Jurnal Matematika dan Statistika serta Aplikasinya, Vol. 13, No. 2, 41-53, 2025
DOI:https://doi.org/10.24252/msa.v13i2.57776

A Binary Logistic Regression Approach to Identifying Determinants of School Dropout among Child Labourers in Sulawesi

Nazlia Rosmini[1]*, Nofita Istiana[2]
[1] *Badan Pusat Statistik, Jakarta, Indonesia*
[2] *Politeknik Statistika STIS, Jakarta, Indonesia*
***Corresponding author**: nofita@stis.ac.id

***Submission date**: 06 July 2025, **Revision**: 20 August 2025, **Accepted**: 03 October 2025

ABSTRACT
Child labour remains a persistent challenge in Indonesia, particularly among children from economically disadvantaged families who are often forced to work at an early age. Balancing work and education create a significant risk of school dropout. This study aims to examine the determinants of school dropout among child labourers in Sulawesi, one of the regions with the highest prevalence of child labour. Using data from the March 2022 National Socio-Economic Survey (Susenas), a binary logistic regression model was employed to analyze the influence of socio-demographic and economic factors. The results show that 16.39% of working children in Sulawesi have discontinued formal education. Key predictors of school dropout include the child's gender, age, household poverty status, and the educational attainment of the household head. These findings highlight the urgent need for targeted educational and social policies to reduce school dropout among vulnerable working children in the region.

KEYWORDS
Child labour, School dropout, Binary logistic regression.

1. INTRODUCTION

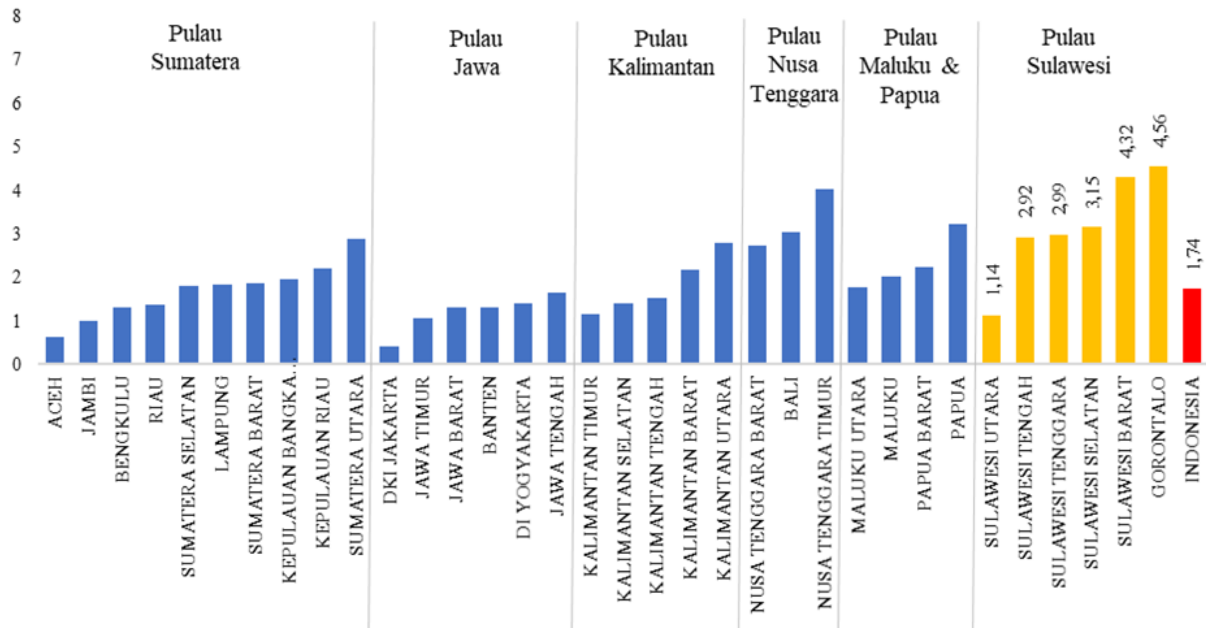
Children are the next generation whose role is essential in supporting future national development. Access to education is a fundamental right that should be guaranteed for every child. Quality education plays a crucial role in preparing competent human resources for Indonesia's progress. However, challenges such as school dropout continue to hinder the realization of children's right to education.

Efforts to reduce school dropout rates are in line with the fourth goal of the Sustainable Development Goals (SDGs), which emphasizes inclusive and equitable quality education for all. According to [1], access to quality education can reduce poverty and support sustainable economic development. Therefore, investing in children's education is one of the government's strategies to improve the quality of human capital.

Ideally, every child should have the opportunity to pursue education without barriers. In reality, various obstacles interfere with children's ability to complete their education—one of the most concerning being school dropout. Dropping out refers to the condition in which children leave school before completing the required learning stages [2]. According to [3], this results in students being unable to obtain a formal certificate of completion. In 2022, as many as 491,311 children in Indonesia dropped out of school at the beginning of the academic year. This includes 252,991 children who left school at the intermediate level and 238,320 children who did not continue to higher levels [4].

