating how combining traditional statistical inference with machine-learning flexibility can enhance poverty modeling and support more targeted, data-driven policy interventions in West Java. This study will not only focus on all regencies and cities in West Java, but we also divide them into three regions. This categorization is based on the 2023 Regional Gross Domestic Product (PDRB) in West Java Province, according to the Badan Pusat Statistik (BPS). High-PDRB region as Quartile I, Medium-PDRB region as Quartile II, Low-PDRB region as Quartile III.

Based on the above explanation, the objectives of this study is to gain an overview of per-capita expenditures an indicator of poverty in West Java Province and its differences across regencies/cities, to model per-capita expenditure in West Java Province by incorporating both fixed effects and random effects, using Linear Mixed Models (LMM) and Mixed Effects Regression Trees (MERT), and to identify the significant influencing variables. Conducted LMM and MERT modeling of per-capita expenditure based on GRDP values, categorizing into high-, medium-, and low-expenditure groups, and to compare the affecting variables identified in these group-specific models with those found in the overall West Java model. compared the performance of MERT and LMM in modeling per-capita expenditure, both in the West Java-wide model and within the expenditure-based subgroups.

2. RESEARCH METHODS

2.1 Dataset

This study uses secondary data, namely the raw data from the 2023 National Socio Economic Survey (SUSENAS). SUSENAS is a routinely conducted survey by Statistics Indonesia (BPS) to collect data on the social and economic conditions of Indonesian households. It serves as a key source for national development planning and evaluation. Unit of analysis is Households at the regency/city level in West Java Province, response variable is average per-capita household consumption expenditure in each regency/city of West Java, Response variables are Gender of household head, Age of household head, Highest educational attainment of household head, Number of household members, Car ownership, Land ownership, Possession of welfare cards, Regency/city (as a region-level random variable). The analysis is based on $N = 25{,}891$ household observations, distributed across 17 regencies and 9 cities in West Java.

Response and predictor variables used in this study can be seen in Table 1. The coverage area in this study is the regencies/cities that are household respondents in West Java Province, namely Bogor Regency, Bekasi Regency, Bandung Regency, Garut Regency, Karawang Regency, Bandung City, Bekasi City, Sukabumi Regency, Cirebon Regency, Cianjur Regency, Indramayu Regency, West Bandung Regency, Ciamis Regency, Depok City, Tasikmalaya Regency, Subang Regency, Majalengka Regency, Sumedang Regency, Kuningan Regency, Purwakarta Regency, Bogor City, Cimahi City, Tasikmalaya City, Cirebon City, Sukabumi City, and Banjar City. The fixed effects in this study include socio-demographic and asset-related variables namely, gender and age of the household head, highest educational attainment, number of household members, car ownership, land ownership, and welfare card possession. These variables are treated as fixed effects because they represent observable household characteristics that are assumed to have a consistent and interpretable influence on per-capita expenditure across all regions.

Meanwhile, regency/city is modeled as a random effect to capture unobserved heterogeneity among regions. Theoretically, households residing in the same administrative area may share contextual similarities such as local labor market conditions, infrastructure access, or policy implementation that influence expenditure patterns but are not directly measured by household-level variables. By incorporating regencies/cities as random effects, the model accounts for this intra-regional correlation and allows the intercepts to vary across areas, thereby providing a more realistic representation of West Java's multilevel socio-economic structure.

Table 1. Description of Variables in Research

Variable	Description	Type	Effects
Y	Average per-capita household consumption expenditure in each regency/city of West Java in Indonesian Rupiah (IDR).	Numeric	Response Variable
X_1	Gender of household head	1: Male 0 :Female	Fixed Effect
X_2	Age of household head	1: 16-29 years old 2: 30-39 years old 3: 40-49 years old 4: 50-59 years old 5: 60-69 years old 6: 70+ years old	Fixed Effect

Variable	Description	Type	Effects
X_3	Highest Degree Ownership	1: Higher education 2: High School or Equivalent 3: Junior High School or Equivalent 4: Elementary School or Equivalent 5: No Diploma	Fixed Effect
X_4	Number of household members	1: 1 member 2: 2 members 3: 3 members 4: 4 members 5: 5 members 6: more than 6 members	Fixed Effect
X_5	Car Ownership	1: Have a Car 0: Don't Have a Car	Fixed Effect
X_6	Land Ownership	1: Have a land 0: Do not have a land	Fixed Effect
X_7	Welfare Card Ownership	1: Have a Welfare Card 0: Don't have a welfare card	Fixed Effect
U_j	Regency/City		Random Effect

2.2 Analysis Method

The analysis methods used are descriptive analysis, inferential and machine learning methods. Descriptive analysis is used to display a general picture of the average household consumption expenditure conditions in West Java Province, while inferential analysis uses a mixed linear method to examine response variables on average household consumption expenditure in West Java Province. Analysis with a machine learning approach uses a mixed effect regression tree model. The stages that need to be carried out in conducting the Linear Mixed Model (LMM) and Mixed effect regression tree (MERT) model analysis can be stated as follows.

1. Collected data from the 2023 National Socio-Economic Survey to collect data on average household income expenditure per capita in West Java Province.
2. Divide the districts/cities into three groups. This group division is based on the value of the Gross Regional Domestic Product (GRDP) in 2023 based on quartile 1, quartile 2, and quartile 3 in West Java Province. GRDP data can be obtained by downloading it from the Central Statistics Agency (BPS) link. The naming of this group uses the names quartile I, quartile II, and quartile III. Areas with high GRDP are defined as quartile I, areas with moderate GRDP as quartile II, and areas with low GRDP as quartile III. The division of areas based on GRDP has a reason, namely the higher the GRDP in the area, the greater the per capita income expenditure.
3. Determine the response variables and predictor variables stated in Table 1 above. The predictor variables consist of fixed effects and random effects as seen in Table 1.
4. The models used in this study are defined as the LMM-West Java model, the LMM-High model, the LMM-Medium model, the LMM-Low model, the MERT-West Java model, the MERT-High model, the MERT-Medium model, and the MERT-Low model.
5. Conduct descriptive data analysis with histogram visualization.
6. Before modeling is done, the data is divided into two categories, namely training data and testing data with a ratio of 70:30. Modeling is done using training data, while model evaluation uses testing data.
7. Conducting data modeling of household income expenditure in West Java Province in 2023 using Linear Mixed Model (LMM) to determine the predictor variables that have a significant effect on income expenditure. The significance level used in this study is 5%.
8. Conduct an error normality test using the Kolmogorov-Smirnov test. The error normality hypothesis test used is as follows

    a. If the assumption of normally distributed errors is violated, the Box-Cox transformation is applied to the response variable.
    b. Conducted LMM Model with transformed data and normality test for residual data.

9. Fit a linear mixed-effects model (LMM) using the transformed data, and then re-check the normality assumption of the residuals.

10. Compare the variables that significantly/influentially affect expenditure and income in each group with the West Java model generated in steps 6–8.
11. Evaluate the best model between MERT and LMM by calculating the RMSE and MAE.

2.3 Linear Mixed Model (LMM)

The distinction between fixed and random factors and their related effects on a response variable is critical in the context of LMMs. We define separate subsections to these topics.

2.3.1 Fixed Factors

The concept of a fixed factor is most used in the context of standard ANOVA or ANCOVA models. A fixed factor is defined as a categorical or classification variable in which all levels or conditions of interest to the investigator are explicitly included in the study. These fixed factors may encompass qualitative covariates such as gender, classification variables implied by a survey sampling design (e.g., region or stratum), study design features (e.g., treatment groups in a randomized clinical trial), or ordinal groupings used in observational research, such as age categories. The levels of a fixed factor are selected to represent specific conditions and are typically analysed to examine contrasts or hypotheses relevant to the objectives of the study [10], [11], [20].

2.3.2 Random Factors

A random factor is a classification variable whose levels are assumed to be randomly sampled from a larger population of potential levels under study. Unlike fixed factors, not all possible levels of the random factor are included in the dataset; instead, the researcher aims to make inferences that generalize beyond the observed levels. In hierarchical or multilevel data structures, such as those found in clustered survey designs or repeated-measures/longitudinal studies, the classification variables identifying Level 2 or Level 3 units such as schools, communities, or individuals are typically treated as random factors. Incorporating random factors into statistical modelling allows the assessment of variation in the response variable attributable to these higher-level units, and enables generalization of findings to the broader population from which these clusters were drawn [1], [17], [20].

Linear mixed models (LMM) are statistical models for continuous outcome variables in which the residuals are normally distributed but may not be response or have constant variance. Study designs leading to data sets that may be appropriately analysed using LMMs studies with clustered data, such as students in classrooms, or experimental designs with random blocks, such as batches of raw material for an industrial process, and (2) longitudinal or repeated measures studies, in which subjects are measured repeatedly over time or under different conditions. LMM is defined as the following

$$
\begin{gathered}
Y_i = X_i\beta + Z_iu_i + \varepsilon_i, \\
u_i \sim N(0, D), \\
\varepsilon_i \sim N(0, R_i),
\end{gathered}
\tag{1}
$$

$$
Y_i = \begin{pmatrix} Y_{1i} \\ Y_{2i} \\ \vdots \\ Y_{ni} \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, u_i = \begin{pmatrix} u_{1i} \\ u_{2i} \\ \vdots \\ u_{qi} \end{pmatrix}, Z = \begin{pmatrix} Z_{1i}^{(1)} & Z_{1i}^{(2)} & \cdots & Z_{1i}^{(q)} \\ Z_{2i}^{(1)} & Z_{2i}^{(2)} & \cdots & Z_{2i}^{(q)} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{n_ii}^{(1)} & Z_{n_ii}^{(2)} & \cdots & Z_{n_ii}^{(q)} \end{pmatrix}, \varepsilon_i = \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \vdots \\ \varepsilon_{qi} \end{pmatrix}.
$$

In Eq. (1), $Y_i$denotes a vector of continuous responses for the i −th subject. The elements of this vector represent repeated or multilevel measurements, with each component corresponding to an individual observation. The fixed-effects design matrix X_i typically includes an intercept term, represented by a column of ones. Furthermore, any time-invariant or subject-specific covariate in X_i will have constant values across all rows for that subject. For theoretical simplicity, it is assumed that the X_i matrices are of full rank i.e., none of the columns (or rows) is a linear combination of the others. However, in practice, the X_i matrices may be rank-deficient, which can result in aliasing or identifiability issues for the fixed-effects parameters in the vector β [4, 20].

