

Simultaneous inference for empirical best predictors in generalized linear mixed models: A poverty study in West Java



NUR FITRIYANI SAHAMONY^{1,2*}, KUSMAN SADIK¹, ANANG KURNIA¹

¹Statistics and Data Science Study Program, School of Data Science, Mathematics and Informatics, IPB University, Indonesia

²Digital Business Study Program, Faculty of Business and Social Science, Binawan University, Jakarta, Indonesia

*Corresponding Author:
nurfitriyanisahamony@gmail.com

Received: 01 July 2025
Revised: 22 September 2025
Accepted: 26 October 2025

Abstract. Accurate poverty mapping at the district and municipal levels remains challenging due to small sample sizes in household surveys, which often result in unstable direct estimates. To address this issue, this study employs microdata from the 2023 National Socioeconomic Survey (SUSENAS) to estimate household-level poverty proportions across 27 districts and municipalities in West Java Province using a binomial Generalized Linear Mixed Model (GLMM) combined with the Empirical Best Predictor (EBP) and Simultaneous Confidence Intervals (SCI). The GLMM framework captures household characteristics and random area effects to account for spatial heterogeneity. Three SCI approaches—Bonferroni correction, Bootstrap-t, and the Simes procedure—were implemented to evaluate EBP uncertainty while controlling the family-wise error rate. Results reveal substantial disparities, with Tasikmalaya (21.7%), Bandung Barat (15.5%), and Cianjur (12.8%) consistently above the provincial average of (6.8%), while urban areas such as Cimahi, Bekasi, and Depok report poverty rates below 2%. All methods achieved full empirical coverage (ECP = 100%), although interval widths differed: Bonferroni produced the widest intervals (AIW = 44.99), Bootstrap-t yielded the narrowest and most efficient (AIW = 29.16), and Simes provided intermediate but highly consistent results (AIW = 33.24). These findings underscore the methodological importance of integrating GLMM, EBP, and SCI for small area estimation while offering practical insights for evidence-based policy development and poverty reduction strategies in Indonesia.

Keywords: Empirical best predictor, generalized linear mixed models, poverty estimation, small area estimation, simultaneous confidence intervals

INTRODUCTION

Poverty remains a pressing global issue, with more than 700 million people worldwide still living below the international poverty line [1]. In Indonesia, although the overall trend shows a decline, the national poverty rate in March 2022 is recorded at approximately 9.54% [2]. This challenge becomes increasingly complex due to regional disparities, which often cause the effectiveness of national programs to be inconsistently reflected at the local level.

Poverty is generally understood not only as income deprivation but also as a multidimensional phenomenon encompassing education, health, and access to basic services. The World Bank and the United Nations Development Programme (UNDP) highlight the multidimensional nature of poverty, which underscores the importance of reliable measurement at both national and subnational levels.

West Java Province, as the most populous region in Indonesia, exemplifies this paradoxical condition. Rapid urbanization and economic growth occur, yet pockets of absolute poverty persist, particularly in rural areas such as Garut District [3]. As of March 2023, the poverty rate in this province reaches 7.62%, equivalent to approximately 3.89 million people, with the poverty line set at IDR 495,229 per capita per month [4]. Although this figure is lower than the national average of 8.57%, several districts such as Bandung Barat and Subang report poverty rates above the provincial mean, at 10.52% and 9.52%, respectively [3]. These facts underscore the existence of significant disparities that require more precise statistical mapping.

In this study, households are classified as poor if their monthly per capita expenditure falls below the 2023 provincial poverty line of West Java, IDR 495,229 [4]. This threshold serves as the basis for categorizing the binary response variable (1 = poor, 0 = non-poor) in



This is an open access article under the terms of the Creative Commons Attribution (CC-BY) License, which permits use and distribution in any medium, provided the original work is properly cited. © 2025 The Authors. Jurnal Natural published by Faculty of Mathematics and Natural Science (FMIPA), Universitas Syiah Kuala.

statistical modeling using Generalized Linear Mixed Models (GLMM).

A common challenge in estimating poverty rates at the district and city level is the small sample size in surveys such as the National Socioeconomic Survey (*Survei Sosial Ekonomi Nasional*, SUSENAS). This limitation frequently results in unstable estimates when using direct estimation approaches. Consequently, Small Area Estimation (SAE) techniques are widely adopted to produce more reliable estimates for small domains. One of the most popular estimators within the SAE framework is the Empirical Best Predictor (EBP), particularly when combined with GLMM [5-7].

However, conventional EBP approaches typically yield confidence intervals independently for each area, without accounting for simultaneous uncertainty across all domains. In decision-making contexts such as aid allocation or intervention prioritization, it is necessary to apply inferential procedures that consider collective uncertainty. Simultaneous Confidence Intervals (SCI) provide a methodological solution that ensures aggregate coverage across regions while reducing the risk of misclassification due to multiple testing [7].

SCI methods such as the Bonferroni correction [8], the Simes procedure, and the bootstrap-t approach [9] each have specific strengths and limitations in terms of precision and conservatism. Nevertheless, to date, there is limited research systematically comparing these three approaches within GLMM frameworks for poverty

estimation, particularly in the Indonesian context. Recent studies such as [7, 9] have advanced simultaneous inference methods within small area estimation frameworks, while [10] introduced new approaches for constructing simultaneous confidence intervals in generalized linear mixed models. More recently, [11, 12] extended SAE applications to modern machine learning frameworks and labor force indicators. Nevertheless, the application of these methods to poverty studies remains scarce in Southeast Asia, particularly in Indonesia, which underscores the novelty of this research.

Therefore, this study aims to implement binomial GLMM to estimate poverty using EBP, construct and compare SCI using the Bonferroni, Simes, and bootstrap-t approaches, and evaluate the performance of these SCI methods based on metrics including Average Interval Width (AIW), Empirical Coverage Probability (ECP), and the classification accuracy of poverty areas. Through a case study in West Java Province, this research is expected to contribute methodological insights into the development of EBP-based simultaneous inference while improving the spatial and statistical accuracy of poverty targeting programs.

METHODOLOGY

This study utilizes raw microdata from the 2023 National Socioeconomic Survey (SUSENAS), published by Statistics Indonesia (*Badan Pusat Statistik*). The dataset includes information on household characteristics across all districts and municipalities in West Java Province.

Table 1. Description of Research Variables

Variables	Description	Type	Effects
Y	Household Poverty Status	1: Poor 0: Non-Poor	Response Variable
X_1	Gender of Household Head	1: Male 0: Female	Fixed Effect
X_2	Age of Household Head	1: 16–29 years 2: 30–39 years 3: 40–49 years 4: 50–59 years 5: 60–69 years 6: 70+ years	Fixed Effect
X_3	Highest Educational Attainment	1: Higher Education 2: Senior High School 3: Junior High School 4: Elementary School 5: No Formal Education	Fixed Effect
X_4	Household Size	1: 1 person 2: 2 persons 3: 3 persons 4: 4 persons 5: 5 persons 6: More than 6 persons	Fixed Effect
X_5	Car Ownership	1: Owns a Car 0: Does Not Own a Car	Fixed Effect
X_6	Land Ownership	1: Owns Land 0: Does Not Own Land	Fixed Effect
X_7	Ownership of Welfare Card	1: Holds a Welfare Card 0: Does Not Hold a Welfare Card	Fixed Effect
U_j	District/Municipality	–	Random Effect

The unit of analysis is the household, with thousands of observations collected from various administrative areas. This study incorporates household-level demographic and socioeconomic characteristics such as education, household size, and asset ownership, which are well-documented determinants of poverty in Indonesia.