One of the major factors linked to school dropout is child labour. As stated by [5], child labour and school dropout have a causal relationship; when child labour increases, so does the rate of children leaving school, and vice versa. Working children are those who perform routine and time-consuming tasks, either for their parents or others, regardless of whether they receive compensation (Tjandraningsih, 2010: 83 in [6]). Child labour becomes a major factor in educational disengagement, often driven by financial pressures within the household. Families with limited economic resources tend to prioritize income over schooling, pushing children into work. These children frequently face challenges in balancing educational responsibilities with work demands, which results in physical and mental fatigue, decreased academic achievement, and ultimately diminished motivation to continue their education.

According to the International Labour Organization (ILO), child labour is defined as the involvement of children in production activities for at least one hour during the reference period [7]. In Indonesia, child labour is further categorized based on age and working hours under Law No. 13 Year 2003: children aged 5–12 who work without time limits, those aged 13–14 who work more than 15 hours per week, and those aged 15–17 who work over 40 hours per week. This study focuses specifically on children who meet these legal definitions of child labour, rather than children who simply help with minor or non-routine tasks. Data from the Ministry of Women's Empowerment and Child Protection (2022) show that 19 of 34 provinces



Source: Ministry of Women's Empowerment and Child Protection 2022

Figure 1. Percentage of child labour by province in Indonesia in 2022

had child labour rates above the national average (see **Figure 1**). Sulawesi Island recorded the highest incidence of child labour in Indonesia. Based on [8], Gorontalo Province had the highest percentage of child labourers at 4.56%, while DKI Jakarta had the lowest at 0.41%. Five out of six provinces on Sulawesi Island exceeded the national average of 1.74%, with Gorontalo and West Sulawesi recording the most significant rates.

Although the national child labour rate decreased slightly by 0.08% from 2021, the number of child labourers who were no longer attending school increased. In 2020, 15.83% of child labourers were out of school, rising to 16.32% in 2022. Sulawesi Island also has dropout rates above the national average across all school levels. Dropout rates at the primary (0.29%), junior secondary (0.28%), and senior secondary (0.34%) levels are higher than national figures, with the highest rate occurring at the senior secondary level. This suggests low public awareness of the importance of completing secondary education and highlights the need for government attention and policy focus.

One such policy is the amendment of Law No. 20 of 2003 concerning the national education system, which emphasizes the acceleration of the 12-Year Compulsory Education Programme. The programme aims to expand access to education at all levels (Ministry of Education and Culture Strategic Plan 2020–2024). Although this programme has been implemented since 2013, challenges persist—particularly high dropout rates in economically vulnerable regions.

Child labour remains the most significant factor causing school dropout [9]. The decision to leave school is influenced by various characteristics, including individual, social, and economic household factors. According to [10], five variables influence this phenomenon: gender, age, number of household members, education of the household head, and place of residence. Meanwhile, [11] identified six key factors: the household head's education level, ownership of an Indonesia Smart Card, number of household members, working status of the child, poverty level, and classification of residence (urban/rural). The incidence of school dropout among child labourers is a binary outcome, where a value of 1 indicates a dropout and a value of 0 indicates continued schooling. Thus, binary logistic regression is the appropriate method to analyse this type of dependent variable.

However, despite a growing body of research at the national level, there is still limited empirical evidence focusing specifically on child labourers in Sulawesi Island—an area with consistently high rates of both child labour and school dropout. Few studies have applied statistical modelling such as binary logistic regression to isolate and measure the influence of socio-economic and demographic factors on school dropout among child labourers at the regional level. This creates a critical research gap that needs to be addressed to support region-specific educational policies and interventions.

Therefore, this study aims to apply a binary logistic regression model to the March 2022 Susenas data to analyse the incidence of school dropout among child labourers on Sulawesi Island. The model will assess the influence of variables such as gender, age, poverty status, and the educational background of the household head on the probability of children dropping out of school.

2. LITERATURE REVIEW

2.1 Child Labour

As stated in [7], work is defined as a wage-generating activity involving at least one hour of labour within the previous week. In 2022, the Central Bureau of Statistics (BPS), UNICEF, Bappenas, and the International Labour Organization (ILO) categorized child labourers aged 5 to 17 in accordance with Law No. 13 of 2003. Under this regulation, children aged 5–12 are classified as child labourers regardless of the number of working hours; those aged 13–14 fall into this category if they work more than 15 hours per week; and individuals aged 15–17 are considered child labourers if their working hours exceed 40 per week. The term "child labour" typically refers to the employment of minors and is often associated with exploitative practices, such as low wages and inadequate legal protection. Tjandraningsih (2010), as cited in [6], describes child labour as the regular involvement of children in work activities, whether for their own family or for others, with or without financial compensation. Additionally, [12] defines child labour as individuals between the ages of 5 and 17 who are engaged in one or more forms of work over a certain duration.

2.2 School Dropout

School dropout refers to students who discontinue their education before completing a particular level of schooling. Gunawan (2010:71), as cited in [13], similarly defines dropout as a status assigned to students who are unable to complete their education at a given educational stage. According to [14], there are two perspectives in defining school dropout: first, students who leave school before earning a diploma; and second, graduates who do not proceed to the next level of education. In this study, the term "school dropout" encompasses both students who withdraw during a level of schooling and those who do not transition to a higher level within the 12-year education system.

The factors influencing children's participation in school can be categorized into internal and external factors. Internal factors include students' abilities, interests, motivation, attitudes, aspirations, and their perceptions of school. External factors consist of parental economic status, views on education, distance to school, student–teacher relationships, and government support [15]. Meanwhile, [16] classifies the causes of school dropout into two broad perspectives: individual and institutional. Individual factors cover aspects such as demographics, attitudes or behaviors, and students' educational backgrounds. Institutional factors refer to the influence of family conditions, community environment, and school-related circumstances.

