DOI: https://doi.org/10.34123/jurnalasks.v17i2.851

Available online at https://jurnal.stis.ac.id/

Jurnal Aplikasi Statistika & Komputasi Statistik
Vol. 17 No. 2, December 2025
e- ISSN: 2615-1367, p-ISSN:2086-4132

Implementation of Twofold HB Beta SAE Model to Estimate Out-of-School Children with Disabilities in Indonesia

Aisha Maharani[1*], Azka Ubaidillah[2], Adhi Kurniawan[3]

[1,2]Politeknik Statistika STIS, Jakarta, Indonesia, [3]King's College London, London, United Kingdom
**Corresponding Author: E-mail address: aishaaranii167@gmail.com*

ARTICLE INFO

Article history:
Received 3 October, 2025
Revised 10 December, 2025
Accepted 23 December, 2025
Published 31 December, 2025

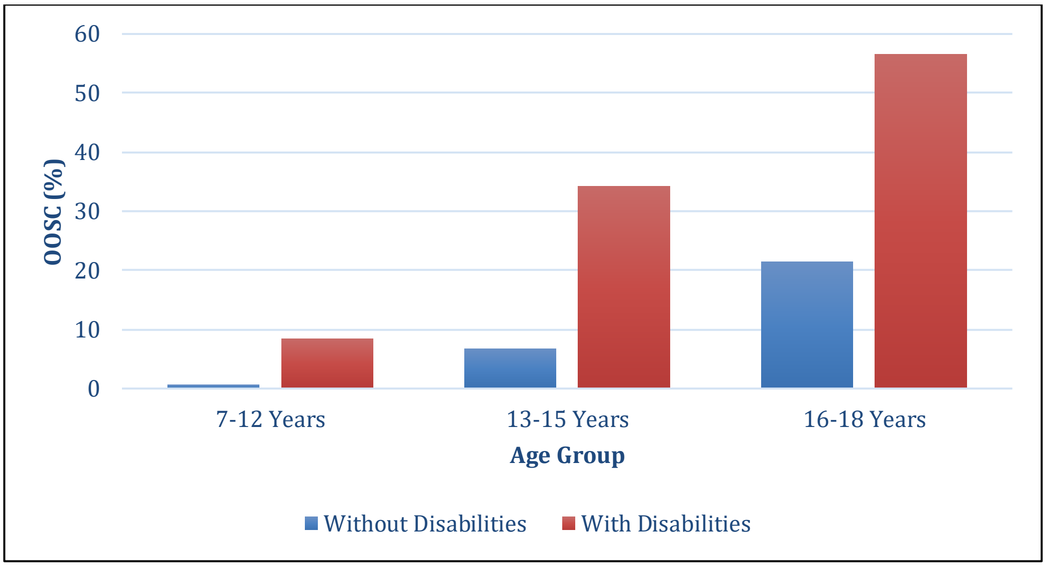
Keywords:
HB Beta; Out-of-School Children; Persons with Disabilities; SAE; Twofold Subarea-Level

Abstract

Introduction/Main Objectives: The high percentage of out-of-school children with disabilities in Indonesia reveals a significant gap in educational participation. **Background Problems:** Due to the absence of disability-focused surveys, accurate data are only available at the national level, which is insufficient to represent regional conditions. **Novelty:** With the increasing demand for small area data, this study estimates the percentage of out-of-school children with disabilities at the provincial and district levels simultaneously, using small area estimation (SAE). **Research Methods:** This study applies SAE using a twofold subarea-level model with a Hierarchical Bayes (HB) beta approach, covering all 34 provinces and 514 districts/cities in Indonesia. This model was developed using data from the National Socio-Economic Survey (Susenas) and the Village Potential Statistics (Podes). **Finding/Results:** The twofold HB beta SAE model achieves higher precision than direct estimation, as shown by lower relative standard errors (RSE) across regions. Furthermore, spatial patterns indicate that the percentage of out-of-school children with disabilities is mostly between 35.36% and 45.34%, with clusters concentrated in Kalimantan and Papua.

1. Introduction

The rapid advancement of technology and information compels societies worldwide to adapt and innovate in response to continuous transformations [1]. Enhancing the quality of human resources is one of the key ways to address this dynamic global transformation [2]. This can be pursued, among other ways, through education [3]. Education is positioned as a primary focus within the Sustainable Development Goals (SDGs), specifically Goal 4, “Quality Education”, which underscores the importance of inclusive and equitable education for all [4]. However, existing empirical evidence indicates that achieving this target remains challenging. UNESCO reported an increase of six million out-of-school children (OOSC) globally, bringing the total to 250 million in 2023 [5]. This figure accounts for 16 percent of the global child population [6]. Within this population of out-of-school children, children with disabilities constitute a particularly vulnerable group. They face greater barriers to accessing education due to physical, social, and institutional challenges. According to UNICEF, issues related to disability are not a new concern and have been a growing area of focus for more than a decade. In 2021, UNICEF released the most comprehensive global statistical analysis on children with disabilities, revealing that approximately 240 million children worldwide live with disabilities. The report highlights significant educational disparities, showing that children with disabilities are 49 percent more likely to have never attended school compared to their peers without disabilities [7]. This evidence underscores that despite global commitments to inclusive and equitable education under SDG 4, the specific educational needs of children with disabilities remain insufficiently addressed.



Source: Education Statistics 2023 (BPS)

Figure 1. The percentage of out-of-school children by age group and characteristics

In the Indonesian context, horizontal inequality in school participation between children with and without disabilities remains a substantial issue. School participation, as an indicator of educational success, can be observed through the percentage of out-of-school children [8]. Out-of-school children refer to school-age children, generally aged 7-18 years, who have never attended school or who dropped out before completing their education [9]. Many children within the 7–18 age group in Indonesia still experience limited access to both formal and informal education [10]. Figure 1 shows the persistent gap in school participation between children with and without disabilities across all age groups in Indonesia. The percentage of out-of-school children is consistently higher among those with disabilities, indicating that school participation among children with disabilities remains significantly lower at every age level.

Persons with disabilities, as members of society, should be entitled to the same rights as others, including the right to education. Ensuring access to education for persons with disabilities is not only a constitutional mandate but also a crucial step toward achieving the fourth SDGs, which emphasizes quality, inclusive, and equitable education for all [11]. The Government of Indonesia has demonstrated a strong commitment through various policies and programs, including equivalency education, educational affirmative action programs, "Wajib Belajar 12 Tahun" program, and the implementation of inclusive and special education designed for learners with disabilities. These initiatives are expected to expand access and reduce educational disparities for persons with disabilities.

Thus, the availability of data on persons with disabilities is crucial for tracking progress toward the SDGs and other development goals. The United Nations also requires every country to collect accurate data on persons with disabilities [12]. Such data can be used to assess whether the rights of persons with disabilities, particularly the right to education, have been fulfilled. In addition, the World Health Organization (WHO) has urged countries to gather disability-related data in order to monitor their living conditions and anticipate potential gaps or inequalities [13].