The β in Eq. (1) is a vector of p unknown regression coefficients (or fixed-effect parameters) associated with the p covariates used in constructing the X_i matrix. The $n_i \times p$ matrix Z_i be design matrix that represents the known values of the q covariates, $Z_{1i}^{(1)}, \ldots, Z^{(q)}$, for the i-th subject. This matrix is very much like the X_i matrix in that it represents the observed values of covariates;however, it usually has fewer columns than the X_i matrix: The columns in the $\mathrm{Z_i}$ matrix represent observed values for the q predictor variables for the q predictor variables for the i-th subject, which have effects on the continuous response variable that vary randomly across subjects. In many cases, predictors with effects that vary randomly across subjects are represented in both the $\mathrm{X_i}$ matrix and the Z_i matrix. In an LMM in which only the intercepts are assumed to vary randomly form subject to subject, the Z_i matrix would simply be a column of 1's. The u_i vector for the i-th subject in Eq. (1) represents a vector of q random effects associated with the q covariates in the Z_i matrix. Recall that by definition, random effects are random variables. We assume that the q random effects in the u_i vector follow a multivariate normal distribution, with mean vector 0 and a variance-covariance matrix denoted by D: $u_i \sim \mathrm{N}\,(0, \mathrm{D})$.

$$D = Var(u_i) = \begin{pmatrix} Var(u_{1i}) & cov(u_{1i}, u_{2i}) & \ldots & cov(u_{1i}, u_{2i}) \\ cov(u_{1i}, u_{2i}) & Var(u_{2i}) & \ldots & cov(u_{2i}, u_{qi}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(u_{1i}, u_{qi}) & cov(u_{2i}, u_{qi}) & \cdots & Var(u_{qi}) \end{pmatrix}.$$

The covariance matrix $D = Var(u_i)$ represents the variance and covariances of the random effects for the i-th subject and is defined as a function of a typically small set of covariance parameters, denoted by the vector θ_D. The elements of θ_D impose structural constraints on the D matrix, allowing for parsimonious yet flexible modeling of the random-effects variance–covariance components [20]. Several structural forms can be assumed for D, including identity (assuming homogeneity and independence), diagonal (allowing for heterogeneous variances), compound symmetry, or unstructured covariance matrices, depending on the design and complexity of the data.

In addition, the vector ε_i from Eq. (1) denotes the residuals associated with the observed responses for the i-th subject, where each element corresponds to a measurement at a particular occasion t. Since the number of observations per subject may vary (e.g., due to unbalanced data or dropouts in longitudinal studies), the length of each εi vector may also differ across subjects [10], [15]. Modeling this within-subject residual variation accurately is essential to ensure valid inference in linear mixed model frameworks.

In contrast to the standard linear model, the residuals associated with repeated observations on the same subject in an LMM can be correlated. We assume that the n_i residuals in the ε_i vector for a given subject, i, random variables that follow a multivariate normal distribution with mean vector 0 and a positive-definite symmetric variance-covariance matrix R_i.

$$R_i = Var(\varepsilon_i) = \begin{pmatrix} Var(\varepsilon_{1i}) & cov(\varepsilon_{1i}, \varepsilon_{2i}) & \ldots & cov(\varepsilon_{1i}, \varepsilon_{2i}) \\ cov(\varepsilon_{1i}, \varepsilon_{2i}) & Var(\varepsilon_{2i}) & \ldots & cov(\varepsilon_{2i}, \varepsilon_{qi}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\varepsilon_{1i}, \varepsilon_{n_i i}) & cov(\varepsilon_{2i}, \varepsilon_{n_i i}) & \cdots & Var(\varepsilon_{n_i i}) \end{pmatrix}.$$

The within-subject covariance matrix R_i= Var (ε_i) characterizes the variance and covariances of the residual errors for subject i. Each element in Ri corresponds to either the variance of a residual at a specific time point or the covariance between residuals at different occasions within the same subject. The general structure of R_i accommodates the possibility that subjects may have differing numbers of repeated measurements (i.e., unbalanced data), and thus the dimension of R_i may vary across subjects.

The elements of the R_i matrix are defined as functions of a typically small set of covariance parameters stored in the vector θ_R. Various covariance structures can be assumed for Ri, such as identity (assuming homoscedasticity and independence), compound symmetry (assuming equal correlation across time points), autoregressive (for temporally ordered data), or fully unstructured matrices. The choice of structure depends on the underlying correlation pattern expected in the repeated measurements or clustered observations [10], [20].

It is typically assumed that the residual vectors $\varepsilon_1, \ \ldots \ , \varepsilon_m$ are mutually inresponse across subjects and are also response of the random effects vectors $u_1, \ldots, u_m$. This assumption simplifies the specification of the full marginal covariance of the outcome vector Yi, which can be written as:

$$Var(Y_i) = Z_i D Z_i^T + R_i.$$

This formulation allows for simultaneous modelling of between-subject variability through the random effects and within-subject or intra-cluster variability through the residual structure.

2.4 Mixed-Effect Regression Tree (MERT)

A statistical model for clustered or hierarchical data generally consists of two key components: a fixed (population-averaged) component and a random (cluster-specific) component. In the context of mixed effects modelling, the fixed component captures the average relationship between predictors and outcomes across all clusters, while the random component accounts for heterogeneity between clusters or subjects. The Mixed Effects Regression Tree (MERT) model extends this framework by integrating a node-invariant linear structure to model random effects and using a regression tree algorithm to model fixed effects.

This approach replaces the conventional linear estimation used in standard Linear Mixed Effects (LME) models with recursive partitioning techniques typically employed in decision trees. The estimation procedure is implemented using the Expectation-Maximization (EM) algorithm to iteratively estimate both fixed and random effects, as originally formalized by McLachlan and Krishnan (1997). By embedding tree-based splitting within the mixed effects modelling structure, MERT enables flexible, non-linear modelling while retaining the capacity to account for correlated or clustered observations [4].

1. EM Algorithm for linear mixed effect model

The LME model is generally written in the following form:

$$y_i = X_i\beta + Z_ib_i + \epsilon_i, \tag{2}$$

$$b_i \sim N(0, D), \epsilon_i \sim N(0, R_i),$$

$$i = 1, \dots, n,$$

where $y_i = [y_{i1}, \dots, y_{in_i}]^T$ is the $n_i \times 1$ vector of responses for the n_i observations in cluster i, $X_i = [x_{i1}, \dots, x_{in_i}]^T$ is the $n_i \times p$ matrix of fixed-effects covariates, $Z_i = [z_{i1}, \dots, z_{in_i}]^T$ is the $n_i \times q$ matrix of random-effects covariates, $\epsilon_i = [\epsilon_{i1}, \dots, \epsilon_{in_i}]^T$ is the $n_i \times 1$ vector of errors, $b_i = [b_{i1}, \dots, b_{iq}]^T$ is the $q \times 1$ unknown vector of observations is $N = \sum_{i=1}^{n} n_i$. The covariance matrix of b_i is D while R_i is the covariance matrix of , ϵ_i. The usual LME model also assumes that b_i and ϵ_i are inresponse and normally distributed and that the between-clusters observations are in response. Hence, the covariance matrix of the vector of observations y_i in cluster i is $V_i = Cov(y_i) = Z_iDZ_i^T + R_i$, and $V = Cov(y) = diag(V_1, \dots, V_n)$, where $y = [y_1^T, \dots, y_n^T]^T$. We will further assume that the correlation is induced solely via the between-clusters variation, that is, R_i is diagonal $\left(R_i = \sigma^2 I_{n_i},\ i = 1, \dots, n\right)$. This assumption is suitable for a large class of clustered data problems. The parameters in LME models can be estimated by the method of maximum likelihood (ML) implemented with the EM algorithm.

The major cycle for the ML-based EM algorithm, as described in Section 2.2.5 of Wu and Zhang (2006), is as follows:

Step 0. Set $r = 0$. Let $\hat{\sigma}_{(0)}^2 = 1$, and $\widehat{D}_{(0)} = I_q$.

Step 1. Set $r = r + 1$. Update $\hat{\beta}_{(r)}$ and $\hat{b}_{i(r)}$

$$\hat{\beta}_{(r)} = \left(\sum_{i=1}^{n} X_i^T \hat{V}_{i(r-1)}^{-1} X_i\right)^{-1} \left(\sum_{i=1}^{n} X_i^T \hat{V}_{i(r-1)}^{-1} y_i\right), \tag{3}$$

$$\hat{b}_{i(r)} = \widehat{D}_{(r-1)\_} Z_i^T \hat{V}_{i(r-1)}^{-1} \left(y_i - X_i\hat{\beta}_{(r)}\right), \qquad i = 1, \dots, n; \tag{1}$$

where $\hat{V}_{i(r-1)} = Z_i \widehat{D}_{(r-1)\_} Z_i^T + \hat{\sigma}_{(r-1)\_}^2 I_{n_i},\ i = 1, \dots, n.$

Step 2. Update $\hat{\sigma}^2_{(r)}$, and $\widehat{D}_{(r)_}$using

$$\hat{\sigma}^2_{(r)} = N^{-1}\sum_{i=1}^{n}\{\hat{\epsilon}^T_{i(r)}\hat{\epsilon}_{i(r)} + \hat{\sigma}^2_{(r-1)\_}[n_i - \hat{\sigma}^2_{(r-1)\_}trace(\hat{V}_{i(r-1)})]\}, \tag{5}$$

$$\widehat{D}_{(r)} = n^{-1}\sum_{i=1}^{n}\{\hat{b}_{i(r)}\hat{b}^T_{i(r)} + [\widehat{D}_{(r-1)} - \widehat{D}_{(r-1)}Z_i^T\hat{V}^{-1}_{i(r-1)}Z_i\widehat{D}_{(r-1)\_}]\}, \tag{6}$$

where $\hat{\epsilon}_{i(r)} = y_i - X_i\hat{\beta}_{(r)} - Z_i\hat{b}_{i(r)}, N = \sum_{i=1}^{n} n_i$.

Step 3. Repeat steps 1 and 2 until convergence.

2. The proposed MERT model is

$$y_i = f(X_i) + Z_ib_i + \epsilon_i, \tag{7}$$
$$b_i \sim N(0, D), \epsilon_i \sim N(0, R_i),$$
$$i = 1, \dots, n,$$

where all quantities are defined as in Section 2.1 except that the linear fixed part $X_i\beta$ in Eq. (1) is replaced by the function $f(X_i)$ that will be estimated with a standard tree-based model. The random part, Z_ib_i, is still assumed to be linear. The MERT algorithm is the ML-based EM-algorithm in which we replace the linear structure used to estimate the fixed part of the model by a standard tree structure. The algorithm is as follows:

Step 0. Set $r = 0$. Let $\hat{b}_{i(0)} = 0$, $\hat{\sigma}^2_{(0)} = 1$, and $\widehat{D}_{(0)} = I_q$.