2.3 Binary Logistic Regression

Binary logistic regression is a statistical method used to model the relationship between a binary (two-category) dependent variable and one or more independent variables. The relationship between the independent variable(s) (X) and dependent variable (Y) is expressed as $E(Y|x)$ which represents the expected value of Y, denoted by $\pi(x)$. This expected value, ranging from 0 to 1, reflects the probability of a particular outcome. In this context, Y is a binary response variable taking values of either 0 or 1. The value $Y = 1$ indicates a successful event with probability $P(Y = 1 \mid x)$, while $Y = 0$ represents failure, with a complementary probability of $1 - P(Y = 1 \mid x)$.

The mathematical form of the binary logistic regression model involving independent variables is

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_j x_j)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_j x_j)} \tag{1}$$

where $\pi(x)$ denotes the probability of a successful outcome, β_0 represents the intercept term, and β_j corresponds to the coefficient associated with the j^{th} independent variable, where $j = 1, 2, \ldots, p$. The term x_j indicates the value of the j^{th} independent variable, and p stands for the total number of independent variables included in the model.

Binary logistic regression is a form of the Generalized Linear Model (GLM), which requires a link function to transform the nonlinear probability function $\pi(x)$ into a linear predictor. Since the logistic function is inherently nonlinear, this transformation is essential for modeling. In the case of binary logistic regression, the link function used is the logit function, which connects the binary outcome probability to a linear combination of predictors. This transformation enables the modeling of binary event probabilities as a function of the independent variables. Thus, to linearize the model, the logit transformation of $\pi(x)$ is applied. The resulting logit model for binary logistic regression is as follows

$$\begin{aligned}
logit\,(\pi) = g(x) &= ln\left(\frac{\pi(x)}{1-\pi(x)}\right) \\
g(x) &= ln\left(\frac{\frac{\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)}{1+\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)}}{1-\frac{\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)}{1+\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)}}\right) \\
g(x) &= ln\left(\frac{\frac{\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)}{1+\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)}}{\frac{1}{1+\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)}}\right) \\
g(x) &= ln\left(\frac{\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)}{1}\right) \\
g(x) &= ln\left(\exp\left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)\right) \\
g(x) &= \left(\beta_0+\beta_1x_1+\ldots+\beta_jx_j\right)
\end{aligned} \tag{2}$$

Following the transformation, the equation is expressed in a linear form. The value of the function $g(x)$ extends over the entire real number line, ranging from negative infinity to positive infinity, contingent upon the values of the predictor variables X.

2.4 Parameter Estimation

In binary logistic regression, parameter estimation is performed using the Maximum Likelihood Estimation (MLE) method. The initial step involves constructing the likelihood function, which represents the probability of the observed data expressed as a function of the unknown parameters. Consequently, the estimator derived is the one that best fits the observed data [17] .

For the binary logistic regression model, where the dependent variable Y takes values 0 or 1, the function $\pi(x)$ denotes the conditional probability $P(Y = 1|x)$, while $1 - \pi(x)$ corresponds to the conditional probability $P(Y = 0|x)$. For an observation(x_i, y_i), if $y_i = 1$, the contribution to the likelihood function is $\pi(x_i)$; if $y_i = 0$, the contribution becomes $1 - \pi(x_i)$, where $\pi(x_i)$ represents the value of $\pi(x)$ evaluated at $= x_i$. Therefore, the probability mass function for the pair (x_i, y_i) is expressed as:

$$\pi(x_i)^{y_i}(1-\pi(x_i))^{1-y_i} \tag{3}$$

Assuming that each observation is independent, the likelihood function for the entire dataset is given by

$$L(\beta) = \prod_{i=1}^{n} \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \tag{4}$$

According to the principle of maximum likelihood, the parameter vector β is estimated by maximizing the likelihood function in **Equation** (4). Mathematically, this is often achieved by maximizing the natural logarithm of the likelihood function, expressed as follows

$$\begin{aligned} l(\beta) = \ln(L(\beta)) = ln\left(\prod_{i=1}^{n} \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i}\right) \\ &= \textstyle\sum_{i=1}^{n} \left(\ln\left[\pi(x_i)^{yi}\right][1 - \pi(x_i)]^{1-yi}\right) \\ &= \textstyle\sum_{i=1}^{n} \left(\ln\left[\pi(x_i)^{yi}\right] + ln[1 - \pi(x_i)]^{1-yi}\right) \\ &= \textstyle\sum_{i=1}^{n} \{y_i \ln[\pi(x_i)] + (1 - y_i)\ln[1 - \pi(x_i)]\} \\ &= \textstyle\sum_{i=1}^{n} y_i \ln\left[\dfrac{\pi(x_i)}{1 - \pi(x_i)}\right] + \sum_{i=1}^{n} \ln[1 - \pi(x_i)] \end{aligned}$$

$$l(\beta) = y_i\left(\beta_0 + \textstyle\sum_{i=1}^{n} \beta_i x_{ij}\right) - \sum_{i=1}^{n} \ln\left[1 + exp\left(\beta_0 + \sum_{i=1}^{n} \beta_i x_{ij}\right)\right] \tag{5}$$

$$\frac{\partial(l(\beta))}{\partial(\beta_0)} = \sum_{i=1}^{n} (y_i - \pi(x_i)) = 0 \tag{6}$$

and

$$\frac{\partial(l(\beta))}{\partial(\beta_j)} = \sum_{i=1}^{n} x_{ij}(y_i - \pi(x_i)) = 0 \tag{7}$$

where $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$. Let n denote the total number of observations, and j the number of estimated parameters.