Research Variables

The response variable (Y_i) is dichotomous, defined as follows:

$$Y_i = \begin{cases} 1, & \text{if monthly per capita expenditure} < \text{IDR } 495,229 \text{ (poor)} \\ 0, & \text{if monthly per capita expenditure} \geq \text{IDR } 495,229 \text{ (non - poor)} \end{cases}$$

This classification is based on the provincial poverty line of West Java as of March 2023, determined by Statistics Indonesia [4]. The predictor variables are encoded categorically and numerically according to their characteristics, as summarized in the Table 1.

Analytical Method

The analytical methods employed in this study consist of both descriptive and inferential techniques. Descriptive analysis was conducted to provide a general overview of the spatial distribution of poverty across districts and cities in West Java Province. Inferential analysis was carried out using a Generalized Linear Mixed Model (GLMM) with a binomial distribution and logit link function, incorporating random effects at the district/city level. This was followed by estimation using the Empirical Best Predictor (EBP) approach and Simultaneous Confidence Interval (SCI)-based inference.

Binomial GLMM Specification

The model used is a two-level hierarchical model, where the first level comprises household-level characteristics, and the second level captures the cluster structure of districts/cities. The general form of the binomial GLMM is expressed as:

$$\text{Logit}(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = \beta_0 + \sum_{k=1}^p \beta_k X_{kij} + u_{0j}$$

Information:

π_{ij} : probability that household i in area j is classified as poor

β_0 : fixed intercept

X_{kij} : household-level predictor variables [e.g., gender of household head, age, education, etc.)

β_k : fixed effect coefficients

$u_{0j} \sim \mathcal{N}(0, \sigma_u^2)$ area-level random effects for district/city j

p : number of fixed-effect covariates included in the model (intercept excluded).

The model parameters were estimated using the Maximum Likelihood method via the `glmer()` function from the `lme4` package in R. The response variable was modeled using a binomial distribution with a logit link, while district/city random effects were assumed to follow a normal distribution with mean zero and unknown variance. Fixed effects were chosen based on theoretical relevance to poverty studies (e.g., gender, education, household size) and statistical significance. Competing model specifications were compared using Akaike Information Criterion

(AIC), Bayesian Information Criterion (BIC), and likelihood ratio tests to ensure optimal fit.

Empirical Best Predictor (EBP)

After model estimation, the Empirical Best Predictor (EBP) approach was used to estimate poverty proportions for each district/city. EBP leverages the hierarchical structure of the GLMM to produce stable and efficient predictions, particularly for areas with small sample sizes [5].

Simultaneous Confidence Intervals for EBP under GLMM

To account for multiple comparisons across small areas and ensure simultaneous inference for Empirical Best Predictors (EBP), we constructed Simultaneous Confidence Intervals (SCI) under the Generalized Linear Mixed Model (GLMM) framework. Three methods were implemented: Bonferroni correction, bootstrap-t approximation, and Simes procedure. These methods aim to control the family-wise error rate (FWER) at a nominal level $1 - \alpha$, enhancing the reliability of small area estimates [7, 13-15].

Bonferroni-Based SCI

Following the classical approach [16], the Bonferroni-adjusted SCI for each domain $d=1, \dots, D$ is defined as:

$$\mathcal{I}_{d,1-\alpha}^{BO} = [\hat{\mu}_d \pm z^{1-\frac{\alpha}{2D}} \cdot \hat{\delta}(\hat{\mu}_d)]$$

where $\hat{\mu}_d$ is the EBP, $\hat{\delta}(\hat{\mu}_d)$ its standard error, and $z^{1-\frac{\alpha}{2D}}$ is the Bonferroni-corrected critical value from the standard normal distribution. This method is computationally simple but tends to be conservative, especially when domains are correlated [17].

Bootstrap-t SCI

The bootstrap-t method (14) approximates the sampling distribution of the maximum standardized statistic. For each bootstrap replicate $b=1, \dots, B$, we compute:

$$S_{Bd}^{*(b)} = \frac{\hat{\mu}_d^{*(b)} - \hat{\mu}_d^{*(b)}}{\hat{\delta}(\hat{\mu}_d^{*(b)})}, \quad S_B^{*(b)} = \max_{d=1, \dots, D} |S_{Bd}^{*(b)}|$$

Then, the $[1 - \alpha]$ -quantile of the empirical distribution of $S_B^{*(b)}$ is used as the critical value $c^{S_0}(1 - \alpha)$, leading to:

$$\mathcal{I}_{d,1-\alpha}^{BT} = [\hat{\mu}_d \pm c^{S_0}(1 - \alpha) \cdot \hat{\delta}(\hat{\mu}_d)]$$

This method is particularly effective when the normal approximation is invalid, or variance heterogeneity is present [18]. The bootstrap-t procedure was implemented with 1000 replications, with resampling stratified at the household level within each district/municipality to preserve intra-area dependence. This replication size was chosen to ensure stable estimation of interval widths and empirical coverage while maintaining computational feasibility.

Simes-Based SCI

The Simes procedure [8] adjusts p-values for multiple testing by ordering individual p-values $p^{(1)} \leq \dots \leq p^{(D)}$ and comparing each to the threshold $\frac{i\alpha}{D}$. A domain d is considered significant if its $p_d \leq \frac{i\alpha}{D}$. SCI bounds are constructed by identifying domains with adjusted significance and estimating intervals for the

accepted ones. This method controls the FWER under positive dependence and is more powerful than Bonferroni under mild correlation [19].

Evaluation Metrics for Simultaneous Confidence Intervals

To assess the performance of Simultaneous Confidence Intervals (SCI) constructed under the Generalized Linear Mixed Model (GLMM), a series of quantitative evaluation metrics were applied. These include Empirical Coverage Probability (ECP), Average Interval Width (AIW), Average Interval Width Variability (AIWV), and Family-Wise Error Rate (FWER). These indicators are standard in the literature for evaluating accuracy, efficiency, and stability of SCI methods [7, 14, 18].

Empirical Coverage Probability (ECP)

ECP measures the proportion of domain-level confidence intervals that successfully capture the true parameter value. It evaluates how well the SCI performs in achieving its nominal confidence level (e.g., 95%).

$$ECP = \frac{1}{D} \sum_{d=1}^D I(\theta_d \in J_{d,1-\alpha})$$

information

θ_d : true value of the parameter in domain d

$J_{d,1-\alpha}$: the $[1 - \alpha]$ level confidence interval for domain d

$I(\cdot)$: indicator function that equals 1 if the condition is true, 0 otherwise

A well-calibrated SCI method should produce an ECP close to the target level.

Average Interval Width (AIW)

AIW is used to assess the efficiency of the SCI. It measures the average length of the intervals across all domains. Narrower intervals generally indicate higher estimation precision.

$$AIW = \frac{1}{D} \sum_{d=1}^D (U_d - L_d)$$

Information

U_d, L_d upper and lower bounds of the SCI for domain d

Average Interval Width Variability (AIWV)

AIWV measures the variability in interval widths across all domains. High variability may indicate inconsistency in the precision of the estimates, while low AIWV reflects stability.

$$AIWV = Var[U_d - L_d]$$

Lower AIWV is generally preferred, as it indicates consistent uncertainty quantification across domains.

Family-Wise Error Rate (FWER)

FWER is the probability of committing at least one Type I error (i.e., at least one true value falling outside its confidence interval). This global error rate is central in multiple inference settings, especially for small area estimation.

$$FWER = P(\exists d \in 1, \dots, D: \theta_d \notin I_{d,1-\alpha})$$

A reliable SCI procedure should maintain FWER at or below the nominal level α . Bonferroni is known to

tightly control FWER, while bootstrap and Simes provide more balanced trade-offs under certain correlation structures [16, 19].