Step 1. Set $r = r + 1$. Update $y^*_{i(r)}, \hat{f}(X_i)_{(r)}$, and $\hat{b}_{i(r)}$

a. $y^*_{i(r)} = y_i - Z_i\hat{b}_{i(r-1)}, i = 1, \dots, n.$

b. Let $\hat{f}(X_i)_{(r)}$ an estimate of $f(X_i)$ obtained from a standard tree algorithm with $y^*_{i(r)}$ as responses and $X_i, i = 1, \dots, n$. As covariates. Note that the tree is built as usual using all N individual observations as inputs along with their covariate vectors.

c. $\hat{b}_{i(r)} = \widehat{D}_{(r-1)}Z_i^T\hat{V}_{i(r-1)}(y_i - \hat{f}(X_i)_{(r)}), i = 1, \dots, n.$ where $\hat{V}_{i(r-1)} = Z_i\widehat{D}_{(r-1)}Z_i^T + \hat{\sigma}^2_{(r-1)}I_{n_i}$, $i = 1, \dots, n.$

Step 2. Update $\hat{\sigma}^2_{(r)}$ and $\widehat{D}_{(r)}$ using

$$\hat{\sigma}^2_{(r)} = N^{-1}\sum_{i=1}^{n}\{\hat{\epsilon}^T_{i(r)}\hat{\epsilon}_{i(r)} + \hat{\sigma}^2_{(r-1)\_}[n_i - \hat{\sigma}^2_{(r-1)\_}trace(\hat{V}_{i(r-1)})]\}, \tag{8}$$

$$\widehat{D}_{(r)} = n^{-1}\sum_{i=1}^{n}\{\hat{b}_{i(r)}\hat{b}^T_{i(r)} + [\widehat{D}_{(r-1)} - \widehat{D}_{(r-1)}Z_i^T\hat{V}^{-1}_{i(r-1)}Z_i\widehat{D}_{(r-1)\_}]\}, \tag{9}$$

where $\hat{\epsilon}_{i(r)} = y_i - \hat{f}(X_i)_{(r)} - Z_i\hat{b}_{i(r)}$.

Step 3. Repeat steps 1 and 2 until convergence.

3. The proposed MERT model is

$$y_i = f(X_i) + Z_ib_i + \epsilon_i, \tag{10}$$
$$b_i \sim N(0, D), \epsilon_i \sim N(0, R_i),$$
$$i = 1, \dots, n,$$

where all quantities are defined as in above except that the linear fixed part $X_i\beta$ in Eq. (1) is replaced by the function $f(X_i)$ that will be estimated with a standard tree-based model. The random part, Z_ib_i, is still assumed to be linear. The MERT algorithm is the ML-based EM-algorithm in which we replace the linear structure used to estimate the fixed part of the model by a standard tree structure. The algorithm is as follows:

Step 0. Set $r = 0$. Let $\widehat{b}_{i(0)} = 0$, $\hat{\sigma}^2_{(0)} = 1$, and $\widehat{D}_{(0)} = I_q$.

Step 1. Set $r = r + 1$. Update $y^*_{i(r)}, \hat{f}(X_i)_{(r)}$, and $\widehat{b}_{i(r)}$

a. $y^*_{i(r)} = y_i - Z_i \widehat{b}_{i(r-1)}, i = 1, \dots, n.$

b. Let $\hat{f}(X_i)_{(r)}$ an estimate of $f(X_i)$ obtained from a standard tree algorithm with $y^*_{i(r)}$ as responses and $X_i, i = 1, \dots, n$. As covariates. Note that the tree is built as usual using all N individual observations as inputs along with their covariate vectors.

c. $\widehat{b}_{i(r)} = \widehat{D}_{(r-1)} Z_i^T \widehat{V}_{i(r-1)} \left(y_i - \hat{f}(X_i)_{(r)}\right), i = 1, \dots, n.$ where $\widehat{V}_{i(r-1)} = Z_i \widehat{D}_{(r-1)} Z_i^T + \hat{\sigma}^2_{(r-1)} I_{n_i}, \; i = 1, \dots, n.$

Step 2. Update $\hat{\sigma}^2_{(r)}$ and $\widehat{D}_{(r)}$ using

$$\hat{\sigma}^2_{(r)} = N^{-1} \sum_{i=1}^{n} \{\hat{\epsilon}^T_{i(r)} \hat{\epsilon}_{i(r)} + \hat{\sigma}^2_{(r-1)\_}[n_i - \hat{\sigma}^2_{(r-1)\_} trace\left(\widehat{V}_{i(r-1)}\right)]\}, \tag{11}$$

$$\widehat{D}_{(r)} = n^{-1} \sum_{i=1}^{n} \{\widehat{b}_{i(r)} \widehat{b}^T_{i(r)} + [\widehat{D}_{(r-1)} - \widehat{D}_{(r-1)} Z_i^T \widehat{V}^{-1}_{i(r-1)} Z_i \widehat{D}_{(r-1)\_}]\}, \tag{12}$$

where $\hat{\epsilon}_{i(r)} = y_i - \hat{f}(X_i)_{(r)} - Z_i \widehat{b}_{i(r)}$.

Step 3. Repeat steps 1 and 2 until convergence.

2.5 Model Evaluation

Model evaluation is measured by Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Root Mean Square Error (RMSE) is a method of summing squares of error or the difference between the real value and the predicted value. RMSE is calculated by squaring the error and then dividing it by the average number. RMSE is stated in Eq. (13). Mean Absolute Error is a method used in measuring the level of accuracy of the prediction model. The formula for MAE is displayed in Eq. (14).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{x}_i)^2}, \tag{13}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |x_i - x|. \tag{14}$$

3. RESULTS AND DISCUSSION

3.1 Overview of Average Household Poverty in West Java Province

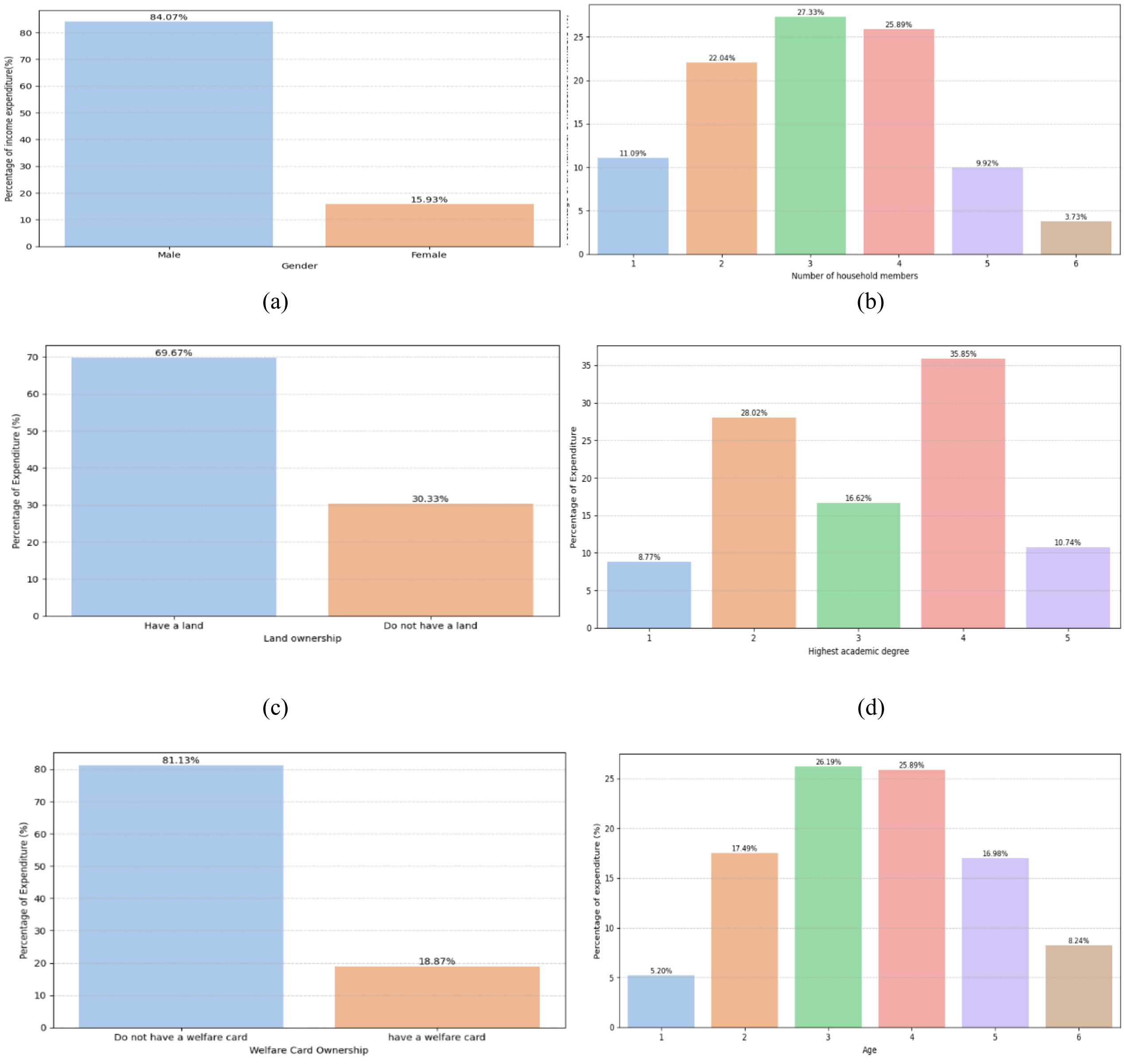
Based on Fig. 1 (a), which presents a bar chart illustrating household consumption expenditure by the gender of the household head, the data shows that 84.07% of total expenditure originates from male-headed households, while only 15.93% comes from female-headed households. This disparity can be interpreted not only in terms of population distribution where male headed households are more prevalent but also as a reflection of structural gender inequality in household welfare. Female headed households often face greater vulnerability to poverty due to limited access to economic resources, such as employment opportunities, productive assets, and financial capital. Consequently, the lower proportion of expenditure among female-headed households may indicate their restricted consumption capacity and higher exposure to economic insecurity. Thus, this finding highlights the importance of gender-sensitive poverty reduction policies, emphasizing targeted support for female-headed households, particularly through programs that expand access to education, employment, and asset ownership.