The values of β that maximize the likelihood function $l(\beta)$ are obtained by taking the derivative of $l(\beta)$ in **Equation** (5) with respect to the parameters β_0 and β_j, for $j = 1, 2, ..., p$. These derivatives are then set equal to zero, as shown in **Equation** (6) and **Equation** (7).

Since the likelihood equations generally do not have closed-form solutions—particularly in models involving nonlinear link functions, such as logistic regression—the estimation of the parameters β is typically performed using iterative numerical methods. One of the most commonly employed techniques is the Newton-Raphson algorithm, which updates parameter estimates at each iteration based on the first and second derivatives (i.e., the gradient and Hessian) of the log-likelihood function.

2.5 Goodness-of-Fit Test Model

The goodness-of-fit test is employed to evaluate the extent to which a statistical model adequately represents the observed data, particularly in its ability to explain the dependent variable. A model is considered a good fit when the discrepancy between predicted values and observed outcomes is minimal. One commonly used method for this purpose in logistic regression is the Hosmer-Lemeshow test. The hypotheses for this test are defined as follows:
H_0: The model fits the data well.
H_1: The model does not fit the data well.
The test statistic is computed using the formula

$$\hat{C} = \sum_{k=1}^{g} \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \sim \chi^2_{(g-2)} \tag{8}$$

In this formula, $\hat{C}$ represents the Hosmer-Lemeshow test statistic. The number of groups formed from the dataset is denoted by g. For each group k, n'_k is the number of observations in that group, and O_k is the total number of observed responses, calculated as the sum of observed outcome values y_j for all subjects within the group, i.e. $O_k = \sum_{j=1}^{c_k} y_j$ where c_k is the number of distinct combinations of predictor variables in the k^{th} group.

The average estimated probability of success in group k is denoted by $\bar{\pi}_k$, computed as $\bar{\pi}_k = \sum_{j=1}^{c_k} \frac{m_j \bar{\pi}_j}{n_k}$.Here, m_j represents the number of individuals with the j^{th} combination of explanatory variables, and $\bar{\pi}_j$ is the estimated probability of success for that combination. The term n_k is the total number of subjects in group k.

The Hosmer-Lemeshow test follows a Chi-square distribution with $g-2$ degrees of freedom. If the calculated test statistic $\hat{C}$ exceeds the critical value of the Chi-square distribution at a given significance level α, or if the corresponding p-value is less than α, then the null hypothesis is rejected, indicating that the model does not fit the data well. Ideally, the test should result in a failure to reject H_0, which would suggest that the model provides an adequate fit for the data and is appropriate for explaining the dependent variable.

2.6 Significance Testing Model

2.6.1 Simultaneous Significance Test

The simultaneous significance test is a statistical procedure used to evaluate whether, collectively, the independent variables in a model have a significant effect on the dependent variable. This test is essential for determining the overall explanatory power of the model. The test is conducted using the Likelihood Ratio Test (LRT), which compares a full model (including independent variables) with a reduced model (excluding them). The hypotheses for the test are formulated as follows:
H_0: $\beta_1 = \beta_2 = \ldots = \beta_j = 0$
H_1: At least one $\beta_j \neq 0; j = 1,\ 2,\ \ldots,\ p$
The likelihood ratio test statistic, denoted by GGG, is calculated using the following formula

$$G = -2ln\left[\frac{l_0}{l_p}\right]\ \chi^2_{(p)} \tag{9}$$

where:
G is the likelihood ratio test statistic
p represents the degrees of freedom, equal to the number of independent variables in the model
l_0 is the likelihood value of the reduced model (i.e., the model without independent variables)
l_p is the likelihood value of the full model (i.e., the model including the independent variables).

The test statistic G follows a Chi-square distribution with p degrees of freedom. The null hypothesis H_0 is rejected if $G >\ \chi^2_{(\alpha;p)}2$, or equivalently, if the p-value is less than the significance level α. A rejection of the null hypothesis implies that, simultaneously, there is sufficient evidence to conclude that at least one independent variable has a statistically significant effect on the dependent variable. Thus, the likelihood ratio test serves as a global assessment of model significance.

2.6.2 Partial Significance Test

The partial significance test is a statistical method used to assess the individual contribution of each independent variable to the dependent variable within a regression model. This test aims to determine whether each independent variable, considered separately, has a statistically significant effect on the response variable. The hypotheses for the test are defined as follows:
H_0: $\beta_j = 0$ H_1: $\beta_j \neq 0$ The test is carried out using the Wald statistic, which is calculated by taking the square of the ratio between the estimated coefficient β_j and its standard error $s(\beta_j)$. Formally, the Wald statistic W_j is defined as

$$W_j = \left(\frac{\beta_j}{s(\beta_j)}\right)^2 \tag{10}$$

and it follows a Chi-square distribution with one degree of freedom. The standard error measures the variability or uncertainty around the estimated coefficient, and by comparing the magnitude of the coefficient relative to its standard error, the Wald test assesses the significance of the variable.