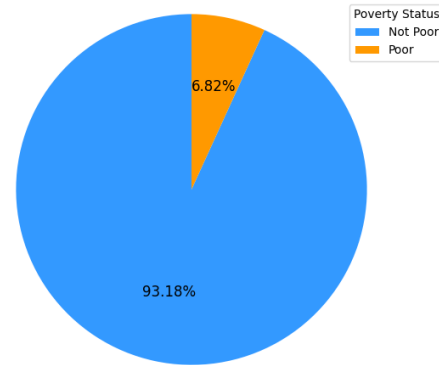


Figure 1. Household Poverty Status in West Java Province

RESULTS AND DISCUSSION

Descriptive Analysis: Household Poverty Status

Prior to statistical modelling, households are classified based on their monthly per capita expenditure. A household is considered poor if its expenditure falls below IDR 495,229, in accordance with the poverty line of West Java Province as of March 2023, as established by Statistics Indonesia [4]. Based on this classification, approximately 6.82% of households in the sample are categorized as poor, while 93.18% are classified as non-poor. This distribution is presented in Figure 1.

Although this overall figure is relatively low, further grouping by district and municipality reveals substantial spatial variation.

Figure 2 shows that several areas exhibit poverty concentrations well above the provincial average. Tasikmalaya District records the highest proportion at 22.04%, followed by Bandung Barat (15.69%) and Cianjur (12.99%). Conversely, Cimahi City (0.51%), Bekasi City (1.34%), and Depok City (1.35%) report significantly lower poverty rates compared to the provincial mean. This variation highlights that poverty reduction policies cannot be uniformly applied across all regions. Priority mapping based on the identification of poverty clusters is necessary. These findings are consistent with prior studies [20] which emphasize the importance of spatial targeting strategies in social protection programs. Furthermore, [21] notes that provincial averages often mask extreme internal disparities at the district and municipal levels. Therefore, advanced multilevel approaches such as Generalized Linear Mixed Models (GLMM) are essential to statistically capture inter-regional heterogeneity.

Additionally, spatial visualization is conducted by grouping regions based on the mean and standard deviation of household poverty proportions. Figure 3 displays the classification into three categories: high (proportion $> \mu + \sigma$), medium ($\mu - \sigma \leq \text{proportion} \leq \mu + \sigma$), and low (proportion $\leq \mu - \sigma$). Most districts and cities

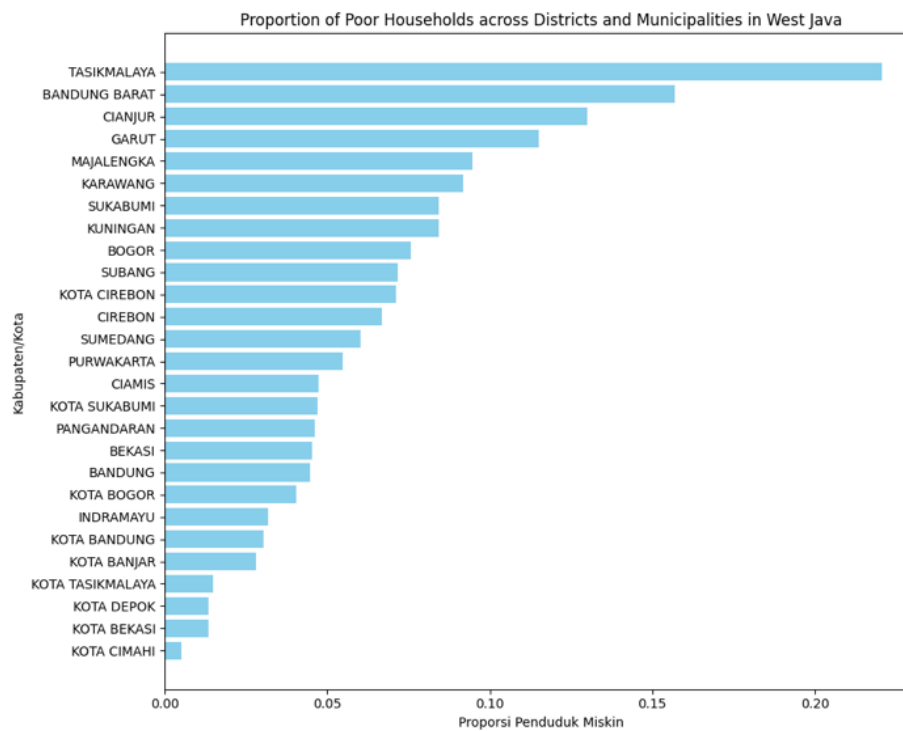


Figure 2. Proportion of Poor Households by District and Municipality
Spatial Distribution of Poor Household Proportions in West Java (Categorized by Mean $\pm$ SD)

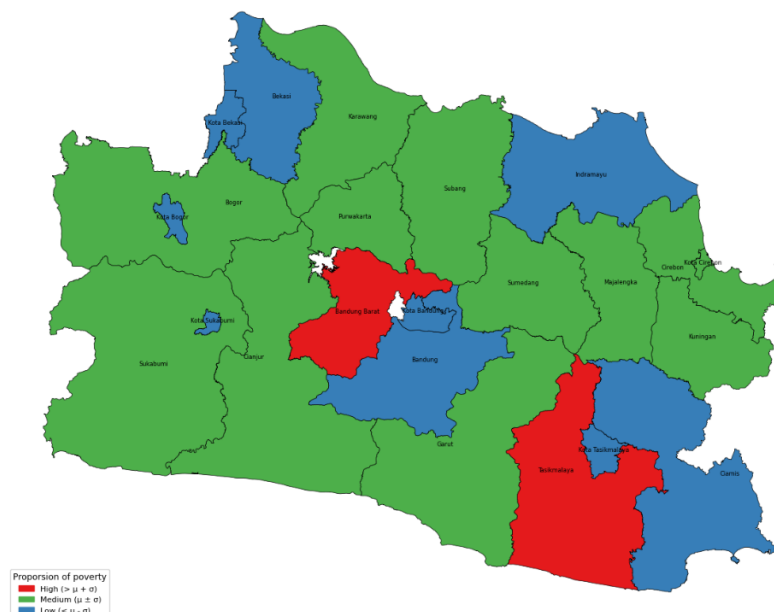


Figure 3. Spatial Distribution Map of Household Poverty Proportions in West Java

fall into the medium category. Areas in the high category, including Tasikmalaya, Cianjur, and Bandung Barat, reflect greater socioeconomic pressure. In contrast, Bekasi City and Tasikmalaya City are classified as low, potentially indicating more effective local policies or more stable economic conditions.

Classification based on the mean $\pm$ standard deviation provides an objective statistical approach for identifying areas with extreme poverty conditions. This method is crucial for supporting evidence-based policymaking, particularly in prioritizing social development interventions and allocating resources more equitably and effectively

Results of Binomial GLMM Modeling

A Generalized Linear Mixed Model (GLMM) with a logit link function is employed to analyze the effects of household characteristics on the probability of poverty status in West Java Province. Parameter estimation is conducted using the maximum likelihood approach through the `glmer()` function from the *lme4* package [22]. The complete results of coefficient estimation are presented in Table 2.