These findings can provide valuable input for gender-responsive poverty reduction policies, particularly through women's economic empowerment and income equality initiatives. Fig. 1 (b) presents a bar chart illustrating expenditure shares by household size. Each bar represents the contribution of households with a specific number of members to total expenditure. The results show that households with 3–4 members account for more than 50% of total expenditure, indicating that they represent the most economically active and productive group in the sample. Specifically, households with 3 members contribute 26.19%, followed closely by those with 4 members at 25.89%. This pattern suggests that households with 3–4 members typically reflect an optimal productive structure, where the balance between working members and dependents allows for higher efficiency in income generation and consumption. In contrast, larger households (with 5–6 members) tend to face a higher economic dependency ratio, which reduces per capita expenditure and increases the likelihood of falling into poverty. Furthermore, this variable household size (X_4) is incorporated into the MERT and LMM models, where it may exhibit a non-linear relationship with poverty indicators. For instance, beyond the optimal size (around 3–4 members), the effect of household size on welfare may become diminishing or even negative, reflecting how overextended household structures can constrain economic resilience and deepen regional disparities in poverty levels. Based on Fig. 1 (d), there is a noticeable relationship between the highest educational attainment of the head of household and household consumption expenditure. Households that the head has only completed elementary school (or equivalent) have the highest expenditure share at 35.85%, surpassing all other education levels. This pattern may indicate that such households allocate a larger proportion of their limited income to basic necessities, reflecting both low income levels and restricted financial literacy. High school graduates record the second-highest expenditure share (28.02%), suggesting that they may already have stable but modest earnings, which support household consumption but leave limited room for savings or investment. Meanwhile, junior high school graduates fall in the intermediate range potentially due to more cautious spending behavior or limited access to steady employment opportunities, resulting in moderate levels of household consumption expenditure.

Interestingly, households headed by individuals with college education show the lowest expenditure share (8.77%), even though they are likely to have higher incomes. This finding suggests a positive relationship between education and financial management, where higher educated individuals may have more diversified income sources, greater savings capacity, and more efficient consumption patterns. In addition, land ownershipoften used as a proxy for household wealth plays a critical role in explaining welfare disparities. As seen in Fig. 1 (e), households that own land account for 69.67% of total expenditure, indicating that asset ownership significantly enhances household economic stability. Therefore, education and land ownership jointly function as key determinants of household welfare, shaping both income-generating capacity and long-term resilience against poverty. This suggests that college graduates tend to manage their finances more efficiently, allocating a greater portion of their income toward savings, investments, or long-term assets rather than daily consumption. Meanwhile, households with no formal education also show relatively low expenditure (10.74%), which can be attributed to income limitations that constrain their spending capacity.

In addition, despite their low income, their expenditure share remains slightly higher than that of college-educated households, reflecting that low-income groups typically spend most of their earnings on essential consumption. Overall, this pattern illustrates that the higher the educational attainment particularly at the tertiary level the lower the per capita expenditure ratio tends to be, suggesting greater financial literacy, budget control, and income diversification among more educated households. Conversely, households with lower educational levels, such as elementary school graduates, spend a larger proportion of their limited income, indicating greater vulnerability to income shocks and poverty risk. These findings reinforce the critical role of education in enhancing economic decision-making, financial planning, and long-term welfare stability. Fig. 1 (c) further supports this narrative by illustrating a bar chart of land ownership, where households that own land account for 69.67% of total expenditure, compared to 30.33% among non–land-owning households. The higher expenditure share among landowners may reflect both higher income levels and additional financial responsibilities, such as land maintenance, taxation, or productive investments (e.g., farming or microenterprises). In socio-economic perspective, land ownership serves as a proxy for wealth, influencing access to credit, investment opportunities, and intergenerational capital accumulation. In the context of asset-based poverty measurement, asset ownership particularly land is recognized as a key determinant of household economic resilience, as it provides a tangible buffer against financial shocks and enhances long-term welfare security. Therefore, the integration of land ownership as an explanatory variable in both the MERT and LMM frameworks is theoretically justified, as it captures structural differences in household welfare beyond monetary income indicators.

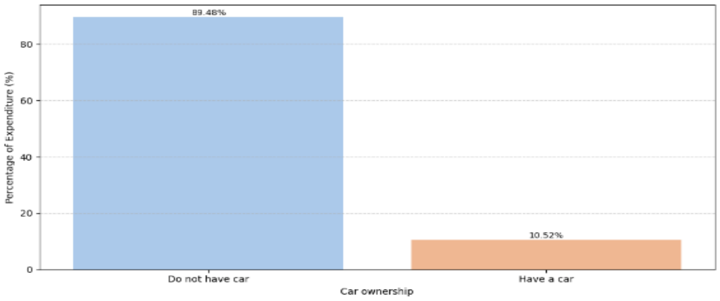
This suggests that college graduates tend to manage their finances more efficiently and allocate more toward savings or investments. Those with no formal education also exhibit relatively low expenditure at 10.74%, likely due to income constraints that force them to minimize spending as much as possible. However, this is still slightly higher than the expenditure of college graduates. Overall, it can be concluded that the higher the level of formal education particularly at the college level the lower the per capita expenditure tends to be. Conversely, households with lower education levels, such as elementary school, tend to spend a greater proportion of their income. This underscores the important role of education in enhancing financial literacy and expenditure efficiency. Fig. 1 (c) is a bar chart representing households that own land (69.67%) and those that do not (30.33%). Households that own land have a significantly higher proportion of expenditure. This is likely due to higher income levels, as land ownership often reflects assets and economic stability. Landowners may also have greater financial burdens due to maintenance costs, land taxes, or productive investments (such as farming or other businesses on the land). On the other hand, households without land tend to have lower expenditures. This indicates more limited economic conditions, leading them to minimize their spending.



(a) (b)

(c) (d)

(e) (f)



(g)

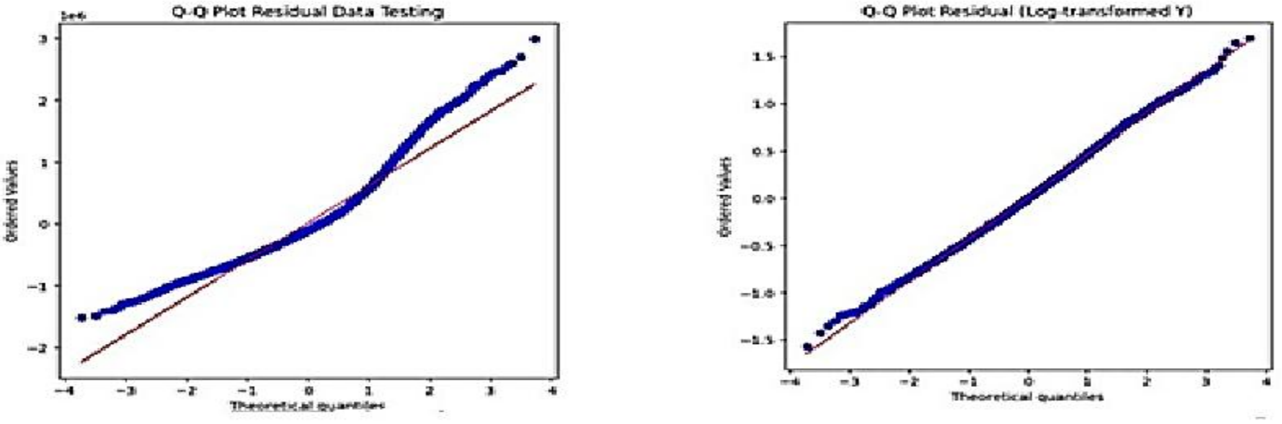
Figure 1. Descriptive Analysis of Household Expenses
(a) Gender, (b) Number of Household Members, (c) Land Ownership, (d) Highest Academic Degree, (e) Car Ownership, (f) Age, (g) Welfare Card Ownership

It could also be due to the absence of additional costs such as land taxes, maintenance, or land installment payments. Moreover, asset ownership influences economic behavior, such as providing greater access to loans, investments, and consumption. Fig. 1 (f) displays bar chart representing the percentage of expenditure based on the age group of the head of household. Based on Fig. 1 (c), the highest proportion of household consumption expenditure occurs in the 40–49 age group (26.19%) and the 50–59 age group (25.89%). This reflects that these ages are the most productive years and the period with the highest financial responsibilities in the household life cycle. Meanwhile, expenditure is relatively low in the 16–29 age group (5.20%), likely because many individuals at this age are still early in their careers or have not yet become heads of households. After the age of 60, the proportion of expenditure drops significantly, with only 8.24% in the 70+ age group, most likely due to declining income after retirement and reduced household needs. Overall, household consumption expenditure follows a bell-curve pattern influenced by the individual's economic life cycle Car ownership, as shown in Fig. 1 (e), distinguishes between households that own a car and those that do not. The majority of total expenditure 89.48% comes from households without a car, indicating that this group dominates the population of surveyed households in West Java.

This finding suggests that car ownership remains relatively limited, reflecting the small size of the middle class in the province. In economic terms, the low rate of car ownership can be viewed as an indicator of non-monetary poverty, highlighting limited access to durable goods and assets that signify economic security. Car ownership often represents not only a symbol of financial capability but also a means of economic mobility, enabling better access to employment, education, and social services. Therefore, the low prevalence of car ownership underscores persistent structural inequalities and the fragility of the middle-income group in West Java. In the broader poverty analysis framework, this variable provides valuable insights into asset-based welfare disparities, complementing income and expenditure measures in identifying households that are economically vulnerable despite not being income-poor. In Fig. 1 (g), it is shown that the majority of per capita expenditure 81.13% comes from households that do not have a welfare card. This means that the distribution of expenditure is uneven, with welfare recipient households contributing less. This suggests a need for policy interventions aimed at improving the economic well-being of welfare recipients. Therefore, these descriptive results justify the inclusion of X_1–X_7 as fixed effects in the MERT and LMM frameworks to capture both individual and regional heterogeneity.

3.2 Normality Test

The Box Cox transformation was applied to stabilize variance and improve the normality of residuals. After transformation, the residuals more closely followed a normal distribution, as shown in the QQ-plot on the right side of Fig. 2. Although the transformation adjusted the scale of the response variable, it did not materially change the sign or relative magnitude of the regression coefficients, meaning that the relationships between predictor variables and expenditure remained consistent.



(a) (b)

Figure 2. (a) Residuals before Transformation and (b) Residuals after Box–Cox Transformation

However, the overall model performance slightly improved, with reduced RMSE and MAE values, suggesting a better model fit after transformation. Therefore, the Box–Cox transformation was retained in subsequent analyses to ensure more reliable inference and model validity.It can be seen that the blue points follow the red diagonal line, indicates that the residual data is normally distributed. In addition, we also conducted a normality test on the residuals using the Kolmogorov-Smirnov test at a 5% significance level. The resulting p-value was 0.0845, which is greater than the significance level. Therefore, there is no significant evidence to reject the assumption of normality. Thus, the residuals are considered to be normally distributed or fail to reject normality.