In Indonesia, the availability of disability-related data is still highly restricted [14]. Data is needed not only to identify existing gaps but also as a basis for formulating evidence-based policies to meet the needs of persons with disabilities and address the inequalities they experience [15]. Data availability also enables monitoring of implemented policies so that their effectiveness in addressing existing issues can be evaluated. BPS Statistics Indonesia as the official data provider, does not conduct a survey specifically dedicated to persons with disabilities. However, disability data are collected by BPS through National Socio-Economic Survey (Susenas), which also includes information on indicators such as the percentages of out-of-school children with disabilities. However, BPS only publishes this indicator at the national level. Consequently, these figures are insufficient to accurately reflect the educational conditions of persons with disabilities across different regions. Each region naturally has its own unique conditions and challenges. Therefore, it is necessary to conduct independent estimation in order to measure the rate of out-of-school children at smaller geographical levels. However, relying solely on Susenas data may lead to estimates with low precision due to insufficient sample sizes [16].

The problem of insufficient sample sizes can be addressed through indirect estimation methods, commonly known as Small Area Estimation (SAE). SAE is expected to enhance the effectiveness of available sample sizes, thereby producing estimates with higher precision without the need for additional sampling [17]. Generally, estimation using SAE is applied at only one area level. However, the demand for small-area data continues to increase, requiring information down to the smallest administrative units, such as districts. In this context, a model capable of producing estimates at multiple small-area levels simultaneously is required. Within SAE, such a model exists, namely the twofold subarea-level model. This model is an extension of the area-level Fay-Herriot model, which allows estimation at two small-area levels simultaneously, namely the subarea and area levels [17]. The twofold subarea-level model has been applied in several previous studies, such as Erciulescu et al. [18], Saadi & Ubaidillah [19], and Yudasena [20]. Other studies have employed the SAE method to estimate OOSC, including Eliezer et al. [21] and Trihandika et al. [22].

The main contribution of this study is to provide a more accurate indicator of OOSC among children with disabilities, which has so far been available only at the national level. Availability of data at the smallest administrative levels can support evidence-based policymaking and serve as a tool to monitor the effectiveness of government programs. Using the twofold HB beta SAE model also facilitates the estimation process, as it allows simultaneous estimation at two administrative levels. This approach helps reduce the time, effort, and costs typically required for separate estimation procedures. Furthermore, this study broadens the perspective on education issues for children with disabilities. Most previous research tends to focus on school dropout rates, whereas the OOSC indicator offers a more comprehensive view as it encompasses both children who have dropped out and those who have never attended school [23]. Therefore, this study aims to provide an overview of both direct and indirect estimates of out-of-school children with disabilities at the provincial and district/city levels in Indonesia, determine the best estimation method, and map the results.

2. Material and Methods

2.1. *Persons with Disabilities*

The concept of disability is highly complex and continues to evolve. The WHO has developed the International Classification of Functioning, Disability, and Health (ICF) framework as a reference for understanding and measuring health and disability [24]. This approach integrates various models of disability and acknowledges the role of environmental and personal factors in identifying disability status, as well as the relevance and impact of associated health conditions [25]. All components of disability and their interactions are illustrated in Figure 2.

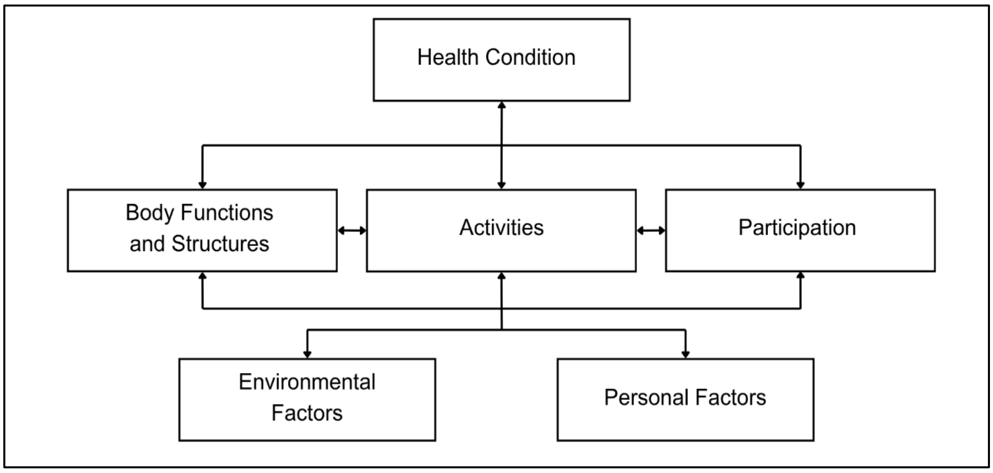


Figure 2. The ICF framework

Based on Figure 2, the ICF integrates health, individual, and social aspects, explaining the causal relationships between impairments and disabilities within a multidimensional model of functioning and health, including environmental and personal factors [26]. The concept of persons with disabilities according to BPS is aligned with the ICF framework. The emphasis in this concept is on long-term functional impairments or limitations that lead to restricted participation in society. Functional impairments or limitations are characterized by an inability, loss, or abnormality in psychological, physiological, or anatomical structures or functions [23].

2.2. Out-of-School Children

Globally, there is no universally agreed definition of out-of-school children [27]. As part of a global initiative, the UNESCO Institute for Statistics and UNICEF define out-of-school children as school-aged children who are not enrolled in any educational setting, whether formal, nonformal, or informal [28]. This group includes children who have never attended school, those of school age who have not yet enrolled, and those who have attended school but dropped out for some reason [29]. BPS defines out-of-school children in the same way as international guidelines, namely school-aged children who are not enrolled in any educational unit [23]. BPS publishes the out-of-school children indicator to facilitate the identification of policy or program interventions specifically targeted at children who are not attending school. The calculation of out-of-school children with disabilities is carried out using the following formula.

$$OOSC\ with\ disabilities_{7-18\ years} = \frac{\begin{array}{c}The\ number\ of\ children\ with\ disabilities\\ aged\ 7-18\ years\ who\ are\ out-of-school\\ during\ a\ specific\ period\end{array}}{\begin{array}{c}The\ number\ of\ children\ with\ disabilities\\ aged\ 7-18\ years\ during\ the\ same\ period\end{array}} \tag{1}$$

2.3. Direct Estimation

According to Rao & Molina, direct estimation is a technique for estimating population parameters based on sample from the relevant area or domain [17]. here are two approaches to direct estimation: model-based and design-based. Generally, direct estimation uses a design-based approach, which involves estimating parameters such as means, totals, or proportions by calculating sample statistics using weights derived from the sampling design employed in the survey. However, direct estimation can also adopt a model-based approach, where population parameters are estimated from the sample without accounting for the sampling design. The model-based approach has a drawback: if the model does not align with the characteristics of the data, the estimates can be inaccurate, even with a larger sample size. Therefore, the design-based approach is more commonly used.

2.4. Small Area Estimation

Small Area Estimation (SAE) is a statistical technique used to estimate parameters in small domains. A domain is considered small if the available sample size within that domain is insufficient for producing precise direct estimates [17]. Domains can be defined as geographic areas, social or demographic groups, or other subpopulations. In this study, out-of-school children with disabilities constitute a subpopulation of the overall population of persons with disabilities. SAE can serve as a solution to address insufficient sample sizes. It is an indirect estimation model with an explicit linking model, meaning that the additional information comes not only from auxiliary variables but also from the variation between areas resulting from area-specific random effects. The random effects refer to the variation between areas that cannot be explained by the auxiliary information. Thus, there are two main concepts in SAE: the fixed effect model, which assumes that the variation in the response variable within small areas can be fully explained by auxiliary information, and the small area random effect, which accounts for variation in small areas that cannot be explained by auxiliary information, known as the area random effect. Based on the availability of auxiliary variables, there are two types of SAE models: area-level models and unit-level models [17]. Most researchers use area-level models because data are often not available at the unit level.