If the computed Wald statistic exceeds the critical value from the Chi-square distribution at the chosen significance level α, or if the corresponding p-value is less than α, then the null hypothesis is rejected. This indicates that the independent variable j has a statistically significant partial effect on the dependent variable. Conversely, if the test statistic does not exceed the critical threshold, the null hypothesis cannot be rejected, implying that the variable does not contribute significantly when considered individually.

2.7 Odd Ratio

The odds ratio (OR) is a statistical measure used to quantify the strength or tendency of the association between two variables. It compares the odds of success (when the variable $x = 1$) to the odds of failure (when $x = 0$). Specifically, the odds for category one (x=1) is expressed as $\frac{\pi(1)}{[1-\pi(1)]}$, while the odds for the reference category ($x = 0$) is $\frac{\pi(0)}{[1-\pi(0)]}$. A larger value of the parameter β_j indicates a stronger influence of the independent variable on the dependent variable. The ratio of these odds between categories is known as the odds ratio, which can be formulated as:

$$\begin{aligned} OR &= \frac{\frac{\pi(1)}{(1-\pi(1))}}{\frac{\pi(0)}{(1-\pi(0))}} \\ &= \frac{\frac{e^{\beta_0+\beta_1}/1+e^{\beta_0+\beta_1}}{1/1+e^{\beta_0+\beta_1}}}{\frac{e^{\beta_0}/1+e^{\beta_0}}{1/1+e^{\beta_0}}} \\ &= \frac{e^{\beta_0+\beta_1}}{e^{\beta_0}} \\ &= e^{(\beta_0+\beta_1)-\beta_0} \qquad = e^{\beta_1} \end{aligned} \tag{11}$$

Thus, the relationship between the odds ratio and the regression coefficient β_j can be succinctly written as:

$$OR = e^{\beta_j} \tag{12}$$

Interpreting the odds ratio: if $OR < 1$, this indicates a negative association between the independent variable and the dependent variable, meaning that an increase in the independent variable decreases the odds of the outcome. Conversely, if $OR > 1$, there is a positive association, meaning an increase in the independent variable increases the odds of the outcome. An odds ratio equal to 1 implies no association between the variables.

3. METHODOLOGY

3.1 Data Source and Sample Characteristics

The data used in this study were obtained from the March 2022 National Socioeconomic Survey (Susenas), conducted by Statistics Indonesia (Badan Pusat Statistik, BPS). The unit of analysis was child laborers aged 5–17 years residing in the Sulawesi region. A total of 800 observations were included in the analysis.

3.2 Analytical Approach

The analytical method was divided into two components: descriptive and inferential analysis. The descriptive analysis aimed to provide a general overview, while the inferential analysis sought to examine factors influencing school dropout among child laborers in Sulawesi in 2022.

Table 1. Independent Variables in the Study

Notation	Variable Name	Category
X_1	Child's Gender	Female*
		Male
X_2	Child's Age	Numerical
X_3	Poverty Status	Non-poor*
		Poor
X_4	Head of Household Education	High School Graduate*
		Did Not Graduate
X_5	Head of Household Employment Status	Employed*
		Unemployed
X_6	Area Classification	Urban*
		Rural

*Note: * indicates reference categories.*

3.3 Analytical Procedure

This study employed binary logistic regression to analyze the factors associated with school dropout among child laborers. Binary logistic regression is appropriate when the dependent variable is binary—in this case, dropout status (1 = dropped out, 0 = did not drop out). The model estimates the probability that a certain event occurs (e.g., dropping out of school) as a function of a set of independent variables. The probability of success, denoted as $\pi(x)$, is transformed using the logit link function to linearize the relationship between predictors and the log odds of the event occurring. The logistic regression model used in this study is expressed as

$$g(x) = ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 \tag{13}$$

where $\pi(x)$ is the probability that the child drops out of school, X_1 to X_6 are the independent variables defined in **Table 1**, β_0 is the intercept term, β_j are the coefficients for each explanatory variable.

The inferential analysis followed these steps:

1. Model Construction. A binary logistic regression model was constructed to examine the relationship between the independent variables and school dropout status among child laborers in Sulawesi.

2. Parameter Estimation. The parameters of the model were estimated using the Maximum Likelihood Estimation (MLE) method. This approach identifies the set of parameter values that maximizes the probability of observing the data. The log-likelihood function was derived and optimized numerically using iterative techniques.

3. Goodness-of-Fit Test. The Hosmer-Lemeshow test was used to evaluate the model's goodness of fit. This test assesses whether the observed event rates match expected event rates in subgroups of the data. A non-significant result (p-value) suggests that the model adequately fits the data.

4. Simultaneous Significance Test. A Likelihood Ratio Test (LRT) was conducted to assess the joint significance of all independent variables in the model. The null hypothesis stated that all coefficients were equal to zero. Rejection of the null indicated that the model provides a better fit than a null (intercept-only) model.

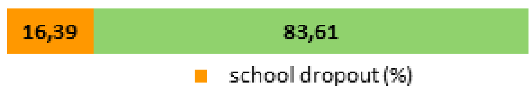
5. Partial Significance Test. To evaluate the contribution of individual variables, the Wald test was used. Each regression coefficient was tested to determine whether it differed significantly from zero, with results evaluated based on the Wald statistic and associated p-values.