3.3 Linear Mixed Model (LMM)

The LMM models used in this study consist of the LMM–West Java model, the LMM–High model, the LMM–Medium model, and the LMM–Low model. The parameter estimation results for these four models, obtained using Python software, are presented in Table 2 below. The variables included are gender of the head of household (X_1), highest educational attainment (X_3), number of household members (X_4), car ownership (X_5), land ownership (X_6), and welfare card ownership (X_7). However, the variable for the age of the head of household (X_2) does not have a significant effect on income expenditure in West Java Province. An adjustment of interpretation was made based on the estimated coefficient results.

Table 2. Estimation Parameter of Linear Mixed Model

	LMM-*West Java*		**LMM-*High***		**LMM-*Medium***		**LMM-*Low***	
Variables	**Coefficient (Std. Error)**	**t-value**	**Coefficient (Std. Error)**	**t-value**	**Coefficient (Std. Error)**	**t-value**	**Coefficient (Std. Error)**	**t-value**
Fixed Effects:								
(Intercept)	14.704 (0.035)	418.85*	14.819 (0.059)	249.8*	14.719 (0.051)	286.349*	14.512 (0.056)	259.918*
Gender of the Head of Household (X_1)	0.044 (0.011)	4.08*	0.028 (0.022)	1.304	0.026 (0.015)	1.714	0.072 (0.021)	3.448*
Age of the Head of Household (X_2)	-0.004 (0.003)	-1.178	0.004 (0.006)	0.630	-0.006 (0.004)	-1.493	-0.013 (0.006)	-2.238
Highest Academic Degree (X_3)	-0.096 (0.004)	-26.6*	-0.095 (0.007)	-13.66*	-0.093 (0.005)	-18.196*	-0.080 0.007	-10.991*
Number of Household Members (X_4)	-0.170 (0.003)	-54.073*	-0.171 (0.006)	-26.61*	-0.163 (0.004)	-36.299*	-0.162 0.006	-27.567*
Car Ownership (X_5)	0.491 (0.015)	33.772*	0.439 (0.028)	15.89*	0.478 (0.020)	24.132*	0.596 (0.035)	17.201*
Land Ownership (X_6)	0.120 (0.008)	15.23*	0.129 (0.016)	8.137*	0.123 (0.011)	10.782*	0.108 (0.015)	7.217*
Welfare Card Ownership (X_7)	-0.064 (0.009)	-7.27*	-0.105 (0.019)	-5.61*	-0.081 (0.012)	-6.457*	-0.030 (0.016)	-1.823*

	LMM-*West Java*		**LMM-*High***		**LMM-*Medium***		**LMM-*Low***	
Variables	**Coefficient (Std. Error)**	**t-value**	**Coefficient (Std. Error)**	**t-value**	**Coefficient (Std. Error)**	**t-value**	**Coefficient (Std. Error)**	**t-value**
Random Effects:								
Groups	**Variance**		**Variance**		**Variance**		**Variance**	
Area	0.022		0.013		0.023		0.011	
Residual	0.014		0.017		0.022		0.014	

The results of the LMM–West Java model indicate that three predictor variables highest educational attainment (X_3), number of household members (X_4), and welfare card ownership (X_7)have a significant negative effect on per-capita household consumption expenditure. The negative coefficient of educational attainment suggests that households with more highly educated heads tend to have lower per-capita consumption, possibly because they allocate a larger share of income to savings or investment rather than daily spending. In contrast, households with lower education levels tend to spend a higher proportion of their limited income on basic needs, reflecting more constrained financial conditions. Similarly, the negative relationship between household size and consumption expenditure implies that larger families face greater resource dilution, leading to reduced per-capita spending and a higher likelihood of poverty. The significant effect of welfare card ownership indicates that households receiving government assistance generally exhibit lower consumption levels, consistent with their classification as economically vulnerable groups. In the LMM–High model, significant predictors include educational attainment (X_3), household size (X_4), car ownership (X_5), land ownership (X_6), and welfare card ownership (X_7), whereas gender and age of the household head do not show significant effects. These results highlight that asset ownership and education play a crucial role in shaping household welfare in high-GRDP regions, reinforcing the view that non-monetary assets such as land and vehicles serve as important buffers against poverty risks.

The variable for welfare card ownership has a negative effect on per capita expenditure. This indicates that households with a welfare card tend to have lower per capita income expenditure compared to households without a welfare card. This conclusion is based on the parameter estimation results presented in Table 2. It is found that in the LMM–High model, the variables that significantly affect per capita income expenditure at the 5% significance level are highest educational attainment (X_3), number of household members (X_4), car ownership (X_5), land ownership (X_6), and welfare card ownership (X_7). However, the predictor variables gender of the head of household (X_1) and age of the head of household (X_2) do not have a significant effect on high-level income expenditure in West Java Province. There are three variables that have a negative effect on per capita income expenditure: highest educational attainment (X_3), number of household members (X_4), and car ownership (X_5). Based on Table 2, it can be concluded that in the LMM–Medium model, the predictor variables that significantly affect per capita income expenditure at the 5% significance level are highest educational attainment (X_3), number of household members (X_4), car ownership (X_5), land ownership (X_6), and welfare card ownership (X_7). However, the predictor variables gender of the head of household (X_1) and age of the head of household (X_2) do not have a significant effect on medium-level income expenditure in West Java Province.

Based on the parameter estimates presented in Table 2, the LMM–Low model identifies six variables that significantly affect per capita household consumption expenditure at the 5% significance level: gender of the head of household (X_1), highest educational attainment (X_3), number of household members (X_4), car ownership (X_5), land ownership (X_6), and welfare card ownership (X_7). In contrast, the age of the household head (X_2) does not have a significant effect. At the low-expenditure level, the factors that reduce per capita consumption are educational attainment, household size, and age of the household head, indicating that lower education, larger family size, and older heads of household are associated with higher poverty risk. Conversely, car ownership and land ownership are associated with higher per capita expenditure, reflecting better household welfare and asset accumulation. Land, in particular, functions as a productive and investment-oriented asset, often serving as a key indicator of long-term economic security [21].

Therefore, land ownership signifies higher economic capacity, providing better access to credit (e.g., through collateral) and investment opportunities. Consequently, land-owning households possess greater ability to spend, invest, or consume more per household member. Similarly, car ownership serves not only as a means of transportation but also as a proxy for wealth. Acquiring and maintaining a car requires substantial financial resources, meaning that such ownership is typically concentrated among high-income households. Furthermore, car ownership can enhance economic productivity by improving access to employment and market opportunities, thereby increasing both income and expenditure. Regarding gender, male-headed households are associated with higher per capita expenditure compared to female-headed households. In many socio-economic contexts, men are more likely to participate in the formal labor market and to have greater access to economic resources, which contributes to higher household spending. Conversely, female-headed households often face greater economic vulnerability, partly due to unequal access to labor markets and the concentration of women in informal, lower-paying sectors. It is important to note that higher expenditure in this context reflects greater purchasing power and economic capacity, rather than excessive consumption. Overall, the results from the LMM models indicate that asset ownership (land and vehicles), education, and household composition are key determinants of household welfare. These findings reinforce the notion that poverty is multidimensional, where material assets and human capital interact to influence household consumption behavior and economic resilience.

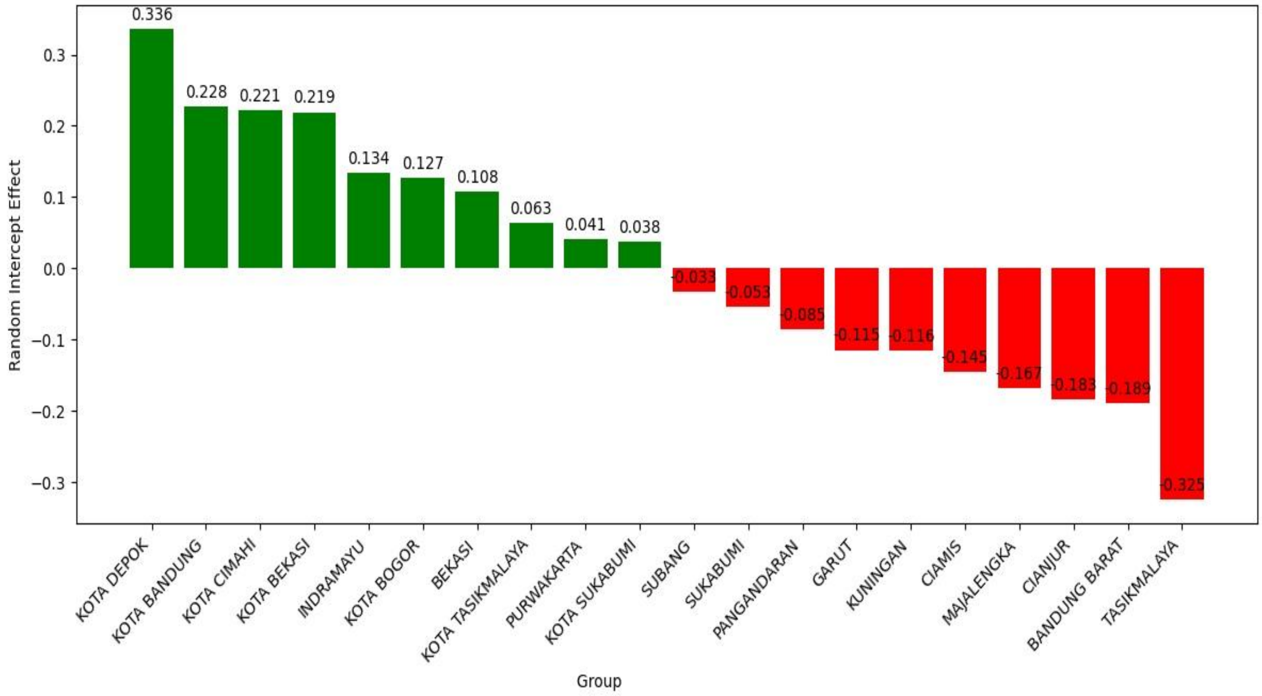


Figure 3. Random Effects of LMM Plot

The difference in economic capacity between male- and female-headed households may stem from gender inequality in access to employment opportunities, income, and ownership of economic assets. Therefore, this finding highlights the need for gender-responsive policies that promote economic equality and empower female-headed households. In the random effects model, the random component captures the unobserved heterogeneity that cannot be explained by the fixed effects alone. In this study, the district/city level within West Java Province is treated as a random effect to account for regional variations in socio-economic conditions that influence household expenditure. The estimated random effects for the West Java model are presented in the following figure.