In the area-level model, auxiliary variables are available only at the level of the area under study. Let $\hat{\theta}_i$ represent a direct and unbiased estimate of the parameter θ_i. Since $\hat{\theta}_i$ contains sampling error, a sampling model can be formulated as follows.

$$\hat{\theta}_i = \theta_i + e_i\ ; i = 1, \dots, m \tag{2}$$

There is a correlation between θ_i and the auxiliary variable for each area, denoted by $\boldsymbol{x_i}$. The linear model, serving as the linking model in the area-level model, is expressed as follows [17].

$$\theta_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + b_i v_i \; ; i = 1, \dots, m \tag{3}$$

By combining the sampling model in equation (2) with the linking model in equation (3), the basic area-level model, commonly known as the Fay-Herriot model, is obtained [17].

$$\hat{\theta}_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + b_i v_i + e_i \; ; i = 1, \dots, m \tag{4}$$

2.5. Twofold HB Beta Model

Essentially, the twofold subarea-level SAE model is an extension of the area-level model that can provide estimates for a given area and further divide it into smaller subareas. This model is referred to as such because auxiliary variables are available at the subarea level. Using this model, estimation can be performed simultaneously at different aggregation levels [17]. Suppose there are m areas, and each area is divided into N_i subareas with $i = 1, \dots, m$. If we aim to estimate the parameter at the area level θ_i and at the subarea level θ_{ij} where $j = 1, \dots, N_i$ and $i = 1, \dots, m$, the equation of the twofold HB beta SAE model can be written as follows [30].

i. $\hat{\theta}_{ij} | \theta_{ij} \sim Beta(a_{ij}, b_{ij})$ (sampling model) (5)

$$a_{ij} = \theta_{ij} \left(\frac{n_i}{deff_i} - 1 \right) \tag{6}$$

$$b_{ij} = (1 - \theta_{ij}) \left(\frac{n_i}{deff_i} - 1 \right) \tag{7}$$

ii. $logit(\theta_{ij}) \sim N(\boldsymbol{x}_{ij}^T \boldsymbol{\beta}, \sigma_u^2)$ (linking model) (8)

iii. $\boldsymbol{\beta}_j \sim N\left(\mu_{\beta_j}, \sigma_{\beta_j}^2\right)$ (prior for $\boldsymbol{\beta}_j$) (9)

iv. $\sigma_u^2 \sim IG(t_1, t_2)$ (prior for σ_u^2) (10)

In this model, parameter estimation can be obtained numerically using the MCMC method with the Gibbs sampling algorithm. The Gibbs sampling estimation is performed for B iterations. The algorithm then produces a sequence of estimated parameters. Accordingly, the parameter estimates at the subarea level can be calculated as follows.

$$\hat{\theta}_{ij}^{HB} = \frac{1}{B} \sum_{k=d+1}^{d+B} \theta_{ij}^{(k)} = \theta_{ij}(.) \tag{11}$$

The estimates at the area level can then be obtained by aggregating the subarea-level estimates using the weights available for each subarea, as follows.

$$\theta_{i.}^{(k)} = \sum_{j=1}^{N_i} W_{ij} \theta_{ij}^{(k)} \tag{12}$$

$$\hat{\theta}_{i.}^{HB} = \frac{1}{B} \sum_{k=d+1}^{d+B} \theta_{i.}^{(k)} \tag{13}$$

2.6. Relative Standard Error

The quality of an estimate can be assessed by its total error, which reflects the difference between the estimated value and the true value. The smaller the total error, the more accurate the estimate, as it deviates less from the true value. Assuming bias is relatively small, accuracy can be approximated by precision. In this context, Relative Standard Error (RSE) is a measure of the precision of an estimate relative to its estimated value, calculated by comparing the standard error to the estimate and expressed as a percentage [23]. RSE can be calculated using the following equation.

$$RSE(\hat{\theta}) = \frac{SE(\hat{\theta})}{\hat{\theta}} \times 100\% \tag{14}$$

where $\hat{\theta}$ is the estimated value for a specific domain and $SE(\hat{\theta})$ is the measure of precision represented by standard error.

Sampling errors in the estimation results need to be considered. BPS sets RSE thresholds to assess the accuracy of an estimate: an estimate with RSE ≤ 25% is considered accurate, an estimate with RSE > 25% and ≤ 50% is considered moderately accurate and should be used with caution, and an estimate with RSE > 50% is regarded as highly inaccurate [23].

2.7. Analysis Method

This study was conducted to obtain estimates of out-of-school children with disabilities in Indonesia. This is motivated by the fact that BPS, as the data provider for out-of-school children with disabilities in Indonesia, is only able to publish data at the national level. Estimating out-of-school children with disabilities could be done using direct estimates from the March 2023 Susenas data. However, obtaining reliable estimates for the population of children with disabilities remains challenging. This is because the survey's minimum sample size was not designed for the disability group, so the small sample size results in estimates with relatively large errors.

One of the solution to address inaccurate estimates due to high relative standard errors (RSE) is to apply indirect estimation methods. Indirect estimation, which utilizes auxiliary information, is known as small area estimation (SAE). As the demand for small-area data continues to increase, requiring information down to the smallest administrative units, this study estimates out-of-school children with disabilities simultaneously at two levels of aggregation: provincial and district/city levels. Accordingly, the proposed model is a twofold subarea-level model using the Hierarchical Bayes (HB) estimation method with a beta distribution. After performing indirect estimation, the results will be evaluated by comparing the RSE values obtained. The estimation method that produces the lowest RSE will be selected as the best method for estimating out-of-school children with disabilities in Indonesia. Based on the best method, the estimated results will then be mapped.

The auxiliary variables must be must be sourced from census data or other datasets that are free from sampling error. At the time of the research, the only available and most recent source meeting this criterion was the 2021 Village Potential Statistics (Podes). There are 20 candidate auxiliary variables obtained from Podes 2021, including the proportion of villages/subdistricts classified as rural (X_1), the proportion of families living under high-voltage transmission lines (X_2), the proportion of families living along riverbanks (X_3), the proportion of families living in slum areas (X_4), the ratio of primary schools per 10,000 population (X_5), the ratio of junior high schools per 10,000 population (X_6), the ratio of senior high schools/vocational schools per 10,000 population (X_7), the ratio of special needs primary schools per 100 children with disabilities (X_8), the ratio of special needs junior high schools per 100 children with disabilities (X_9), the ratio of special needs senior high schools per 100 children with disabilities (X_{10}), the ratio of hospitals per 10,000 population (X_{11}), the ratio of community health centers per 10,000 population (X_{12}), the ratio of PKK per 10,000 population (X_{13}), the ratio of youth organizations per 10,000 population (X_{14}), the ratio of customary institutions per 10,000 population (X_{15}), the ratio of micro and small industries per 10,000 population (X_{16}), the ratio of financial institution facilities per 10,000 population (X_{17}), the ratio of cooperatives per 10,000 population (X_{18}), the ratio of economic facilities and infrastructure per 10,000 population (X_{19}), and the ratio of issued poverty certificates per 100 families (X_{20}).