6. Odds Ratio Calculation. For interpretation purposes, the odds ratios (OR) were computed by exponentiating each regression coefficient ($OR = e^{\beta_j}$). These values indicate how the odds of school dropout change with a one-unit increase in the corresponding independent variable, holding other variables constant.

4. RESULT & DISCUSSION

Child labor is subject to legal limitations regarding both age and working hours, as it increases the risk of school dropout. This issue remains prevalent in Indonesia and serves as a significant indicator of the quality of education. School dropout rates reflect the number of children who discontinue their education, underscoring the importance of safeguarding children's rights, particularly the right to education.

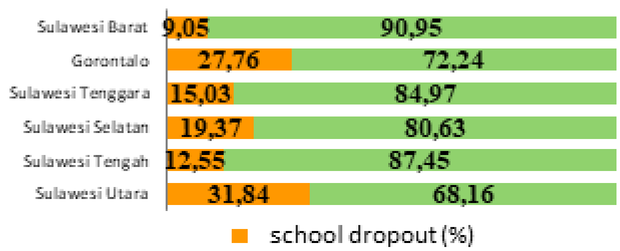
Based on processed data from the March 2022 National Socio-Economic Survey (Susenas), a total of 16,807 children aged 5 to 17 years were sampled. Among these children on Sulawesi Island, 4.75 percent (800 children) were identified as child laborers, while the remaining 95.25 percent (16,007 children) were not engaged in child labor. From this group of child laborers, the incidence of school dropout was analyzed. A significant proportion of child laborers on Sulawesi Island were found to have



Source: Susenas, March 2022 (processed data)

Figure 2. Percentage of School Dropout among Child Laborers on Sulawesi Island in 2022

dropped out of school (**Figure 2**). The data indicated that 16.39 percent of child laborers had discontinued their education, whereas 83.61 percent remained in school. These dropout cases were not concentrated in a single province but were dispersed across all provinces on Sulawesi Island (**Figure 3**). Sulawesi Barat recorded the lowest proportion of school dropouts among



Source: Susenas, March 2022 (processed data)

Figure 3. Percentage of School Dropout among Child Laborers by Province on Sulawesi Island in 2022

child laborers at 9.05 percent—making it the only province with a rate below 10 percent. In contrast, Sulawesi Utara had the highest dropout rate at 31.84 percent, followed by Gorontalo, Sulawesi Selatan, Sulawesi Tenggara, and Sulawesi Tengah, respectively.

The child laborers who dropped out of school were more likely to be male (24.53%), older in age, and from non-poor households (16.52%). Additionally, they typically lived in households where the head had not completed senior high school (19.65%), where the head was employed (22.20%), and resided in rural areas (17.29%) (**Table 2**). To examine the factors influencing the dropout status of child laborers on Sulawesi Island, a suitable inferential analysis was conducted using binary logistic regression. The model parameters were estimated using the Maximum Likelihood Estimation (MLE) method, yielding the following equation

$$\hat{g}(x) = -14,138 + 1,446X_1 + 0,803X_2 + 0,761X_3 + 1,063X_4 + 0,074X_5 + 0,114X_6.$$

A goodness of fit test was subsequently performed to evaluate whether the model adequately represented the data concerning school dropout among child laborers on Sulawesi Island in 2022. The Hosmer-Lemeshow test was employed for this purpose. The model was deemed a good fit if the test failed to reject the null hypothesis. The results are presented in **Table 3**. The chi-square statistic of 13.145 was lower than the critical value of $\chi^2_{(8)} = 15,507$, and the corresponding p-value of 0.107 exceeded the significance level $\alpha = 0,05$ (**Table 3**). Therefore, the decision was to fail to reject the null hypothesis, indicating that the logistic regression model was an appropriate fit to explain the school dropout status of child laborers in Sulawesi.

Table 2. Overview of School Dropout among Child Laborers on Sulawesi Island in 2022

Variable	Category	Still in School (%)	School Dropout (%)
Child's Gender	Female	96.24	3.76
	Male	75.47	24.53
Child's Age	5–12	99.11	0.89
	13–14	51.75	48.25
	15–17	25.38	74.62
Poverty Status	Non-poor	83.48	16.52
	Poor	84.14	15.86
Head of Household Education	Completed Senior High School	91.11	8.89
	Did Not Complete Senior High School	80.35	19.65
Head of Household Employment Status	Employed	77.80	22.20
	Unemployed	93.95	6.05
Area Classification	Urban	85.35	14.65
	Rural	82.71	17.29

Source: Susenas, March 2022 (processed data).

Table 3. Hosmer–Lemeshow Goodness of Fit Test

Chi-square	df	p-value
13.145	8	0.107

Source: Susenas, March 2022 (processed data).

A likelihood ratio test was then conducted to assess whether at least one of the independent variables significantly influenced school dropout. The *G*-statistic of 404.910 exceeded the critical value of $\chi^2_{(6)} = 12,592$, and the p-value of 0.000 was below the 0.05 threshold (**Table 4**). As a result, the null hypothesis was rejected, confirming that at least one independent variable had a statistically significant effect on the dropout status of child laborers. Following the significant result in the overall model test,

Table 4. Simultaneous Significance Test

Chi-square	df	p-value
404.910	6	0.000

Source: Susenas, March 2022 (processed data).

a partial significance test was performed using the Wald statistic to identify which specific independent variables significantly influenced dropout status. Four out of six independent variables showed p-values below 0.05, indicating statistical significance. These variables were child's gender, age, poverty status, and the education level of the household head. In contrast, household head's employment status and area classification did not exhibit significant effects on school dropout among child laborers in Sulawesi.