Based on Fig. 3, the top 10 and bottom 10 areas in West Java are presented according to the Linear Mixed Model (LMM) results. The horizontal axis represents the random intercept values for each area, reflecting variations in household consumption expenditure after controlling for fixed effects. In this model, regional differences are treated as random effects to capture unobserved heterogeneity across districts and cities. Each bar in the figure indicates the relative deviation of an area's expenditure level from the provincial average, highlighting regions with notably higher or lower consumption patterns.

Positive random effects (green bars) indicate regions where households tend to have higher household consumption expenditure than the provincial average, even after controlling for individual characteristics. Conversely, negative random effects (red bars) represent areas where households spend less than average, which may reflect lower income levels, reduced living costs, or more conservative consumption behavior. For example, urban areas such as Depok City and Bandung City show positive random effects, suggesting relatively higher household purchasing power and stronger local economies. In contrast, rural areas such as Tasikmalaya and Ciamis Regencies exhibit negative effects, indicating more limited economic capacity.These spatial disparities highlight the importance of region-specific poverty alleviation strategies. Areas with persistently negative random effects could be prioritized for targeted fiscal transfers or empowerment programs, while regions with positive effects may contribute more to regional revenue or serve as economic growth centers.

3.4 Mixed-Effect Regression Tree (MERT)

Mixed Effect Regression Tree (MERT) analysis in this study comprises four models: MERT Jawa Barat, MERT High, MERT Medium, and MERT Low. Each MERT model is analyzed using its respective decision tree.

3.4.1 MERT-*West Java* Interpretation

The decision tree model illustrates the determinants of household consumption expenditure based on various socioeconomic variables. The root node shows that car ownership (X5) is the most influential factor affecting household spending. Households without a car tend to have lower levels of consumption expenditure, with an average predicted value of 13.94. In contrast, households that own at least one car exhibit higher consumption levels, averaging 14.47. This finding indicates that car ownership serves as a proxy for higher economic capacity and is positively associated with greater household spending.

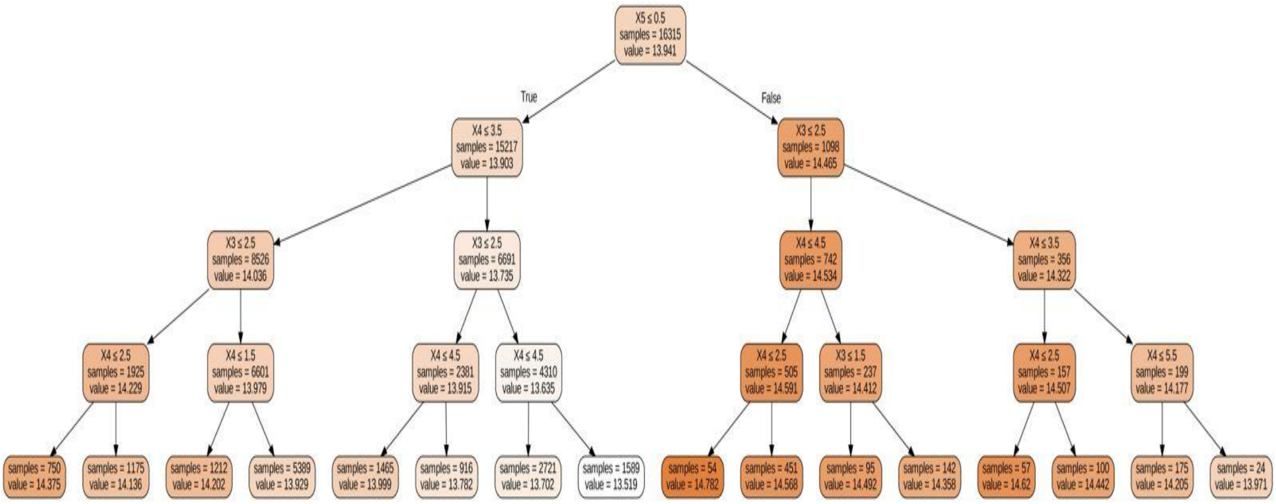


Figure 4. Decision Tree of MERT-*West Java*

Among households without a car, the number of household members (X4) and education level (X3) further influence the level of consumption. When the number of household members is relatively small, the average expenditure is around 13.90. Within this group, differences in education level also lead to variations in spending. Households with higher education levels tend to allocate more resources to consumption, suggesting that education contributes to better income opportunities and higher purchasing power, even among non-car-owning households. For households with car ownership, both education and household size continue to play significant roles in determining expenditure. When the education level of the household head is lower, average consumption expenditure ranges from 14.53 to 14.59. However, when education levels increase and the number of household members is relatively small , the average expenditure slightly decreases to around 14.32. This implies that although education improves income potential, households with higher education may also manage their consumption more efficiently or allocate income to non-consumption investments.

Overall, the decision tree suggests that car ownership is the strongest indicator of household welfare, reflecting both economic capacity and lifestyle differences. Education level and household size are secondary factors that further refine consumption patterns. Households with better education and moderate family size tend to maintain balanced consumption, while those with car ownership generally exhibit higher spending. These findings highlight that consumption expenditure is not only influenced by income-related indicators but also by demographic and behavioral characteristics that reflect household preferences and living standards.

3.4.2 MERT-*High* Interpretation

Based on the MERT–High model presented in Fig. 5, the predictor variable highest educational attainment serves as the first split in the model, confirming that education is the most important factor differentiating households in terms of welfare. If the education level of the household head is low, the household follows the left branch with a predicted value of approximately 26,981. Conversely, if the household head has a higher education level, the model proceeds to a split based on the gender of the household head.

The results show that female-headed households with higher education are associated with higher household consumption expenditure. This finding suggests that education plays a compensatory role for women, helping to offset gender-based disadvantages in labor market access, income opportunities, and control over economic resources. Educated female heads of household are more likely to participate in formal employment, manage household finances effectively, and allocate spending toward productive uses, which collectively enhance household welfare.

From a broader perspective, this pattern highlights the importance of education as a key driver of gender-inclusive poverty reduction. It emphasizes that empowering women through access to education not only narrows welfare disparities between male- and female-headed households but also contributes to more equitable and sustainable economic development in West Java.

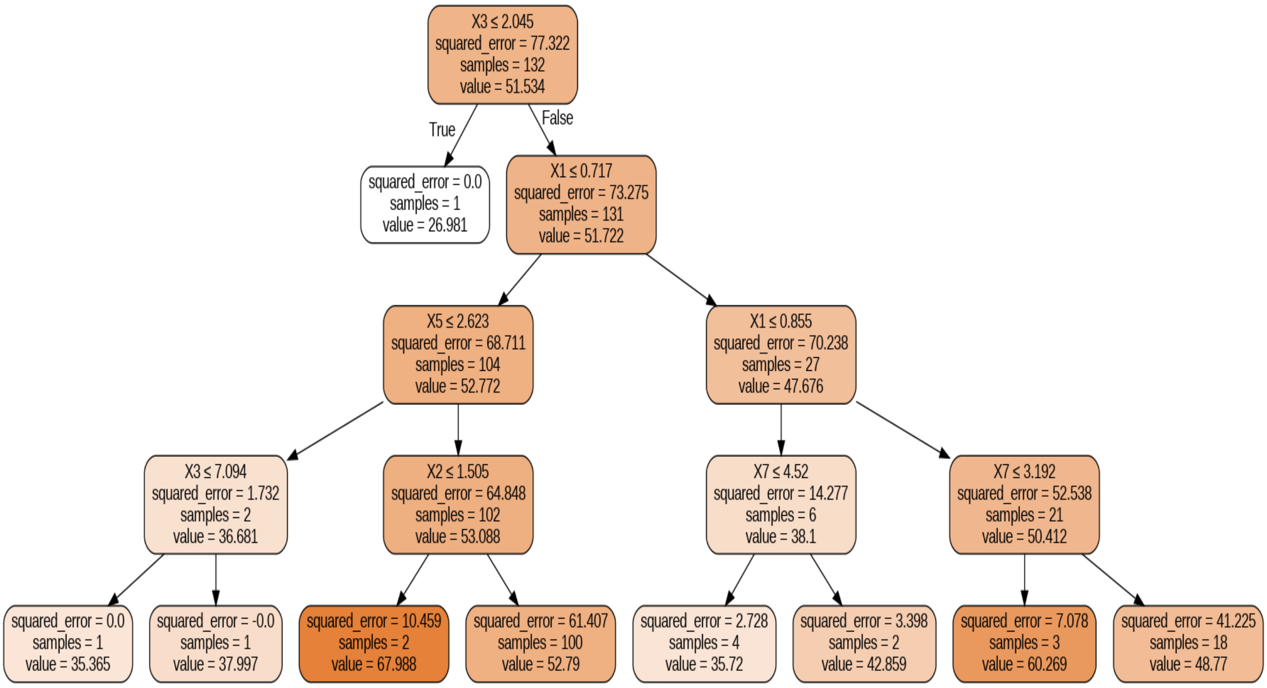


Figure 5. Decision Tree of MERT-*High*

3.4.3 MERT-Medium Interpretation

Based on the MERT–Medium model presented in Fig. 6, the first and most influential predictor variable is the gender of the household head (X_1). This indicates that gender plays a key role in differentiating household welfare patterns in West Java. The model shows that female-headed households generally have lower predicted household consumption expenditure compared to male-headed households, reflecting persistent gender-based disparities in economic opportunity and access to resources.

Within the female-headed household group, education (X_3) and asset ownership particularly land (X_4) and car ownership (X_5)are the most important determinants of expenditure. Female household heads with higher education and ownership of productive assets tend to exhibit higher predicted consumption levels, underscoring the role of education and asset accumulation in reducing female household poverty and strengthening financial resilience. Conversely, low-educated women without assets tend to fall into the lowest expenditure group, suggesting their limited capacity to generate income and invest in household welfare.