3. Results and Discussion

3.1. Direct Estimate

The direct estimation of the number of out-of-school children with disabilities was obtained through a design-based direct estimation approach. The calculation process utilized sampling weights from the March 2023 Susenas and was conducted across 34 provinces in Indonesia. Figure 3 presents the visualization of the direct estimates of out-of-school children with disabilities at the provincial level in Indonesia for 2023.

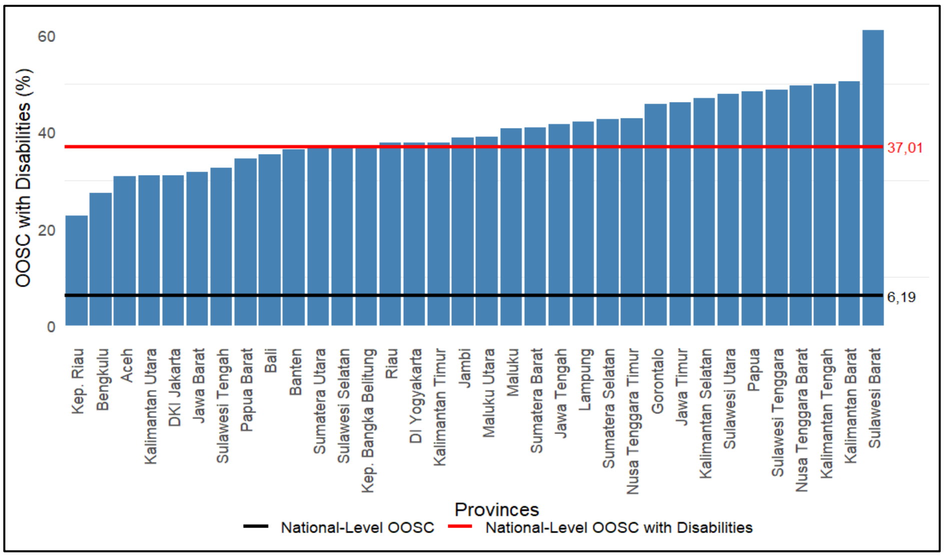


Figure 3. Direct estimates of out-of-school children with disabilities at the provincial level in Indonesia, 2023

Figure 3 shows the out-of-school children with disabilities across all provinces in Indonesia in 2023. The black horizontal line represents the national percentage of out-of-school for all children, regardless of disability status, which is 6.19%, while the red line indicates the national percentage of out-of-school for children with disabilities, which is 37.01%. It is evident that all provinces are above the national rate of 6.19%, indicating that children with disabilities face greater gaps and vulnerabilities in access to education compared to the general population. Moreover, more than half of the provinces have percentage of out-of-school for children with disabilities that exceed the national rate for this group. Only eleven provinces fall below the national percentage of out-of-school for children with disabilities. For a clearer overview of the direct estimates of out-of-school children with disabilities at the provincial level in Indonesia, descriptive statistics, along with their relative standard errors, are summarized in Table 1.

Table 1. Descriptive statistics of direct estimates of out-of-school children with disabilities at the provincial level in Indonesia, 2023

Descriptive Statistics	Direct Estimate (%)	RSE (%)
Minimum	22.64	13.51
First Quartile	35.68	24.29
Median	38.98	29.95
Mean	40.06	36.54
Third Quartile	46.05	43.65
Maximum	61.03	91.06

Based on Table 1, the lowest percentage of out-of-school children with disabilities is 22.64 percent, while the highest is 61.03 percent. The considerable gap between the highest and lowest values indicates significant disparities in education for children with disabilities. In addition to examining the direct estimates, it is also important to assess the accuracy of these estimates using the RSE values. According to Table 1, the highest and lowest RSE values for the direct estimates are 91.06 percent and 13.51 percent, respectively. The average RSE is 36.54 percent indicates that the estimates are not entirely accurate and, therefore, cannot yet serve as a strong basis for policy formulation.

The issue of out-of-school children with disabilities also needs to be analyzed at the district/city level. This aligns with the implementation of regional autonomy, whereby local governments are granted the rights and authority to manage their own affairs in accordance with the principles of decentralization. One of these affairs is education, which is primarily the responsibility of district/city governments. This means that policies related to education and interventions for out-of-school children are highly dependent on the capacity and commitment of the local government. By understanding the condition of out-of-school children with disabilities at a smaller level, namely districts/cities, local governments can formulate policies that are more responsive, data-driven, and tailored to local needs.

Direct estimates of out-of-school children with disabilities have also been conducted for 514 districts/cities in Indonesia. However, direct estimates at the district/city level are only available for 451 out of the 514 districts/cities. The remaining 63 districts/cities could not be directly estimated due to the absence of samples of children with disabilities in these areas. For a clearer understanding, the following is a visualization of districts/cities based on the availability of samples of children with disabilities.

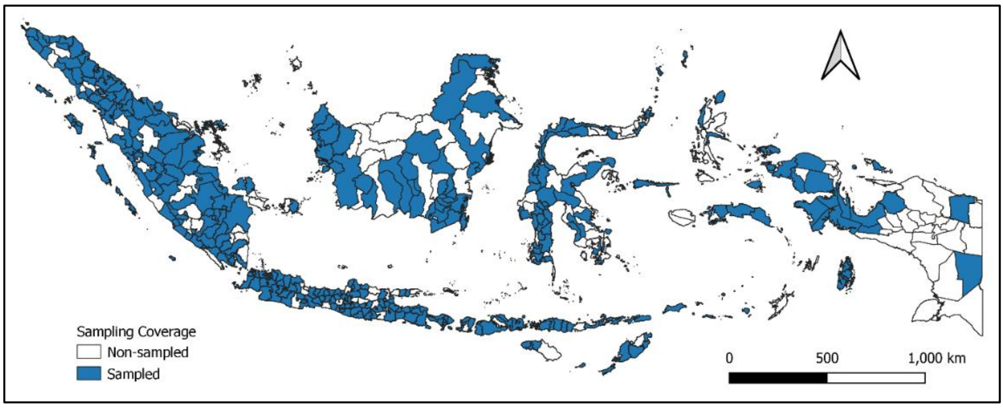


Figure 4. Map of districts/cities based on sample availability

As shown in Figure 4, there are several areas without samples of children with disabilities (non-sampled areas), indicated in white. A total of 63 districts/cities fall into the non-sampled category, and they are scattered across the country, though they tend to cluster in the northern part of Kalimantan and the western part of Papua. The existence of non-sampled areas reflects the limitations of surveys in reaching children with disabilities. This situation results in most regions lacking an adequate number of samples or even having no samples at all. Excluding these non-sampled areas, the following presents a summary of the descriptive statistics of the direct estimates of out-of-school children with disabilities at the district/city level in Indonesia.

Table 2. Descriptive statistics of direct estimates of out-of-school children with disabilities at the district/city level in Indonesia, 2023

Descriptive Statistics	Direct Estimate (%)	RSE (%)
Minimum	0	0
First Quartile	0	31.14
Median	18.77	60.58
Mean	30.36	58.47
Third Quartile	52.77	90.34
Maximum	100	137.99

Based on Table 2, the direct estimates of out-of-school children with disabilities range from a minimum of 0 percent to a maximum of 100 percent. There are 170 districts/cities with an estimated value of 0 percent, indicating that there were no recorded cases of children with disabilities being out of school. However, due to the limitations of the survey used, this value cannot be taken as conclusive evidence of the population condition of children with disabilities. In other words, it should not be interpreted as meaning that there are absolutely no out-of-school children with disabilities in those areas. Similarly, in regions with an estimated value of 100 percent, it does not imply that all children with disabilities are out of school. These extreme values rather reflect the limitations of the data in reaching small population groups.