The gender of the child was found to significantly influence dropout status. The odds ratio for male child laborers was exp(1.446) = 4.245, suggesting that boys were 4.245 times more likely to drop out of school compared to girls. This result is consistent with findings by [18] and [19], which also highlighted the influence of gender on school dropout.

Age was another significant predictor. The odds ratio for this variable was exp(0.803) = 2.232, indicating that each additional year of age increased the probability of dropping out by 2.232 times. This aligns with [20] and [21], which noted that as children grow older, they are more likely to leave school, often to support household work or income.

Poverty status also had a statistically significant impact. The odds ratio of 2.140 implied that children from poor households were 2.14 times more likely to drop out than those from non-poor households. This result corroborates the findings of [11], which emphasized the adverse effect of poverty on educational attainment.

The education level of the household head (KRT) was similarly significant. Children whose household heads had not

Table 5. Wald Test Results (Partial Significance Test)

Variable	Coefficient	S.E.	Wald	P-value	Odds Ratio
Intercept	-14.138	1.209	136.679	0.000	0.000
Child's Gender					
Female (ref)					
Male	1.446	0.380	14.480	0.000*	4.245
Child's Age	0.803	0.072	125.970	0.000*	2.232
Poverty Status					
Non-poor (ref)					
Poor	0.761	0.367	4.301	0.038*	2.140
Head of Household Education					
Completed Senior High School (ref)	1.063	0.447	5.660	0.017*	2.894
Did Not Complete Senior High School					
Head of Household Employment Status					
Employed (ref)					
Unemployed	0.074	0.402	0.033	0.855	1.076
Area Classification					
Urban (ref)					
Rural	0.114	0.364	0.097	0.755	1.120

Source: Susenas, March 2022 (processed data).
*Note: (ref) denotes reference category; * indicates significance at* $\alpha = 0.05$.

completed senior high school were 2.894 times more likely to drop out than those whose heads had completed at least that level. This suggests that lower parental education correlates with reduced awareness of the importance of education. Studies by [11] and [22] support this finding, indicating that higher parental education is associated with better school participation outcomes for children.

Conversely, the employment status of the household head did not significantly influence dropout status. This occurs because the employment status of the household head only indicates whether they are engaged in some form of work. However, this status does not adequately reflect the adequacy, stability, or quality of the income generated. Consequently, even if the head of the household is employed, this factor alone may not significantly influence a working child's decision to drop out of school.

Lastly, area classification (urban vs. rural) was also not a significant factor. This may be due to the flexible nature of rural child labor, such as helping on family farms, which often occurs outside of school hours. Thus, even though rural children may engage in work, they may still be able to attend school.

5. CONCLUSION

The analysis identified four variables that significantly influenced school dropout among child laborers in Sulawesi Island in 2022: the child's gender, age, household poverty status, and the educational attainment of the household head. The probability of school dropout was found to be higher among male children, those of older age, children from economically disadvantaged households, and those whose household heads had not completed senior high school. These findings underscore the multifaceted nature of school dropout, which is not only shaped by individual characteristics but also by household socioeconomic conditions and parental educational background.

Given these results, it is imperative that policymakers develop targeted interventions aimed at mitigating the risk of school dropout among vulnerable children. One strategic approach would involve providing inclusive education and vocational training programs tailored for both boys and girls, designed to accommodate their working schedules without compromising their educational participation. Furthermore, there is a critical need to enhance monitoring mechanisms for child labor, especially in cases where children are engaged in physically demanding or hazardous work, to ensure that such labor does not hinder their right to education.

In addition, efforts should be made to reduce the financial burden associated with secondary and higher education, as cost remains a significant barrier for children from low-income families. Reducing or subsidizing educational expenses could improve school retention, particularly among at-risk populations. Equally important is the need to raise parental awareness regarding the long-term benefits of education. Outreach programs targeting parents with low educational attainment could foster greater support for their children's schooling and counteract negative perceptions about formal education.

For future research, it is recommended to expand the analytical framework by incorporating additional variables that may affect school attendance and dropout, such as school infrastructure, child health conditions, educational aspirations, parental income levels, and access to transportation. Utilizing primary data collection methods can provide more detailed insights into the experiences of child laborers. Methods such as interviews or focus group discussions help reveal structural and personal barriers that secondary data often overlook. This approach allows for a deeper understanding of the challenges children face in continuing their education and supports the creation of more effective interventions.

REFERENCES

[1] N. G. Mankiw, D. Romer, and D. N. Weil, "A Contribution to the Empirics of Economic Growth," *The Quarterly Journal of Economics*, vol. 107, no. 2, pp. 407–437, May 1992. [Online]. Available: https://academic.oup.com/qje/article-lookup/doi/10.2307/2118477

[2] V. Vebrianto, "Pengaruh Keluarga Dalam Pendidikan Anak," Jaya Sakti, 1995.

[3] W. N. Utami and A. Rosyid, "Identifikasi faktor penyebab siswa putus sekolah di tingkat sekolah dasar wilayah duri kepa," in *Prosiding Seminar dan Diskusi Nasional Pendidikan Dasar*, 2020. [Online]. Available: https://digilib.esaunggul.ac.id/identifikasi-faktor-penyebab-siswa-putus-sekolah-di-tingkat-sekolah-dasar-wilayah-duri-kepa-20850.html

[4] W. Wandi, "Banyak Anak Putus Sekolah, Ketua DPR Ingatkan Pemerintah Wajib Beri Akses Pendidikan," *InfoPublik*, 2023.