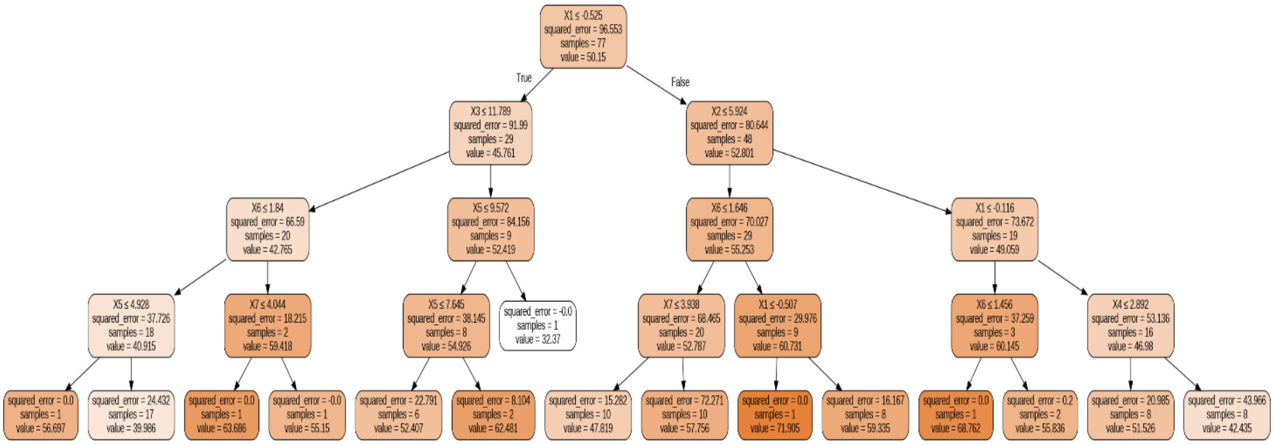


Figure 6. Decision Tree of MERT-*Medium*

In contrast, among male-headed households, educational attainment and car ownership emerge as the dominant factors. Male heads with higher education, car ownership, and smaller household size demonstrate the highest predicted consumption expenditure, suggesting that the combination of human capital, asset ownership, and manageable dependency ratios enhances economic stability. As the number of household members increases, per capita consumption expenditure declines, indicating that larger families face greater financial pressure and potential vulnerability to poverty. Statistically, the MERT–Medium model achieves lower RMSE and MAE values than the LMM counterpart, indicating better predictive accuracy and reduced estimation error.

Beyond statistical performance, MERT provides stronger interpretability by revealing nonlinear and interaction effects such as how education amplifies the impact of asset ownership differently across gender groups that conventional linear models may overlook. From a policy perspective, these findings emphasize the importance of integrated poverty alleviation strategies that address both human capital development and asset-based empowerment. Expanding educational access particularly for women and promoting asset ownership (e.g., through microcredit, land certification, or vehicle financing programs) can significantly enhance household welfare. Moreover, gender-sensitive social protection programs should prioritize female-headed and asset-poor households, ensuring that poverty reduction efforts are both inclusive and sustainable.

3.4.4 MERT-Low Interpretation

The MERT–Low model, presented in Fig. 7, illustrates the decision rules that differentiate household groups with low household consumption expenditure in West Java. The decision tree consists of several components the root node, internal nodes, branches (splits), and leaf nodes each representing the model’s process of grouping households with similar socio-economic characteristics. The root node marks the most influential variable in explaining differences in household consumption patterns among lower-income groups. In this model, the gender of the household head (X_1) serves as the initial and most decisive variable.

This indicates that gender plays a crucial role in shaping household welfare at the lower end of the expenditure distribution. Female-headed households are predominantly concentrated in the low-consumption group, reflecting structural gender inequalities such as limited access to formal employment, restricted control over assets, and a greater dependence on informal or unstable income sources. This finding reinforces the notion that female-headed households are more vulnerable to poverty and should be prioritized in poverty alleviation programs.

Among female-headed households, education level (X_3) and welfare card ownership (X_7) further differentiate welfare conditions. Households in which the head has only attained junior or senior high school education and also holds a welfare card exhibit the lowest predicted consumption expenditure. This pattern suggests that limited educational attainment constrains income-generating capacity, while welfare card ownership signals reliance on government assistance serving as an indicator of poverty targeting. These results are consistent with the human capital theory, which emphasizes that education enhances skills, productivity, and access to better-paying jobs. Conversely, low education perpetuates economic vulnerability and dependence on social protection programs.

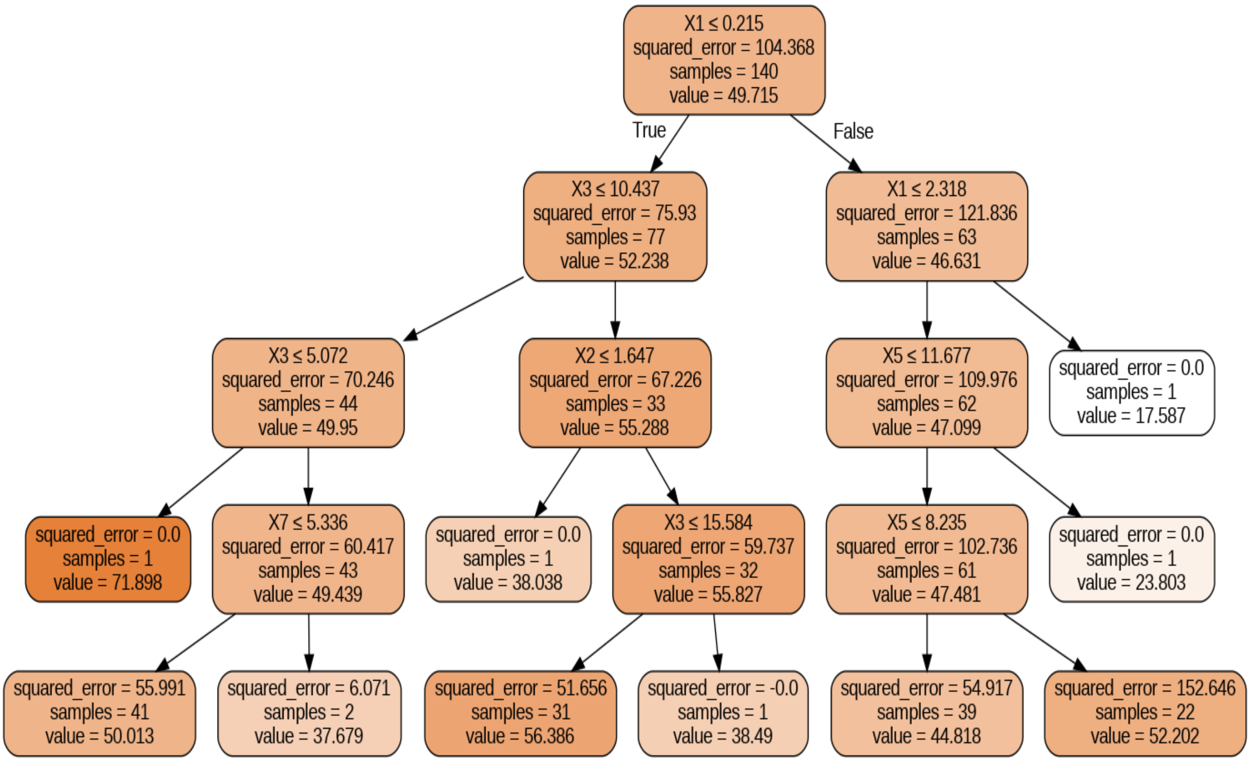


Figure 7. Decision Tree of MERT-*Low*

On the right side of the decision tree, male-headed households dominate. Within this group, the interaction between age of the household head (X_2) and car ownership (X_5) defines expenditure differences. Younger male household heads tend to have lower predicted consumption (around 17,587), which may reflect early career instability and lower income.

As household heads age and accumulate experience, their expenditure levels increase, particularly among those who own a car, suggesting higher purchasing power and greater access to productive opportunities. Car ownership, as a proxy for wealth, not only represents financial capacity but also facilitates income generation by improving mobility and access to employment or markets.

Model evaluation results presented in Table 3 further demonstrate the empirical strength of MERT compared to the Linear Mixed Model (LMM). The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) for the MERT–West Java model are lower than those for the LMM–West Java model, indicating that MERT provides more accurate predictions of household consumption expenditure. Beyond statistical accuracy, MERT also offers superior interpretability, as it captures nonlinear relationships and interactions among socio-economic factorsparticularly how education, gender, and asset ownership jointly shape poverty outcomes. In summary, the MERT–Low model highlights that education, gender, and asset ownership are the most critical determinants of household welfare among low-income groups. These findings suggest that efforts to reduce poverty should focus on: improving access to quality education, especially for women; strengthening asset ownership and productive capacity; and refining the targeting of social protection programs such as welfare cards to reach households most in need. Collectively, these interventions can enhance household resilience and support more inclusive economic growth in West Java.

The overall findings from the MERT–High, MERT–Medium, MERT–Low, and MERT–West Java models collectively reveal a consistent structural pattern in household consumption expenditure across socio-economic groups in West Java. Education, gender, and asset ownership emerge as the most decisive factors influencing household welfare, supported by the additional role of social protection indicators such as welfare card ownership. Across all model variants, the household head's education level stands out as the strongest determinant of expenditure, affirming the critical role of human capital in enhancing income capacity and reducing vulnerability to poverty.

Higher educational attainment is consistently associated with greater household consumption, reflecting improved access to stable employment and higher earning potential. The gender of the household head also serves as a significant differentiating factor, especially in the MERT–Low and MERT–Medium models. Female-headed households are systematically linked with lower predicted expenditure, indicating persistent gender disparities in economic opportunities, asset control, and access to productive resources. This pattern underscores the gendered nature of poverty and highlights the need for inclusive economic empowerment programs targeting women, particularly those leading households.

Asset ownership, particularly of land and cars, functions as a tangible indicator of economic stability and wealth accumulation. Ownership of these assets not only reflects material well-being but also enhances the household's ability to generate income, access financial services, and withstand economic shocks. In contrast, households lacking productive assets remain more susceptible to non-monetary forms of poverty, such as limited access to mobility and productive infrastructure. Moreover, the welfare card variable consistently captures the dimension of social protection targeting, indicating that households receiving welfare cards generally belong to the lower expenditure group.

While this demonstrates that government poverty alleviation programs are reaching the intended beneficiaries, it also emphasizes that such interventions primarily address short-term consumption rather than long-term income generation. Therefore, integrating education and asset development into social protection strategies is essential for sustained poverty reduction. From a methodological perspective, model evaluation using RMSE and MAE confirms that the MERT framework outperforms the Linear Mixed Model (LMM) across all household consumption expenditurecategories. The lower prediction errors in MERT indicate superior model accuracy, while its tree-based structure enhances interpretability, allowing clearer identification of key poverty determinants and their interactions. Thus, MERT provides not only a more robust statistical tool but also a more insightful analytical framework for understanding the multidimensional nature of poverty. In summary, the MERT analyses demonstrate that enhancing education, promoting gender equity, expanding asset ownership, and optimizing social protection targeting are central to improving household welfare in West Java.