Most predictor variables show significant effects on household poverty status. The gender of the household

Table 2. Estimated Parameters of the Binomial GLMM

Variables	Estimate	Std. Error	z value	Pr(> z)	Significance
[Intercept]	-6.16590	0.22684	-27.182	< 2e-16	***
X1	-0.31525	0.08829	-3.571	0.000356	***
X2	0.01329	0.02380	0.558	0.576551	
X3	0.42115	0.02986	14.104	< 2e-16	***
X4	0.69039	0.02250	30.679	< 2e-16	***
X5	-2.78211	0.35674	-7.799	6.26e-15	***
X6	-0.59030	0.05642	-10.463	< 2e-16	***
X7	0.13699	0.06091	2.249	0.024505	*
random effect					
Group Name	Variance	Standard Deviation			
District/City	0,585	0,7648			
[Intercept]					

*) Significant at $\alpha = 5\%$

Table 3. Model fit indices for the binomial GLMM

Index	Value
AIC	10574.0
BIC	10647.5
Marginal R ²	0.334
Conditional R ²	0.434
ICC	0.151
RMSE	0.237
Log-loss	0.202

head (X₁) has a significant negative effect ($\beta = -0.315$; $p < 0.001$), indicating that households headed by males are less likely to be poor compared to those headed by females, consistent with previous findings [23] on the relationship between gender inequality and economic vulnerability.

The age of the household head (X₂) has a positive but non-significant coefficient ($\beta = +0.013$; $p = 0.576$), suggesting that, after controlling for other variables, age does not have a direct impact on poverty status. This aligns with similar results reported in multilevel studies on energy poverty [24].

Educational attainment (X₃) exhibits a significant positive effect ($\beta = +0.421$; $p < 0.001$), indicating that lower levels of education are associated with higher probabilities of poverty. This finding corroborates prior research [25, 26] emphasizing the strong correlation between limited education and low household welfare. Household size (X₄) also shows a significant positive effect ($\beta = +0.690$; $p < 0.001$), implying that households with more members tend to face higher consumption burdens and an increased risk of poverty, a pattern similarly observed by [27] in mapping vulnerable households in West Nusa Tenggara Province.

Car ownership (X₅) and land ownership (X₆) each have significant negative effects ($\beta = -2.782$ and $\beta = -0.590$;

$p < 0.001$), suggesting that households possessing assets such as vehicles and land demonstrate better economic resilience, as noted by [28, 29].

Ownership of a welfare card (X₇) has a significant positive effect ($\beta = +0.137$; $p = 0.024$), confirming that recipient households are more likely to belong to targeted low-income groups, as highlighted in evaluative studies [30]. The random effect of district or municipality is also significant, with an estimated variance of 0.585 and a standard deviation of 0.765. This indicates the presence of spatial heterogeneity that cannot be explained solely by individual household factors, reinforcing the importance of employing multilevel mixed models. The influence of area-level random effects is essential in the context of small area estimation, consistent with assertions by [31, 32] that GLMMs are effective in capturing systematic inter-area variation in welfare indicators. Overall, these results demonstrate that household poverty status is determined by a combination of demographic, social, and asset ownership factors, along with spatial effects captured through the random intercept structure.

To evaluate the adequacy of the model specification, we further examined several fit indices. The final model converged successfully and provided lower AIC and BIC compared to simpler alternatives. Table 3 presents the diagnostic measures, including AIC, BIC, R², ICC, RMSE, and log-loss.

The final model converged successfully and demonstrated adequate fit (Table 3). The model yielded AIC = 10574.0 and BIC = 10647.5. Fixed effects explained 33.4% of the variance (marginal R² = 0.334), while the inclusion of district/city random effects increased the explained variance to 43.4% (conditional R² = 0.434). The intra-class correlation (ICC = 0.151) indicated that 15.1% of the variation in poverty status was attributable to differences between districts, confirming the relevance of the hierarchical specification. Additional performance indices (RMSE =

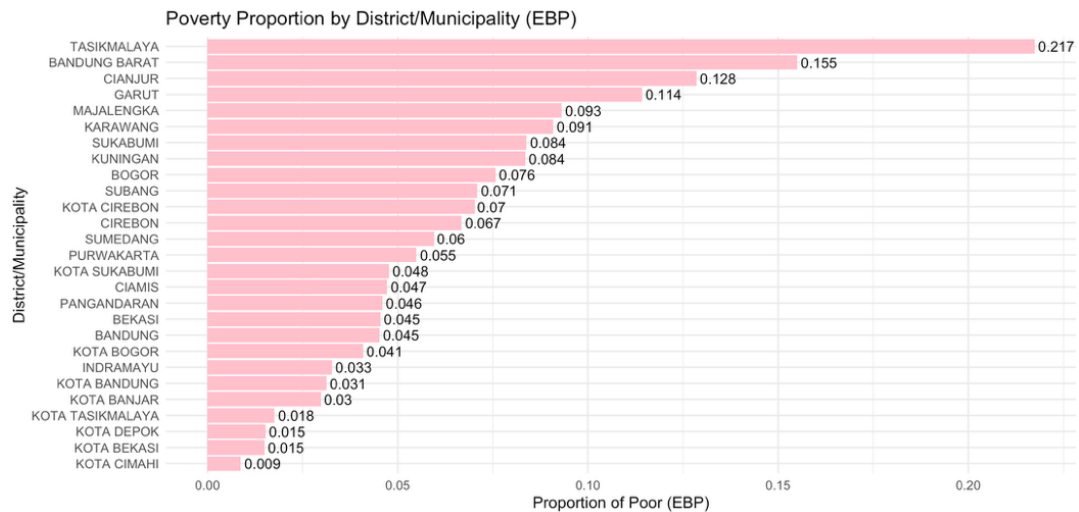


Figure 4. Poverty Proportions by District/City Based on Empirical Best Predictor (EBP)

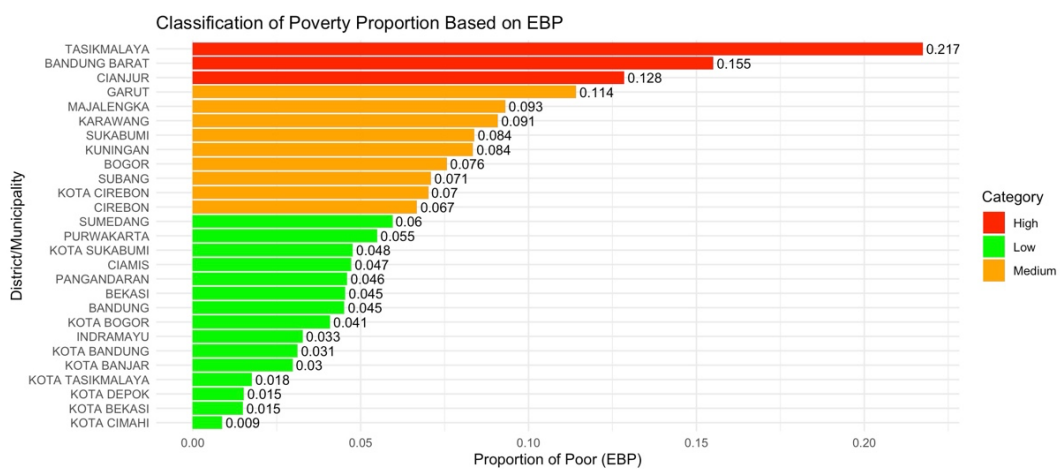


Figure 5. Classification of Poverty Proportions Based on EBP Estimates

0.237, log-loss = 0.202) also supported the adequacy of the model.

Poverty Proportion Estimation Using the Empirical Best Predictor (EBP)

The results of the binomial logit GLMM estimation are used to compute the probability that each household falls below the poverty threshold in each district and municipality, employing the Empirical Best Predictor (EBP) approach. The EBP leverages the hierarchical structure of the GLMM to produce more stable predictions, particularly in areas with small sample sizes. Unlike direct estimation methods, EBP simultaneously accounts for household-level characteristics and area-level random effects, resulting in estimates that are more informative and reliable.