Table 2 also shows that the minimum relative standard error (RSE) is 0 percent. An RSE value of zero does not indicate that the estimate is highly accurate or free from error; instead, it signals an anomaly in the data. Generally, an RSE of zero occurs when data variation cannot be calculated, usually due to very small and unrepresentative sample sizes. Moreover, the average RSE of 58.47 percent and the median RSE of 60.58 percent indicate that, overall, the quality of direct estimates at the district/city level is very inaccurate. An RSE exceeding 50 percent suggests that the estimates carry high uncertainty or are highly unreliable, making them unusable and uninterpretable. This is further reinforced by the maximum RSE value of 137.99 percent, which indicates extremely high estimation uncertainty, far beyond statistically acceptable limits.

Table 3. Number of regions by RSE category

Category	District/City	Province
RSE ≤ 25%	62	11
25% < RSE ≤ 50%	51	16
RSE > 50%	159	7
NA	242	0

Based on the direct estimation results (Table 3), only 62 districts/cities demonstrate good accuracy. In fact, there are 242 districts/cities with RSE values reported as not available (NA). At the provincial level, most estimates are also relatively inaccurate. Only 11 provinces have accurate estimates, as indicated by RSE values less than or equal to 25 percent.

3.2. Auxiliary Variable Selection

Before entering the SAE modeling stage, auxiliary variables are selected by first examining the relationship between candidate auxiliary variables and the response variable. The selection of auxiliary variables is crucial because choosing the appropriate auxiliary variables is an important component for the success of indirect estimation using SAE [17]. One of the methods to statistically confirm the existence of relationships between variables is through Pearson correlation testing. In this study, the twofold HB beta SAE model is applied, so the candidate auxiliary variables are tested for their correlation with the logit-transformed direct estimates. Pearson correlation serves as an effective preliminary screening tool to identify variables that exhibit meaningful linear associations with the transformed response. It is considered sufficient at this stage because the purpose is only to filter out clearly irrelevant predictors before applying more comprehensive model-based evaluations [31]. Based on the correlation test results, six auxiliary variables were found to have a significant correlation with the logit of the direct estimator.

After conducting the correlation tests, it is necessary to ensure that there is no correlation among auxiliary variables to prevent multicollinearity in the model. Multicollinearity detection is carried out by examining Pearson correlation values. Collinearity occurs when the correlation between auxiliary variables exceeds 0.8 [32]. The results of multicollinearity detection among the auxiliary variables indicate that no variable has a correlation greater than 0.8 with another variable. Therefore, no auxiliary variables were eliminated, and all can proceed to the next stage.

The final step is variable selection using the backward elimination method. This approach is consistent with the study by Mellinda & Sumarni, which stated that backward elimination in SAE modeling can produce better results [33]. The variable selection process ultimately retained four variables: the ratio of junior high schools per 10,000 population (X_6), the ratio of public health centers per 10,000 population (X_{12}), the ratio of youth organizations per 10,000 population (X_{14}), and the ratio of issued poverty letters/SKTM per 100 families (X_{20}). These four variables are able to reflect various dimensions of life that are relevant in explaining the conditions of out-of-school children.

The ratio of junior high schools per 10,000 population (X_6) represents the educational dimension, while the ratio of public health centers per 10,000 population (X_{12}) represent the health dimension. Adequate availability of and access to schools and healthcare facilities can support the learning process of children optimally, both in terms of physical and mental readiness [34]. The ratio of youth organizations per 10,000 population (X_{14}) represent the social dimension. The organization is oriented toward achieving social welfare for the community, thereby serving as an inclusive and open space for all groups, including children with disabilities. Such social support can enhance the self-confidence and courage of children with disabilities to interact with the outside world, including in the context of attending school. The ratio of issued poverty letters per 100 families (X_{20}) represent the economic dimension. Economic limitations can increase the risk of children with disabilities being out of school [35].

3.3. Development SAE using Twofold HB Beta Model

In parameter estimation using the HB beta approach, calculations are based on the posterior distribution, which is derived from the prior distribution and the likelihood function. The resulting posterior equation is highly complex, making direct integration extremely difficult. One method that can address this complexity is the Markov Chain Monte Carlo (MCMC) method. The MCMC method produces an accurate posterior distribution once it reaches equilibrium, that is, when the algorithm has converged [36]. In this study, several trials were conducted to achieve convergence, specifically using 60,000 iterations, 30 update iterations, a thinning interval of 25, and a burn-in of 15,000.

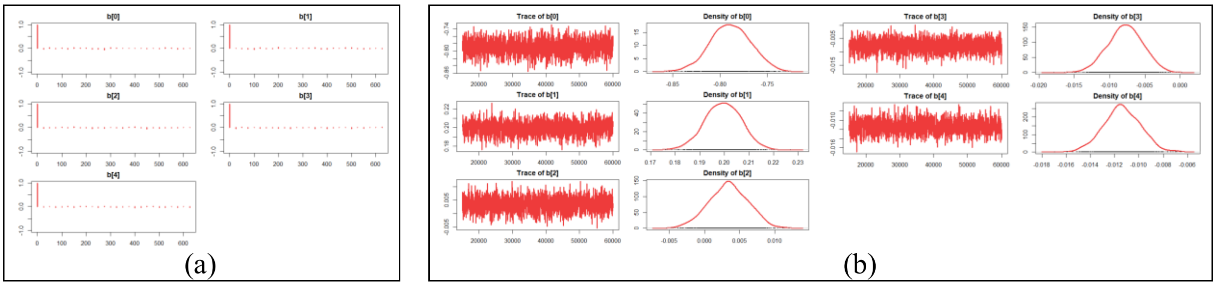


Figure 5. (a) Autocorrelation plot; (b) Trace plot and density plot

According to the Figure 5, convergence was confirmed by autocorrelation plots showing a clear cut-off, stationary trace plots without periodic patterns, and density plots for each parameter resembling a normal distribution. After obtaining convergence results for the four auxiliary variables included in the model, the following presents the estimated parameter coefficients from the twofold HB beta SAE model for out-of-school children with disabilities, with subareas defined as districts/cities and areas defined as provinces.

Table 4. Estimated parameter coefficients of the twofold HB beta SAE model

Estimated Parameter Coefficients	Mean	SD	2.5%	97.5%
b_0	-0.789	0.021	-0.835	-0.748
b_1	0.200	0.007	0.187	0.215
b_2	-0.003	0.002	-0.009	-0.001
b_3	-0.008	0.003	-0.013	-0.003
b_4	0.011	0.001	0.008	0.014

Based on Table 4, it can be seen that the parameter estimates associated with the variables used, namely X_6, X_{12}, X_{14}, and X_{20}, are significant in the twofold HB beta model. Significance in the model is assessed using the credible interval ranging from 2.5 percent to 97.5 percent, where a variable is considered significant if the credible interval within this range does not include zero. An overview of the indirect estimation results of out-of-school children with disabilities using the twofold HB beta SAE model is presented in Table 5.