[5] R. D. Aisy and F. Nailufar, "Pengaruh Tingkat Kemiskinan Dan Angka Putus Sekolah terhadap Pekerja Anak di Indonesia," *j. ekonomika indonesia*, vol. 11, no. 1, p. 21, Jun. 2022. [Online]. Available: https://ojs.unimal.ac.id/ekonomika/article/view/7727

[6] M. B. Rustam, "Analisis Faktor-Faktor yang Mempengaruhi Status Sekolah Pekerja Anak Usia Wajib Belajar 9 Tahun di Kota Makassar," Ph.D. dissertation, Universitas Muhammadiyah Makassar, Makassar, 2020. [Online]. Available: https://monograph.unismuh.ac.id/dokumen/detail/10501/

[7] D. S. K. dan Ketenagakakerjaan, "Booklet Pekerja Anak di Indonesia 2022 Sebelum dan Semasa Pandemi COVID-19," Badan Pusat Statistik, Booklet 04100.2302, 2022.

[8] K. P. P. dan Perlindungan Anak, "Persentase pekerja anak di Indonesia 2022," Kementerian Pemberdayaan Perempuan dan Perlindungan Anak, Tech. Rep.

[9] Y. R. I. Asmara and I. W. Sukadana, "Mengapa Angka Putus Sekolah Masih Tinggi? (Studi Kasus Kabupaten Buleleng Bali)." *Jurnal Harian Regional*, vol. 5, no. 12, pp. 1347–1585, 2016.

[10] N. D. Sinaga, R. Rais, and H. Sain, "Model Regresi Logistik Biner Untuk Menentukan Faktor Yang Berpengaruh Terhadap Anak Putus Sekolah Di Sulawesi Tengah," *JIMT*, vol. 13, no. 1, Dec. 2016. [Online]. Available: https://bestjournal.untad.ac.id/index.php/JIMT/article/view/7494

[11] A. Hakim, "Faktor Penyebab Anak Putus Sekolah," *j. pendidik.*, vol. 21, no. 2, pp. 122–132, Sep. 2020. [Online]. Available: http://jurnal.ut.ac.id/index.php/jp/article/view/907

[12] Badan Pusat Statistik, "Indikator Pekerjaan Layak di Indonesia Tahun 2020-2022," BPS RI, 2022.

[13] A. Y. Budi Lestari, F. Kurniawan, and R. Bayu Ardi, “Penyebab Tingginya Angka Anak Putus Sekolah Jenjang Sekolah Dasar (SD),” *JISD*, vol. 4, no. 2, p. 299, Jun. 2020. [Online]. Available: https://ejournal.undiksha.ac.id/index.php/JISD/article/view/24470

[14] M. Dahlan, “Problematika Putus Sekolah Dan Pengangguran (Analisis Sosial Pendidikan),” Aug. 2019. [Online]. Available: https://osf.io/h6vp7

[15] Sopyan, A. D. M. P. Agustang, and A. Agustang, “Dinamika Putus Sekolah di Kalangan Pekerja Anak di Kota Makassar,” 2022, publisher: Unpublished. [Online]. Available: https://rgdoi.net/10.13140/RG.2.2.18768.79361

[16] R. W. Rumberger, “Why Students Drop Out of School and What Can be Done,” 2001.

[17] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*, 1st ed. Wiley, Sep. 2000. [Online]. Available: https://onlinelibrary.wiley.com/doi/book/10.1002/0471722146

[18] T. A. Perhati, I. Indahwati, and B. Susetyo, “Identifikasi Karakteristik Anak Putus Sekolah Di Jawa Barat Dengan Regresi Logistik,” *Indonesian Journal of Statistics and Applications*, vol. 1, no. 1, pp. 56–65, 2017.

[19] D. Cahyawati, O. Dwipurwani, and R. Ruswanti, “Penentuan Peluang Anak Putus Sekolah Pendidikan Dasar di Kabupaten Ogan Ilir mengggunakan Regresi Logistik Biner (Studi Kasus pada Kelompok Masyarakat Miskin),” *jmi*, vol. 10, no. 2, p. 69, Nov. 2016. [Online]. Available: http://jurnal.unpad.ac.id/jmi/article/view/10249

[20] R. A. Satti and M. Jamil, “Socio-Economic Determinants of School Dropouts: An Evidence from Households in Pakistan,” *iRASD J. Eco.*, vol. 3, no. 3, Dec. 2021. [Online]. Available: https://journals.internationalrasd.org/index.php/joe/article/view/449

[21] A. P. Arum and M. R. Adry, “Analisis Determinan Partisipasi Sekolah di Provinsi Jambi,” *Ecosains*, vol. 9, no. 1, p. 10, May 2020. [Online]. Available: http://ejournal.unp.ac.id/index.php/ekosains/article/view/115495

[22] Y. Setiadi, “Faktor-Faktor yang Mempengaruhi Putus Sekolah Jenjang SMA/Sederajat di Provinsi DKI Jakarta Tahun 2019,” *PSPMM*, vol. 3, May 2021. [Online]. Available: http://prosiding.himatikauny.org/index.php/prosidinglsm/article/view/144