Figure 4 presents the predicted poverty proportions (EBP estimates) for all 27 districts and cities in West Java Province. The findings reveal substantial spatial disparities. Tasikmalaya District records the highest poverty proportion at 21.7%, followed by Bandung Barat (15.5%) and Cianjur (12.8%). In contrast, urban areas such as Cimahi City (0.9%), Bekasi City (1.5%), and Depok City (1.5%) show considerably lower poverty levels. This gap emphasizes that the provincial average

of 6.82% does not reflect conditions in specific areas. Differences exceeding a twenty-fold range between districts and cities reinforce the argument that poverty reduction policies based solely on average indicators risk neglecting poverty hotspots. This observation is consistent with previous studies [21] which show that average indicators often obscure internal inequalities, and with [20] which highlights the importance of spatial approaches in targeting social protection programs.

In this context, EBP estimates serve as a critical instrument for designing equitable, evidence-based policymaking. As noted by [5], the EBP is highly effective for small area estimation, especially when sample sizes vary across regions. These authors demonstrate that the EBP significantly reduces estimation variance compared to direct methods, improving the efficiency and reliability of results. Moreover [31] emphasize that the EBP is particularly well-suited to capturing spatial heterogeneity and measuring poverty at the local level.

This information not only provides a more accurate spatial profile but also forms the basis for classifying poverty areas according to the provincial average threshold or supporting further inference using

SCI Comparison: Bootstrap-t, Bonferroni, and Simes

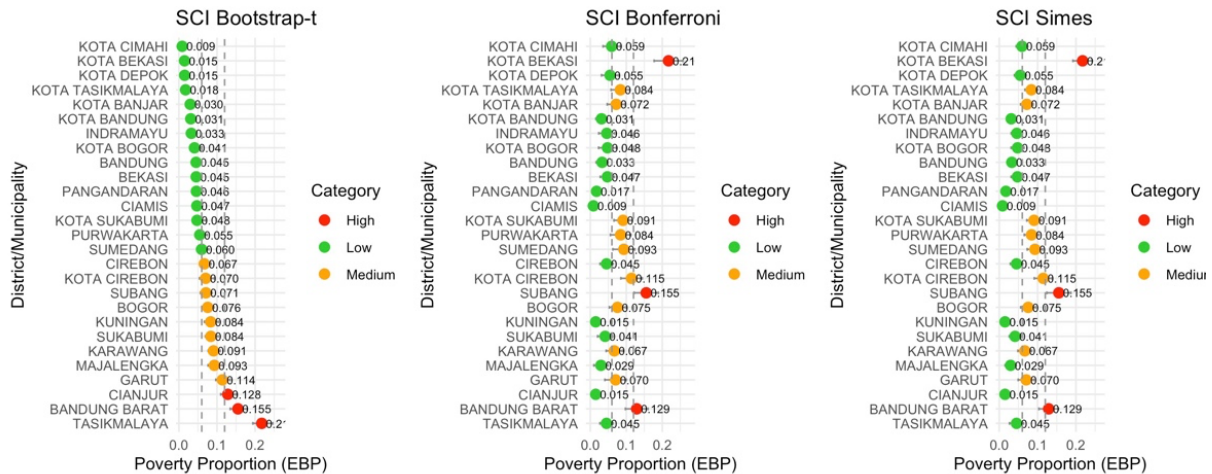


Figure 6. Comparison of Simultaneous Confidence Intervals Using Bonferroni, Simes, and Bootstrap-t Methods.

Simultaneous Confidence Intervals (SCI), which are discussed in the following sections.

Classification of Poverty Proportions Based on EBP Point Estimates

As an initial step in identifying priority areas for poverty alleviation, poverty proportions are classified according to the EBP point estimates for each district and city. The classification uses fixed thresholds defined as follows:

- High: $EBP \geq 0.12$
- Medium: $0.06 \leq EBP < 0.12$
- Low: $EBP < 0.06$

This classification applies directly to the EBP point estimates derived from the GLMM, without considering estimation variance or uncertainty. The purpose of this approach is to offer a practical and easily interpretable mapping framework, particularly in policy contexts that rely on predetermined thresholds.

Figure 5 presents the classification results, with colors indicating the three categories: red for high, orange for medium, and green for low poverty areas.

The results show that:

- High category areas include Tasikmalaya District (21.7%), Bandung Barat (15.5%), and Cianjur (12.8%), reflecting acute poverty clusters that should be prioritized for intervention.
 - Medium category areas comprise districts and cities such as Majalengka, Garut, Karawang, and Cirebon City. While not extreme, these areas still warrant attention in poverty reduction strategies.
 - Low category areas are primarily urban centers such as Cimahi City, Bekasi City, and Depok City, which benefit from urbanization and more developed socioeconomic infrastructure.
- This EBP-based classification facilitates more responsive poverty reduction strategies by segmenting areas according to measurable poverty risk levels. This approach aligns with recommendations by [20, 21], who advocate using

explicit spatial indicators in designing social assistance programs.

Simultaneous Inference of Poverty Proportions Using Simultaneous Confidence Intervals (SCI)

As a further step in analyzing poverty proportions based on EBP, simultaneous inference is performed using Simultaneous Confidence Intervals (SCI) to identify districts and municipalities where poverty proportions are statistically significantly higher than the provincial average. In this study, the provincial threshold is calculated as the mean EBP proportion across all areas, amounting to 6.82%. SCI are constructed using three statistical methods: Bonferroni Correction, Bootstrap-t, and the Simes Procedure.

Figure 6 presents a comparison of the three SCI approaches, with error bars showing EBP point estimates alongside their lower and upper confidence bounds for each area. Each method demonstrates distinct strengths and limitations in terms of interval width, detection sensitivity, and statistical conservatism.

The Bonferroni approach is the most conservative, as it tightly controls the family-wise error rate (FWER), producing the widest intervals. Consequently, only three areas—Bandung Barat, Subang, and Kota Bekasi—are identified as significantly above the provincial poverty threshold. This characteristic is consistent with findings by [16, 17], who emphasize that Bonferroni's strict error control often comes at the cost of under-detecting true signals.

In contrast, the Simes procedure generates shorter, more efficient intervals by applying a step-up adjustment for FWER control [8]. Similar to Bonferroni, it identifies Bandung Barat, Subang, and Kota Bekasi as significant, but with narrower intervals that enhance detection power. However, as noted by [19], this increased sensitivity may raise the risk of false positives under dependency structures.

The Bootstrap-t method represents a compromise between conservatism and efficiency. By resampling at