Table 3. Model Evaluation of LMM and MERT

Model	**RMSE**	**MAE**
LMM-*West Java*	663,208.1060	480,599.9187
LMM-*High*	673,164.50	673,164.50
LMM-*Medium*	651,747.12	474,460.40
LMM-*Low*	595,015.03	411,191.46
MERT-*West Java*	631,940.4383	477,850.2285
MERT-*High*	711,891.8172	570,050.1674
MERT-*Medium*	624,978.4396	470,751.5320
MERT-*Low*	557,844.2004	407,822.8447

The model evaluation presented in Table 3 reports the computed RMSE and MAE values for each model. Based on these metrics, the MERT–West Java model exhibits lower RMSE and MAE values compared to the LMM–West Java model, indicating better predictive accuracy in modeling per-capita expenditure as an indicator of poverty. Therefore, MERT can be considered to provide a relatively more accurate fit for the data than LMM. However, the observed differences should be interpreted as indicative rather than conclusive, since no formal statistical test of model performance differences was conducted. These findings provide strong empirical evidence for policymakers to design integrated poverty alleviation strategies that move beyond short-term assistance toward building long-term economic resilience and social inclusion.

4. CONCLUSION

In conclusion, Linear Mixed Model (LMM) and Mixed Effect Regression Tree (MERT) provided valuable into the determinants of household consumption expenditure in West Java. However, the MERT model demonstrates superior predictive performance and interpretability, as reflected by its lowes RMSE and MAE values and its ability to capture nonlinear relationships and household heterogeneity. While LMM effectively estimates average effects, MERT offers a more detailed understanding of subgroup-specific patterns, making it particularly useful for identifying vulnerable household profiles. These findings highlight the importance of education, asset ownership, and targeted social protection in poverty alleviation policies. Overall, MERT represents a more adaptive and informative framework for designing evidence-based strategies to reduce household poverty.

Author Contributions

Nur Fitriyani Sahamony: Conceptualization, Methodology, Software, Validation, Writing-Original Draft. Asysta Amalia Pasaribu: Data curation, Resources, Draft Preparation, Writing-Review and Editing. Bagus Sartono: Formal Analysis, Validation, Khairil Anwar Notodiputro: Software, Visualization, **S**upervision. All authors discussed the results and contributed to the final manuscript.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgment

Authors express their sincere gratitude to the Statistics and Data Science Program of the School of Data Science, Mathematics, and Informatics at IPB University, Bina Nusantara University, and Binawan University for their valuable support in this work.

Declarations

The authors declare no competing interest.

Declaration of Generative AI and AI-assisted Technologies

AI-assisted technology was used to support sentence restructuring and clarity improvements. The authors confirm that the underlying ideas, arguments, data analyses, and conclusions are original and were not generated by AI. All AI-assisted edits were critically reviewed and validated by the authors.

REFERENCES

[1] Mouteyica AEN, Ngepah N. EXPLORING HEALTH OUTCOME DISPARITIES IN AFRICAN REGIONAL ECONOMICS COMMUNITIES: A MULTILEVEL LINEAR MIXED-EFFECT ANALYSIS. *BMC Public Health*. 2025;25(1):175.doi: https://doi.org/10.1186/s12889-025-21306-5

[2] Liu Q, al. e. DEVELOPMENT AND VALIDATION OF PREDICTION MODELS FOR INCIDENT REVERSIBLE COGNITIVE FRAILTY BASED ON SOCIAL-ECOLOGICAL PREDICTORS USING GENERALIZED LINEAR MIXED MODEL AND MACHINE LEARNING ALGORITHMS. *Journal Of Applied Gerontology*. 2025;44(2):255-66.doi: https://doi.org/10.1177/07334648241270052

[3] Mhelembe T, Ramroop S, Habyarimana F. DETERMINING THE RISK FACTORS OF MALARIA AND ANEMIA IN CHILDREN BETWEEN 6 AND 59 MONTHS USING THE JOINT GENERALIZED LINEAR MIXED MODEL ON THE 2021 NIGERIA MALARIA INDICATOR SURVEY DATASET. *Frontiers in Public Health*. 2025;12:doi: https://doi.org/10.3389/fpubh.2024.1503884

[4] Scandola M, Tidoni E. RELIABILITY AND FEASIBILITY OF LINEAR MIXED MODELS IN FULLY CROSSED EXPERIMENTAL DESIGNS. *Advances in Methods and Practices in Psychological Science*. 2024;7: doi: https://doi.org/10.1177/25152459231214454.

[5] Santi VM, Notodiputro KA, Indahwati I, Sartono B. RESTRICTED MAXIMUM LIKELIHOOD ESTIMATION FOR MULTIVARIATE LINEAR MIXED MODEL IN ANALYZING PISA DATA FOR INDONESIAN STUDENTS. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*. 2022;16(2):607-14.doi: https://doi.org/10.30598/barekengvol16iss2pp607-614

[6] Diz-Rosales N, Lombardía MJ, Morales D. POVERTY MAPPING UNDER AREA-LEVEL RANDOM REGRESSION COEFFICIENT POISSON MODELS. *Journal of Survey Statistics and Methodology*. 2024;12:404-34.doi: https://doi.org/10.1093/jssam/smad036

[7] Li H, Avendaño SM, Bak MYS. ANALYZING THE EFFECTS OF A REPEATED READING INTERVENTION ON READING FLUENCY WITH GENERALIZED LINEAR MIXED MODELS. *Evaluation & the Health Professions*. 2025;48(1):112-28.doi: https://doi.org/10.1177/01632787241257450

[8] Gutierrez S, Meza E, Glymour MM, Torres JM. MY PARENT, MYSELF, OR MY CHILD: WHOSE EDUCATION MATTERS MOST FOR TRAJECTORIES OF COGNITIVE AGING IN MIDDLE AGE. *American Journal of Epidemiology*. 2024;193(5):695-706.doi: https://doi.org/10.1093/aje/kwad108

[9] Perić M, Filipović S. FOREIGN DIRECT INVESTMENTS AND LABOUR FORCE INDICATORS IN TRANSITION ECONOMIES: LINEAR MIXED-EFFECTS MODELS IMPACT ANALYSIS. *Sociológia-Slovak Sociol*. 2021;53(3).doi: https://doi.org/10.31577/sociologia.2021.53.3.9

[10] Lee K. ASSOCIATIONS BETWEEN CUMULATIVE POVERTY AND CHILDREN'S ACADEMIC OUTCOMES: ABSOLUTE VERSUS RELATIVE POVERTY. *British Journal of Social Work*. 2025.doi: https://doi.org/10.1093/bjsw/bcaf001

[11] Eyasu AM, Zewotir T, Dessie ZG. IMPACT OF CROP COMMERCIALIZATION ON MULTIDIMENSIONAL POVERTY IN RURAL ETHIOPIA: PROPENSITY SCORE APPROACH. *Frontiers in Public Health*. 2025;12.doi: https://doi.org/10.3389/fpubh.2024.1412670.

[12] Deveci N, Uğurlu M, Uğurlu GK, Şahin EK, Erdoğan EÇ, Çayköylü A. THE EFFECTS OF PSYCHOLOGICAL FLEXIBILITY AND RESILIENCE ON PSYCHOPHARMACOLOGICAL TREATMENT RESPONSE IN PATIENTS WITH OBSESSIVE-COMPULSIVE DISORDER. *International Clinical Psychopharmacology*. 2024;39(3):181–6. doi: https://doi.org/10.1097/YIC.0000000000000499

[13] Jenkins MR, al. e. BEHAVIORAL AND PSYCHOLOGICAL SYMPTOMS OF DEMENTIA AND ALZHEIMER'S DISEASE PROGRESSION IN DOWN SYNDROME. *Journal of Neurodevelopmental Disorders*. 2025;17(1):19. doi: https://doi.org/10.1186/s11689-025-09604-w

[14] Nurfadilah K, Aidi MN, Notodiputro KA, Susetyo B. MULTILEVEL REGRESSIONS FOR MODELING MEAN SCORES OF NATIONAL EXAMINATIONS. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*. 2024;18(1):0323-32.doi: https://doi.org/10.30598/barekengvol18iss1pp0323-0332

[15] Rezapour M, al. e. ANALYZING GAIT DYNAMICS AND RECOVERY TRAJECTORY IN LOWER EXTREMITY FRACTURES USING LINEAR MIXED MODELS AND GAIT ANALYSIS VARIABLES. *Bioengineering*. 2025;12(1):67.doi: https://doi.org/10.3390/bioengineering12010067

[16] Smout S, Champion K, O'Dean S, Teesson M, Gardner L, Newton N. ANXIETY, DEPRESSION AND DISTRESS OUTCOMES FROM THE HEALTH4LIFE INTERVENTION FOR ADOLESCENT MENTAL HEALTH: A CLUSTER-RANDOMIZED CONTROLLED TRIAL. *Nature Mental Health*. 2024;2(7):818–27.doi: https://doi.org/10.1038/s44220-024-00246-w

[17] Busanga JK, Njuho P. LINEAR MIXED MODEL EFFECTS IN META ANALYSIS OF AGRICULTURAL DATA. *International Journal of Agricultural and Statistical Sciences*. 2024;20(2).doi: https://doi.org/10.59467/IJASS.2024.20.511

[18] Caviedes J, Ibarra JT, Calvet-Mir L, Álvarez-Fernández S, Junqueira AB. INDIGENOUS AND LOCAL KNOWLEDGE ON SOCIAL-ECOLOGICAL CHANGES IS POSITIVELY ASSOCIATED WITH LIVELIHOOD RESILIENCE IN A GLOBALLY IMPORTANT AGRICULTURAL HERITAGE SYSTEM. *Agricultural Systems*. 2024;216. doi: https://doi.org/10.1016/j.agsy.2024.103885.

[19] Osorio-Gallardo T, Bijma P. COMPARISON OF LINEAR MIXED MODELS FOR GENETIC FEATHER SCORE ANALYSIS IN LAYING HENS KEPT IN RECURRENT TESTING FACILITIES. *Poultry Science*. 2025:104833.doi: https://doi.org/10.1016/j.psj.2025.104833

[20] Brown H, Prescott R. APPLIED MIXED MODELS IN MEDICINE.: John Wiley & Sons; 2015.doi: https://doi.org/10.1002/9781118778210

[21] A. Setiawan. HUKUM PERTANAHAN: (PENGATURAN, PROBLEMATIKA DAN REFORMASI AGRARIA): LaksBang *Justitia;* 2019.