Table 5. Descriptive statistics of indirect estimate of out-of-school children with disabilities in Indonesia, 2023

Descriptive Statistics	District/City (%)	Province (%)
Minimum	12.43	28.07
First Quartile	36.70	38.12
Median	42.28	42.20
Mean	42.70	41.88
Third Quartile	48.58	46.06
Maximum	82.49	60.71

Based on Table 5, the indirect estimates of out-of-school children with disabilities in Indonesia in 2023 using the twofold HB beta SAE model show an average of 42.70 percent at the district/city level and 41.88 percent at the provincial level. At the district/city level, the lowest estimate is 12.43 percent in Banda Aceh City, while the highest is 82.49 percent in Lamandau Regency. At the provincial level, the lowest estimate is 28.07 percent in Aceh Province, and the highest is 60.71 percent in West Sulawesi Province.

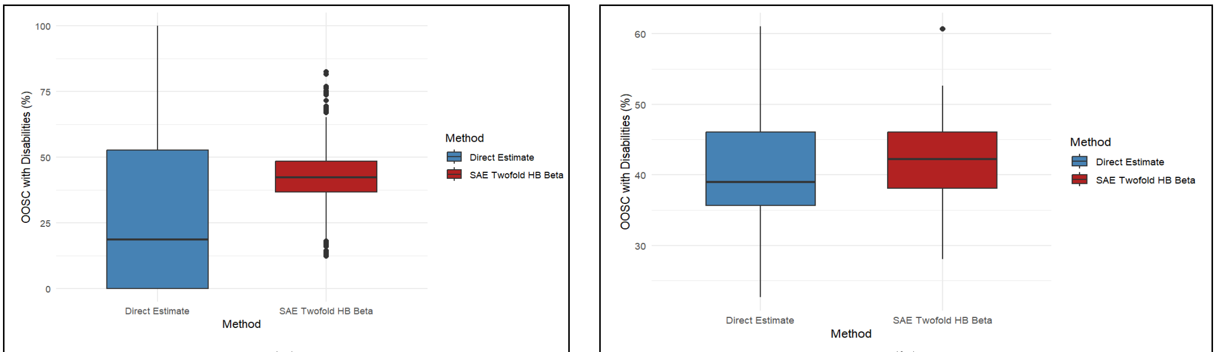
3.4. Evaluation

The evaluation was carried out by comparing the estimates of out-of-school children with disabilities obtained from both direct estimation and indirect estimation using the twofold HB beta SAE model. The indirect estimates from the twofold HB beta SAE model are considered reliable and unbiased when their values are not significantly different from those of the direct estimates [37]. The following presents a comparative statistical summary of the estimates of out-of-school children with disabilities derived from direct and indirect estimation.

Table 6. Descriptive statistics of direct and indirect estimate of out-of-school children with disabilities in Indonesia, 2023

Descriptive Statistics	District/City		Province	
	Direct Est. (%)	Indirect Est. (%)	Direct Est. (%)	Indirect Est. (%)
Minimum	0	12.43	22.64	28.07
First Quartile	0	36.70	35.68	38.12
Median	18.77	42.28	38.98	42.20
Mean	30.36	42.70	40.06	41.88
Third Quartile	52.77	48.58	46.05	46.06
Maximum	100	82.49	61.03	60.71
NA	63	0	0	0

Table 6 shows a noticeable difference between the direct estimates and the indirect estimates using the twofold HB beta SAE model in estimating out-of-school children with disabilities, particularly at the district/city level. The most striking differences are observed in the minimum and maximum values. At the provincial level, no striking differences are observed; however, there is still a change in the average value, from 40.06 percent to 41.88 percent.



(a) (b)

Figure 6. Boxplot of direct and indirect estimates of out-of-school children with disabilities at (a) the district/city level; (b) the provincial level in Indonesia, 2023

Further examination through the data distribution using the boxplot in Figure 6 reveals that the estimates from the twofold HB beta SAE model still fall within the main distribution range of the direct estimates, indicating that the differences do not exceed the extreme limits of data dispersion. The notable differences arise from the limitations of the direct estimation method, which is highly influenced by the number and representativeness of the sample in reflecting the actual population conditions. Direct estimates tend to be less accurate due to these sample constraints, whereas the SAE model is able to correct such issues. To determine the most appropriate estimation method for estimating out-of-school children with disabilities, an evaluation must be conducted based on the accuracy level of the estimates. In this study, the measure of accuracy used is the relative standard error (RSE). Presented below is the number of regions classified by RSE category.

From the comparison of RSE categories in Table 7, it can be observed that the use of the twofold HB beta SAE model is able to reduce RSE values and address the problem of non-sampled areas. Even at the provincial level, all 34 provinces have RSE values below 25 percent, indicating that the estimates are accurate. Most notably, the SAE method also succeeded in providing estimates for all areas without a single missing value. These findings reinforce that the twofold HB beta SAE model is not only superior in producing more accurate estimates but also capable of generating estimates for all areas, including non-sampled ones. Its ability to lower RSE and produce comprehensive estimates makes this method the most appropriate choice for addressing the problem of limited sample data.

Table 7. Number of regions by RSE category

Category	District/City		Province	
	Direct Est. (%)	Indirect Est. (%)	Direct Est. (%)	Indirect Est. (%)
RSE ≤ 25%	62	70	11	34
25% < RSE ≤ 50%	51	406	16	0
RSE > 50%	159	38	7	0
NA	242	0	0	0

3.5. Mapping of the Estimated OOSC with Disabilities

To observe the distribution of out-of-school children with disabilities (OOSC) in Indonesia more clearly, a thematic map is used. Since there is no specific guideline for classifying OOSC with disabilities, the classification is carried out using three groups generated through the natural breaks classification (Jenks optimization method). The natural breaks classification method divides groups by minimizing within-group variation and maximizing between-group variation, making it suitable for data that are not normally distributed [38]. The use of the same classification for OOSC with disabilities at both the district/city and provincial levels aims to enable better comparison across aggregation levels. Based on the natural breaks classification with three groups, the resulting breakpoints are 35.36 percent and 45.34 percent. The first group consists of areas with OOSC with disabilities of less than 35.36 percent. The second group consists of areas with OOSC values ranging from a minimum of 35.36 percent to a maximum of 45.34 percent. Finally, the third group consists of areas with OOSC values above 45.34 percent.

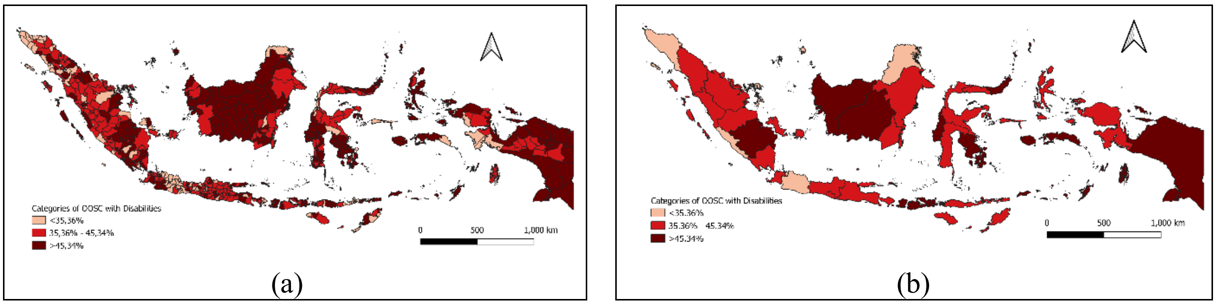


Figure 7. Thematic map based on estimates of out-of-school children with disabilities at (a) the district/city level; (b) the provincial level in Indonesia, 2023

Higher percentages of out-of-school children with disabilities are represented by darker colors, and vice versa. Based on Figure 7, most areas, both at the provincial and district/city levels, fall within the 35.36 to 45.34 percent category, while 195 districts/cities and 9 provinces have values exceeding 45.34 percent. This indicates that the issue of out-of-school children with disabilities is not confined to specific regions but is a widespread challenge across Indonesia. The dominance of areas in the 35.36 to 45.34 percent range reflects the systemic nature of limited educational access for children with disabilities, while the existence of regions in the highest category underscores more concerning conditions that demand greater attention and more serious interventions in policy planning and implementation.