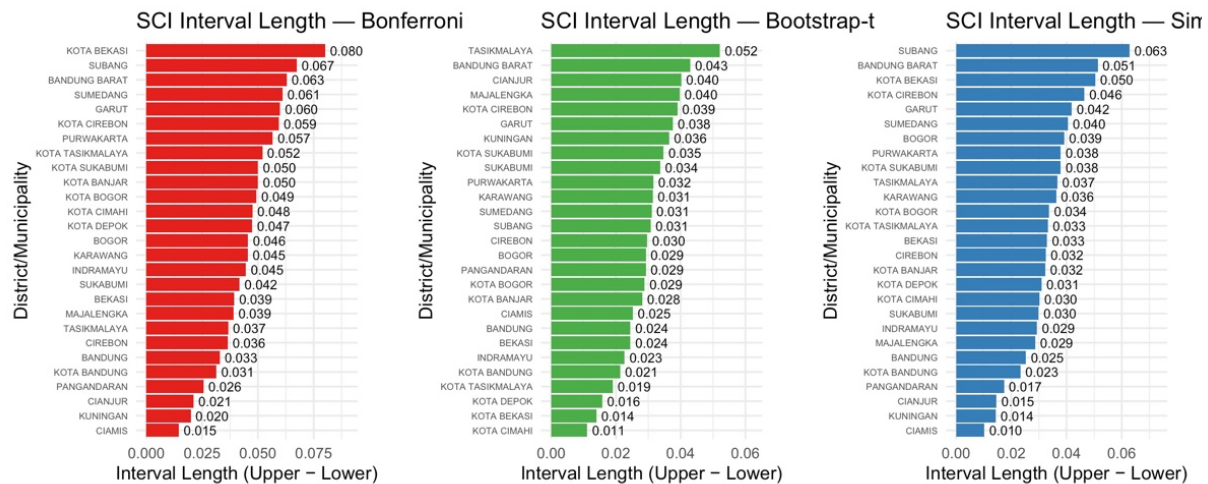


Figure 7. Comparison of SCI Interval Lengths Across Three Methods

the household level within districts, it adapts to empirical distributional features and is particularly suitable for hierarchical models [14, 33]. Using this approach, Tasikmalaya, Bandung Barat, and Cianjur were detected as significantly high-poverty areas, highlighting its ability to capture rural poverty clusters that might be overlooked by more conservative procedures.

Overall, Bandung Barat consistently appears across all three methods, confirming its robustness as a poverty hotspot. The divergence in identified areas underscores the methodological trade-offs: Bonferroni prioritizes strict error control, Simes emphasizes detection power, and Bootstrap-t balances both. These findings reinforce recent arguments by [7, 9] that simultaneous inference procedures are essential for reliable decision-making in small area poverty estimation. From a policy perspective, Bonferroni may be preferable when minimizing misclassification is paramount, Simes when broader coverage is desired, and Bootstrap-t when a balanced strategy is required.

Comparison of SCI Interval Lengths Across Methods

Figure 7 illustrates the comparison of SCI interval lengths across all 27 districts and cities in West Java Province, based on the Bonferroni, Bootstrap-t, and Simes approaches. Intervals are calculated using EBP estimates derived from the binomial GLMM.

The visualization shows that Bonferroni consistently produces the longest intervals across nearly all areas, with the maximum length reaching 0.073 in Tasikmalaya. This reflects the method's conservative nature, which favors wider intervals to guarantee 100% simultaneous coverage [16].

Figure 7 presents the comparison of interval lengths generated by the Bonferroni, Bootstrap-t, and Simes procedures across districts in West Java. The results indicate clear differences in efficiency and variability among the three approaches. The Bonferroni method consistently produced the widest intervals, with the maximum length observed in Kota Bekasi (0.080), followed by Subang (0.067) and Bandung Barat (0.063). This pattern reflects the conservative nature of

Bonferroni, which prioritizes strict control of the family-wise error rate but at the expense of estimation precision, as previously noted by [16, 17]. In contrast, the Bootstrap-t procedure yielded moderate and relatively stable interval lengths, with Tasikmalaya (0.052), Bandung Barat (0.043), and Cianjur (0.040) recording the longest values. These results confirm the adaptability of the bootstrap framework in handling complex data structures, supporting earlier findings by [9, 14] regarding its robustness in small area estimation. Meanwhile, the Simes procedure produced the narrowest intervals, with a maximum width of 0.063 in Subang and a minimum of 0.010 in Ciamis. Although this efficiency highlights the power of the Simes method, its tendency toward shorter intervals may increase the risk of undercoverage when spatial dependence exists [8, 19]. Overall, these findings illustrate the methodological trade-offs among SCI approaches: Bonferroni ensures conservative error control, Simes provides highly efficient yet less conservative inference, and Bootstrap-t offers a balanced compromise between stability and efficiency. This evidence aligns with [7], who emphasize that the choice of SCI method substantially shapes both the statistical accuracy and the policy relevance of small area estimates.

Classification of Poverty Proportions and SCI Results

Figures 8–10 display the classification of poverty proportions based on Empirical Best Predictor (EBP) estimates alongside the simultaneous inference results using Bootstrap-t (Figure 8), Simes (Figure 9), and Bonferroni (Figure 10). The EBP classification provides an initial ranking of districts, whereas the SCI procedures offer statistically validated identification of areas with poverty levels significantly above the provincial

Under the Bootstrap-t procedure, Tasikmalaya (21.7%), Bandung Barat (15.5%), and Cianjur (12.8%) were identified as significantly high-poverty areas. These results are in line with official *Badan Pusat statistics* (March 2023), which report Tasikmalaya at 12.13% and Bandung Barat at 10.52%, both above the provincial average of 7.62%. The detection of Cianjur, although not

EBP (left) and SCI Bootstrap-t (right)

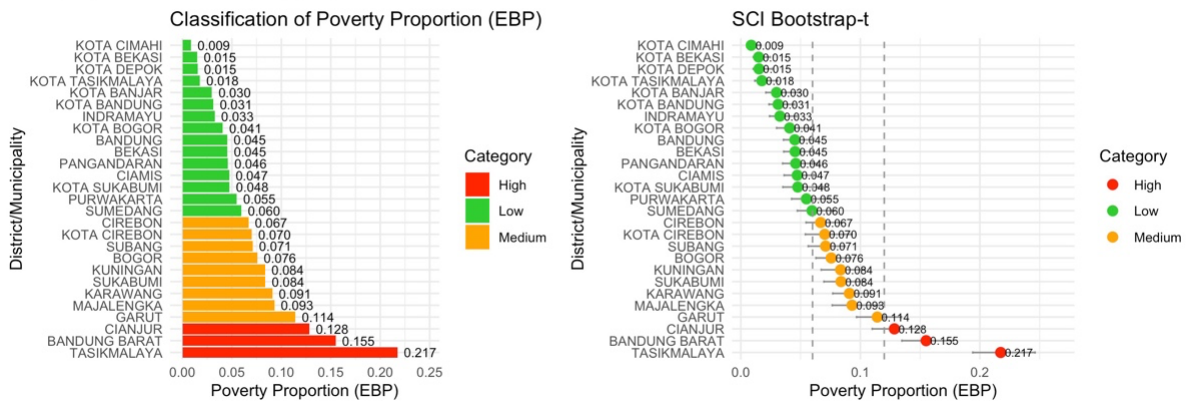


Figure 8. Classification of Poverty Proportion (EBP) and SCI Bootstrap

EBP (left) and SCI Simes (right)

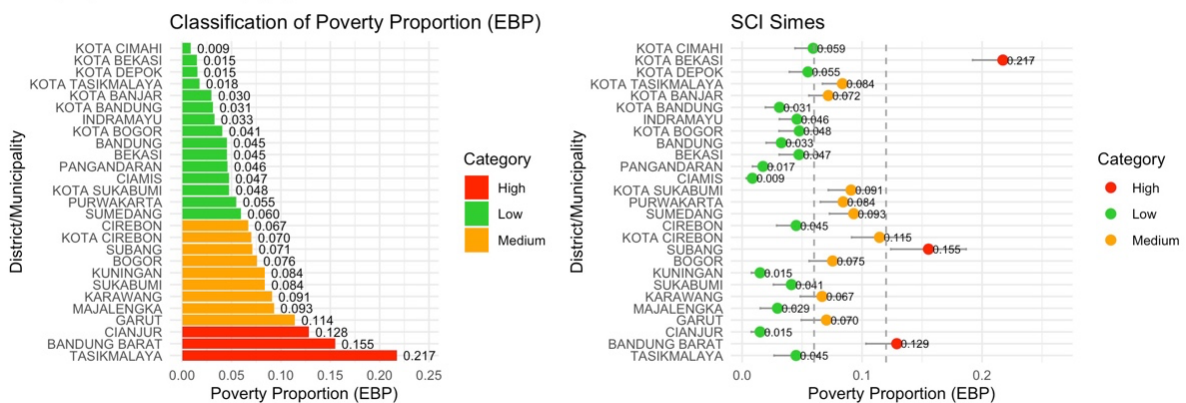


Figure 9. Classification of Poverty Proportion (EBP) and SCI Simes

EBP (left) and SCI Bonferroni (right)

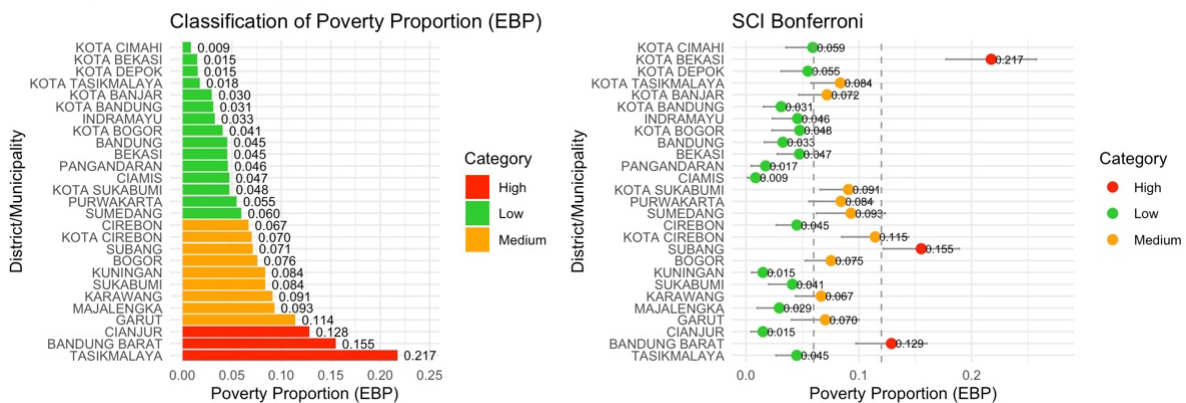


Figure 10. Classification of Poverty Proportion (EBP) and SCI Bonferroni

Table 4 Evaluation Summary of SCI Methods

Methods	ECP (%)	AIW (permil)	AIWV (permil ²)	FWER
Bonferroni	100	44.99	15.26	1.000
Bootstrap-t	100	29.16	9.210	0.657
Simes	100	33.24	11.64	0.741

highlighted in BPS publications, reflects the sensitivity of Bootstrap-t in capturing rural heterogeneity and small-sample domains.

The Simes procedure identified Tasikmalaya, Bandung Barat, and Subang (9.5%) as significantly above the

provincial threshold. The inclusion of Tasikmalaya (BPS: 12.13%) and Bandung Barat (BPS: 10.52%) is consistent with official statistics, while Subang (BPS: 9.52%) also lies above the provincial mean, reinforcing the detection. However, the tendency of Simes to classify

additional borderline districts illustrates its higher detection power but also its potential vulnerability to false positives in urban areas.

The Bonferroni procedure produced conservative intervals and consistently classified Tasikmalaya and Bandung Barat as significantly high-poverty areas, confirming their status as hotspots both in model-based estimates (EBP: 21.7% and 15.5%) and in BPS data (12.13% and 10.52%). However, it failed to identify Cianjur and Subang, despite their relatively elevated poverty rates (EBP: 12.8% and 9.5%; BPS: 9.17% and 9.52%). This under-detection underscores Bonferroni's strict control of the family-wise error rate.

Overall, Tasikmalaya and Bandung Barat were consistently identified as significant across all three SCI methods, supported by both EBP estimates and BPS official statistics. The differences across methods highlight the methodological trade-offs: Bonferroni offers strict error control but risks overlooking critical areas, Simes enhances detection power but may over-classify, and Bootstrap-t provides the most balanced compromise by aligning closely with BPS statistics while remaining sensitive to rural heterogeneity.

Evaluation of SCI Performance: ECP, AIW, and AIWV

Table 4 summarizes the evaluation of SCI methods based on three criteria: empirical coverage probability (ECP), average interval width (AIW), and interval width variability (AIWV).

All three methods achieved satisfactory coverage levels, each recording an Empirical Coverage Probability (ECP) of 100%. This indicates that all intervals consistently captured the true parameter values across all simulation replicates, demonstrating excellent coverage performance.

In terms of interval efficiency, the Bonferroni approach produced the widest intervals, as reflected by the highest Average Interval Width (AIW) of 44.99%, suggesting a highly conservative estimation strategy. However, this conservatism comes at the expense of interpretability and statistical precision. In contrast, the Bootstrap-t method yielded the narrowest intervals (AIW = 29.16%), indicating a more efficient and precise estimation. The Simes procedure exhibited intermediate performance, with an AIW of 33.24%, providing a balance between coverage and efficiency.

Regarding consistency, the Bootstrap-t method also demonstrated the lowest Average Interval Width Variance (AIWV = 9.210%²), reflecting more stable interval lengths across all areas. This consistency is particularly relevant for applications requiring uniform uncertainty quantification. Conversely, the Bonferroni method had the highest AIWV (15.26%²), further emphasizing its tendency toward greater variability and less parsimonious intervals.

In terms of error control, the Bonferroni procedure achieved the strictest Family-Wise Error Rate (FWER = 1.000), highlighting its strong control over the probability of making at least one Type I error.

Meanwhile, Bootstrap-t and Simes demonstrated lower FWERs of (0.657 and 0.741), respectively, indicating that these methods allow more flexibility in balancing error control with interval precision.

These findings align with the results reported by [14, 17], who underscored the inherent trade-offs between interval width, coverage probability, and error control in small area estimation frameworks [14, 17]. While Bonferroni remains a robust option when strict error control is prioritized, the Bootstrap-t and Simes methods provide attractive alternatives for generating narrower, more consistent intervals, particularly in settings involving Generalized Linear Mixed Models (GLMM) with spatial dependence among domains. More efficient and consistent intervals when applying GLMM-based small area estimation.

CONCLUSION

This study demonstrates the implementation of binomial GLMM combined with EBP and three SCI approaches to estimate poverty proportions across districts and municipalities in West Java Province using SUSENAS 2023 data. The results confirm substantial spatial heterogeneity, with Tasikmalaya, Bandung Barat, and Cianjur consistently classified as high-poverty areas, while urban centers such as Cimahi, Bekasi, and Depok fall into the low-poverty category.

All SCI methods achieved full empirical coverage, yet their performance differed: Bonferroni provided the most conservative intervals, Bootstrap-t offered the most efficient and stable inference, and Simes delivered compact intervals with higher detection power. Among household characteristics, education, household size, and asset ownership emerged as significant predictors of poverty, reinforcing their role as key determinants in policy design.

The methodological contribution of this study lies in integrating GLMM, EBP, and SCI to improve the reliability of small area poverty estimation. Practically, the findings provide evidence-based insights for policymakers to design targeted poverty alleviation strategies. Future research could extend this work by incorporating spatial correlation structures or Bayesian hierarchical models to further enhance inference accuracy in complex socioeconomic contexts.

ACKNOWLEDGMENT

The authors gratefully acknowledge the Central Bureau of Statistics of the Republic of Indonesia for granting access to the SUSENAS 2023 microdata used in this study. The authors also express their sincere appreciation to the reviewers and editorial team of Jurnal Natural for their constructive feedback that has contributed to improving the manuscript. This research was funded solely by the authors without external financial support. All findings, opinions, and conclusions presented herein are entirely the responsibility of the authors and do not necessarily reflect the views of their affiliated institutions.