As shown in the thematic map in Figure 7, the highest values of out-of-school children with disabilities tend to cluster in Kalimantan and Papua. This is closely related to geographical factors, which serve as a major barrier to educational access in these areas. Many regions in eastern Indonesia, such as Papua, Maluku, and several remote parts of Kalimantan, are characterized by challenging topography, including mountainous areas or small isolated islands. These conditions hinder access to educational services, particularly for children with disabilities who require mobility support and special accommodations [39]. Katheryn Bennet, representing UNICEF Indonesia, also stated that the high rate of out-of-school children in Indonesia is caused by various factors, such as economic pressures that push children into work, limited geographical access, and other barriers [40]. For children with disabilities, these barriers are often exacerbated by the family's low economic condition, which prevents them from affording additional needs such as special transportation, learning aids, or companion costs. As a result, schooling is no longer seen as a main priority, and many of them instead choose to work to help support their family's economy [41].

The Government of Indonesia has also introduced various policies to address the issue of out-of-school children with disabilities. However, in practice, implementation on the ground remains uneven. For instance, in the provision of inclusive education, many regular schools still lack disability-friendly facilities such as wheelchair ramps, special toilets, or visual/auditory aids. In addition, the shortage of Special Education Support Teachers poses a serious challenge in supporting the learning process of children with disabilities [42]. Furthermore, data collection on children with disabilities remains suboptimal, hindering evidence-based and well-targeted policy planning.

The estimation results indicate that addressing out-of-school children with disabilities requires more than education policies alone. A multisectoral approach is necessary, involving education, social affairs, health, infrastructure, and other sectors. Key measures include improving disability-friendly infrastructure, providing social protection such as financial aid and learning support, and enhancing educational quality through special education teachers, training, and inclusive curricula. These efforts must be coordinated across national and regional agencies and supported by accurate, up-to-date data to monitor progress, evaluate policies, and design programs that reflect the real needs of the field.

4. Conclusion

Overall, the twofold HB beta SAE model is the most suitable method for estimating the number of out-of-school children with disabilities in Indonesia at both provincial and district/city levels, as it addresses sample limitations and improves estimation accuracy. The estimated proportion ranges from 35.36 to 45.34 percent, with the highest concentrations observed in Kalimantan and Papua. These results provide evidence for local governments to support inclusive education policy formulation and evaluation. More targeted regional analysis and improved inter-agency data integration are needed, especially in areas with complex socio-geographic conditions.

Ethics approval

Not required.

Acknowledgments

Not required.

Competing interests

All the authors declare that there are no conflicts of interest.

Funding

This study received no external funding.

Underlying data

This research uses secondary data from BPS Statistics Indonesia.

Credit Authorship

Aisha Maharani: Conceptualization, Methodology, Writing, Editing, Visualization. **Azka Ubaidillah:** Reviewing, Validation, Supervision. **Adhi Kurniawan:** Reviewing, Editing, Validation.

References

[1] D. R. Taneo, S. E. P. R. A. Ratu, and D. R. . Lodang, "Persepsi peserta didik terhadap penggunaan teknologi pembelajaran pada mata pelajaran seni budaya di Sekolah Menengah Kejuruan Negeri 2 Kupang," *Metta J. Ilmu Multidisiplin*, vol. 4, no. 3, pp. 132–143, Oct. 2024, doi: https://doi.org/10.37329/metta.v4i3.3607.

[2] N. Larasati and I. Rindaningsih, "The quality improve of education: implementation, development, and human resources," *Int. J. Multidiscip.*, vol. 1, no. 1, 2024, [Online]. Available: https://journal.antispublisher.com/index.php/ijmi

[3] UNDP, *Human development report 1990*. New York: Oxford University Press, 1990.

[4] Bappenas, "Sustainable development goals 4 : pendidikan berkualitas," https://sdgs.bappenas.go.id/17-goals/goal-4/.

[5] UNESCO, "UNESCO: 250 million children now out of school," Sep. 2023. Accessed: Jan. 20, 2025. [Online]. Available: https://news.un.org/en/story/2023/09/1140882

[6] UNESCO, "250 million children out-of-school: what you need to know about UNESCO's latest education data," Sep. 2023. Accessed: Jun. 13, 2025. [Online]. Available: https://www.unesco.org/en/articles/250-million-children-out-school-what-you-need-know-about-unescos-latest-education-data

[7] BPS, *Statistik pendidikan 2023*, vol. 12. Badan Pusat Statistik, 2023.

[8] I. Rosmilawati and M. F. T. Mutaqin, "Eradicating out-of-school-children: strategies and best practices from alternative education context," *JPPI (Jurnal Penelit. Pendidik. Indones.*, vol. 10, no. 2, p. 364, Jun. 2024, doi: 10.29210/020244055.

[9] Y. Anggrayni, "Major influencing factors of out-of-school in West Java Province, Indonesia," *Asia-Japan Res. Acad. Bull.*, vol. 5, no. 12, Jul. 2024, doi: https://doi.org/10.34389/asiajapanbulletin.5.0_12.

[10] I. R. Marfu'ah, D. N. Yasmara, and M. N. F. Al Amin, "Analisis program pendidikan inklusif bagi disabilitas di Kota Kediri dengan perspektif problem tree analysis," *J. Sos. Dan Hum.*, vol. 1, no. 4, pp. 178–189, May 2024, doi: 10.62017/arima.

[11] United Nations, "Convention on the rights of persons with disabilities and optional protocol," 2006.

[12] E. A. Lauer and A. J. Houtenville, "Estimates of prevalence, demographic characteristics and social factors among people with disabilities in the USA: a cross-survey comparison," Feb. 01, 2018, *BMJ Publishing Group*. doi: 10.1136/bmjopen-2017-017828.

[13] United Nations, "Laporan PBB menyoroti kesenjangan data disabilitas di Indonesia," 2022. [Online]. Available: https://indonesia.un.org/id/209911-laporan-pbb-menyoroti-kesenjangan-data-disabilitas-di-indonesia

[14] B. Berlian and A. Firdaus, "Good governance melalui kebijakan berbasis bukti reformulasi sistem peradilan pidana nasional," *J. Yuridis*, vol. 9, no. 1, pp. 27–36, Jun. 2022, doi: 10.35586/jyur.v9i1.3847.

[15] A. H. Hakim and N. Hajarisman, "Pendugaan rata-rata pengeluaran per kapita menurut Kabupaten/Kota di Provinsi Jawa Barat melalui empirical best linear unbiased prediction dalam pendugaan area kecil," *Bandung Conf. Ser. Stat.*, vol. 2, no. 2, pp. 474–481, Aug. 2022, doi: 10.29313/bcss.v2i2.4747.