REFERENCE

- [1]. World Bank. "Poverty and Inequality Platform (PIP)," *World Bank*. Available: <https://www.worldbank.org> (accessed Jun. 30, 2025).
- [2]. Badan Pusat Statistik (BPS). 2022. *Persentase Penduduk Miskin Maret 2022*. (Jakarta: BPS-Statistics Indonesia).
- [3]. BPS Provinsi Jawa Barat. 2022. *Profil Kemiskinan Provinsi Jawa Barat*. (Bandung: BPS Provinsi Jawa Barat).
- [4]. BPS Provinsi Jawa Barat. "Kemiskinan Maret 2023 Provinsi Jawa Barat." *BPS Provinsi Jawa Barat*. Available: <https://jabar.bps.go.id> (accessed Jun. 30, 2025).
- [5]. Rao J. N. K.; Molina I. 2015. *Small area estimation*. 2nd ed. (Hoboken, NJ: Wiley).
- [6]. Molina, I.; Rao, J. N. K. 2010. Small area estimation of poverty indicators. *Can. J. Stat.* **38**(3) 369–385. DOI:<https://www.jstor.org/stable/27896031>.
- [7]. Reluga, K.; Lombardía, M.-J.; Sperlich, S. 2023. Simultaneous inference for empirical best predictors with a poverty study in small areas. *J. Am. Stat. Assoc.* **118**(541) 583–595. DOI: <https://doi.org/10.1080/01621459.2021.1942014>
- [8]. Simes R. J. 1986. An improved Bonferroni procedure for multiple tests of significance. *Biometrika*. **73**(3) 751–754. DOI: <https://doi.org/10.1093/biomet/73.3.751>.
- [9]. Zhang L.; Hall P.; Maiti T. 2022. Bootstrap t procedures for simultaneous confidence intervals for complex indicators in small area estimation. *Stat. Probab. Lett.* **188**(1) 109446.
- [10]. Wang Y.; et al. 2022. Constructing statistical intervals for small area estimates under GLMM. *BMC Med. Res. Methodol.* **22** 178.
- [11]. Frink N.; Schmid T. 2024. Small area estimation with generalized random forests: Estimating poverty rates in Mexico. *arXiv preprint arXiv:2406.03861*. Available: <https://arxiv.org/abs/2406.03861> (accessed Jun. 24, 2025).
- [12]. Bugallo M.; et al. 2025. Small area estimation of labour force indicators under unit-level models. *J. R. Stat. Soc. A (Stat. Soc.)*. **188**(1) 241–262.
- [13]. Butar F. B.; Lahiri P. 2003. On measures of uncertainty of empirical Bayes small-area estimators. *J. Stat. Plan. Inference*. **112**(1–2) 63–76.
- [14]. Hall P.; Maiti T. 2006. On parametric bootstrap methods for small area prediction. *J R Stat Soc Series B Stat Methodol.* **68**(2) 221–238.
- [15]. Reluga; Katarzyna. 2020. Simultaneous and post-selection inference for mixed parameters. Doctoral Thesis. DOI: 10.13097/archive-ouverte/unige:138615.
- [16]. Romano J. P.; Wolf M. 2005. Exact and approximate stepdown methods for multiple hypothesis testing. *J. Am. Stat. Assoc.* **100**(469) 94–108.
- [17]. Hobza T.; Novák M.; Pustejovská P. 2018. The performance of the Bonferroni and step-down procedures under spatial dependence. *Revstat Stat. J.* **16**(1) 53–76.
- [18]. Chatterjee S.; Lahiri P.; Li H. 2008. Parametric bootstrap approximation to the distribution of EBLUP and related prediction intervals in linear mixed models. *Ann. Stat.* **36**(3) 1221–1245.
- [19]. Benjamini Y.; Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**(4) 1165–1188.
- [20]. Pratiwi S.; Setianingrum L. 2022. Pola spasial dan tren kemiskinan di Indonesia tahun 2015–2020. *J. Perenc. Pengemb. Kebijak.* **2**(3) 220–229.
- [21]. Suharyo W.; Akhmadi; Hastuti; Filaili R.; Budiati S.; Munawar W. 2005. *Developing a poverty map for Indonesia: A tool for better targeting in poverty reduction and social protection programs – Book 4: Field verification*. (Jakarta: The SMERU Research Institute).
- [22]. Bates D.; Mächler M.; Bolker B.; Walker S. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**(1) 1–48.
- [23]. Muhammad F.; Muchtar M.; Sihombing P. R. 2024. The nexus of gender inequality and poverty rate in Indonesia. *Perspektif: Perspektif: J. Ekon. Manaj. Univ. Bina Sarana Inform.* **22**(1).
- [24]. Hasanujjaman M.; Omar M. A. 2022. Household and non-household factors influencing multidimensional energy poverty in Bangladesh: Demographics, urbanization and regional differentiation via a multilevel modeling approach. *Energy Res. Soc. Sci.* **92** 102083. DOI: <https://doi.org/10.1016/j.erss.2022.102083>.
- [25]. Pangeran M.; Arham M. A.; Dai S. I. 2022. The effect of education and health performance on poverty in Indonesia. *Eur. J. Res. Dev. Sustain.* **3**(6) 119–124. <https://www.scholarzest.com>.
- [26]. Soseco T. 2021. Household size, education, and household wealth in Indonesia: Evidence from quantile regression. *J. Ekon. Indones.* **10**(3) 281–297. DOI: <https://doi.org/10.52813/jei.v10i3.72>.
- [27]. Wahyuni W.; Gazali M.; Hidayaturohman U. 2022. Pengelompokan dan pemetaan karakteristik kemiskinan di Provinsi Nusa Tenggara Barat menggunakan self organizing map (SOM) dan biplot. *Syntax Literate*. **7**(11) 15587–15605.
- [28]. Bryant G.; Spies-Butcher B.; Stebbing A. 2024. Comparing asset-based welfare capitalism: Wealth inequality, housing finance and household risk. *Hous. Stud.* **39**(2). DOI: <https://doi.org/10.1080/02673037.2022.2056150>.
- [29]. Fauziyah E.; Awang S. A.; Suryanto P.; Achmade B. 2025. Inequality and poverty of privately owned forests farmers in rural areas of Indonesia. *For. Sci. Technol.* **21**(1) 1–14. DOI: <https://doi.org/10.1080/21580103.2024.2409219>.
- [30]. Fikri A. T.; Afifudin; Sekarsari R. W. 2023. Evaluasi implementasi program bantuan sosial dalam menekan angka pengemis di Kota Malang (studi kasus di Dinas Sosial Kota Malang). *Respon Publik*. **17**(9) 11–21. <https://jim.unisma.ac.id/index.php/rpp/article/view/21914>.

- [31]. Tzavidis N.; Zhang L.-C.; Luna-Hernández Á.; Schmid T.; Rojas-Perilla N. 2018. From start to finish: A framework for the production of small area official statistics. *J. R. Stat. Soc. A (Stat. Soc.)* **181**(4) 927–979. DOI: <https://doi.org/10.1111/rssa.12364>.
- [32]. Hall P.; Maiti T.; Parnell K. 2011. Confidence regions for empirical best linear unbiased predictors in small area estimation. *The Annals of Statistics*. **39**(6) 2862–2882.
- [33]. Efron B.; Tibshirani R. J. 1994. *Ann. Stat.* (New York: Chapman & Hall/CRC).

How to cite this article:

Sahamony, N.F.; Sadik, K.; Kurnia., A. 2025. Simultaneous inference for empirical best predictors in generalized linear mixed models: A poverty study in West Java. *Jurnal Natural*, 25(3), 168-180.