[16] J. N. K. Rao and I. Molina, *Small area estimation*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2015.

[17] A. L. Erciulescu, N. B. Cruze, and B. Nandram, "Model-based county level crop estimates incorporating auxiliary sources of information," *J. R. Stat. Soc. Ser. A Stat. Soc.*, vol. 182, no. 1, pp. 283–303, Jan. 2019, doi: 10.1111/rssa.12390.

[18] R. Saadi and A. Ubaidillah, "Pendugaan area kecil angka partisipasi kasar perguruan tinggi di Indonesia menggunakan model subarea twofold dengan pendekatan hierarchical bayes," *Semin. Nas. Off. Stat.*, vol. 2022, no. 1, pp. 123–136, Nov. 2022, doi: https://doi.org/10.34123/semnasoffstat.v2022i1.1217.

[19] R. F. Yudasena, "Penerapan SAE twofold HB beta dengan kecamatan sebagai subarea dan kabupaten/kota sebagai area (studi kasus estimasi persentase anak miskin tingkat Kecamatan dan Kabupaten/Kota di Jawa Tengah pada tahun 2022)," *Jakarta Politek. Stat. STIS*, 2024.

[20] W. R. Eliezer *et al.*, "Small area estimation for percentage of out-of-school children aged 7-17 years in Sumatera Island, 2023," *J. Mat. Stat. dan Komputasi*, vol. 21, no. 1, pp. 215–233, Sep. 2024, doi: 10.20956/j.v21i1.36043.

[21] L. F. Trihandika, A. Ubaidillah, A. Az-Zahra, A. Amirudin, M. Rusydiana, and Z. Maharani, "Small area estimation anak tidak sekolah di Pulau Kalimantan tahun 2023," *Limits J. Math. Its*

Appl., vol. 21, no. 2, p. 273, Oct. 2024, doi: 10.12962/limits.v21i2.21497.

[22] BPS, *Buku 4 Konsep dan definisi susenas maret 2023*. Jakarta, 2022.

[23] UNICEF, *Definition and Classification of Disability*. New York, 2014.

[24] H. Anttila, S. Tallqvist, M. Muñoz, S. Leppäjoki-Tiistola, O. Mäkitie, and S. Hiekkala, "Towards an ICF-based self-report questionnaire for people with skeletal dysplasia to study health, functioning, disability and accessibility," *Orphanet J. Rare Dis.*, vol. 16, no. 1, Dec. 2021, doi: 10.1186/s13023-021-01857-7.

[25] S. van der Veen, N. Evans, M. Huisman, P. Welch Saleeby, and G. Widdershoven, "Toward a paradigm shift in healthcare: using the International Classification of Functioning, Disability and Health (ICF) and the capability approach (CA) jointly in theory and practice," *Disabil. Rehabil.*, vol. 45, no. 14, pp. 2382–2389, 2023, doi: 10.1080/09638288.2022.2089737.

[26] S. Mitra, S. K. Mishra, and R. K. Abhay, "Out-of-school girls in India: a study of socioeconomic-spatial disparities," *GeoJournal*, vol. 88, no. 1, pp. 341–357, Feb. 2023, doi: 10.1007/s10708-022-10579-7.

[27] UNICEF and UNESCO Institute for Statistics, *Operational manual: global out-of-school children initiative*. UNICEF, 2015. [Online]. Available: www.unicef.org/education/

[28] A. Shanker, D. Marian, and C. Swimmer, "Effective interventions aimed at reaching out-of-school children a literature review," Kathmandu, 2015. Accessed: Apr. 23, 2025. [Online]. Available: https://files.eric.ed.gov/fulltext/ED573790.pdf

[29] M. Torabi and J. N. K. Rao, "On small area estimation under a sub-area level model," *J. Multivar. Anal.*, vol. 127, pp. 36–55, May 2014, doi: 10.1016/j.jmva.2014.02.001.

[30] W. Ren, J. Li, A. Erciulescu, T. Krenzke, and L. Mohadjer, "A variable selection method for small area estimation modeling of the proficiency of adult competency," pp. 689–713, 2022, [Online]. Available: https://www.mdpi.com/2571-905X/5/3/41

[31] D. Gujarati, D. Porter, and S. Gunasekar, *Basic Econometrics*. 2013.

[32] Mellinda and C. Sumarni, "Estimation of total fertility rate (TFR) using small area estimation (SAE) in Nusa Tenggara Timur (NTT) Province," *Proc. Int. Conf. Data Sci. Off. Stat.*, vol. 2021, no. 1, pp. 796–809, 2022, doi: https://doi.org/10.34123/icdsos.v2021i1.107.

[33] G. S. Becker, *Human capital: a theoretical and empirical analysis with special reference to education (third edition)*, 3rd ed. The University of Chicago Press, 1993.

[34] UNESCO Institute for Statistics and UNICEF, *Fixing the broken promise of education for all: findings from the global initiative an out-of-school children*. Montreal: UNESCO Institute for Statistics, 2015. doi: 10.15220/978-92-9189-161-0-en.

[35] A. Ubaidillah, "Small area estimation dengan pendekatan hierarchical bayesian neural network untuk pemetaan kemiskinan di Kota Jambi (Tesis S2)," *Inst. Teknol. Sepuluh Novemb.*, 2014, [Online]. Available: https://www.researchgate.net/publication/340172483

[36] Eurostat, *Guidelines on small area estimation for city statistics and other functional geographies*. European Union, 2019. doi: 10.2785/763467.

[37] A. Shafira, F. Kristiani, and B. Yong, "Penerapan metode klasifikasi perangkat lunak arcmap pada pemetaan penyebaran penyakit dengue di Bandung," *Limits J. Math. Its Appl.*, vol. 20, no. 1, p. 39, Mar. 2023, doi: 10.12962/limits.v20i1.9226.

[38] Bappenas, *Pedoman teknis penyusunan rencana aksi tujuan pembangunan berkelanjutan (TPB) / sustainable development goals (SDGs)*, II. Kedeputian Bidang Kemaritiman dan Sumber Daya Alam, Kementerian Perencanaan Pembangunan Nasional / Badan Perencanaan Pembangunan Nasional, 2020. Accessed: Apr. 23, 2025. [Online]. Available: https://sdgs.bappenas.go.id/website/wp-content/uploads/2020/10/Buku-Pedoman-Rencana-Aksi-SDGs.pdf

[39] R. Luthfia, "Mengatasi masalah anak tidak sekolah melalui pendidikan nonformal," PSKP Kemendikdasmen. [Online]. Available: https://pskp.kemendikdasmen.go.id/kabar/detail/mengatasi-masalah-anak-tidak-sekolah-melalui-pendidikan-nonformal

[40] M. I. Nasrallah, "Opportunities and challenges to support out- of-school children and youth through accelerated learning programmes: case study of Jordan," 2022, [Online]. Available: https://inee.org/evidence/e-cubed.

[41] P. Lydiasari, "Sekolah inklusi di Indonesia baru 14,8 persen yang memiliki guru pembimbing khusus," Jakarta, Apr. 2024. Accessed: Jun. 13, 2025. [Online]. Available: https://jakarta.suaramerdeka.com/pendidikan/13412344129/sekolah-inklusi-di-indonesia-baru-148-persen-yang-memiliki-guru-pembimbing-khusus