The Nonlinear Impact of Education, Capital, and Labor on Regional Income Distribution in Indonesia

Sanusi Fattah*

Universitas Hasanuddin, Makassar, Indonesia
*sanusi_fattah@fe.unhas.ac.id

Abstract. This study explores how education, physical capital, and labor interact in shaping income inequality across Indonesia's 34 provinces between 2004 and 2023. Traditional linear models often miss the nuances here—they tend to assume that these factors affect all regions in the same way, regardless of their stage of development. To dig deeper, this research applies a Generalized Additive Model (GAM), which captures nonlinear patterns and thresholds that change depending on regional conditions. The results paint an interesting picture: primary, junior, and senior high school education consistently help narrow inequality, while tertiary education follows an inverted U-shaped trend. In other words, higher education can initially widen income gaps before eventually helping to reduce them as more graduates enter the workforce. Physical capital shows diminishing equalizing power—it's most effective in capital-scarce regions and less so as investment levels grow. Labor, on the other hand, forms a U-shaped relationship with inequality: it reduces disparities up to a point, after which its benefits start to taper off once regions hit their absorptive limits. Overall, the GAM model performs notably better than standard linear fixed-effect approaches, explaining 82.4% of the deviance compared to 68.1%. These findings highlight the need for more tailored policies that reflect each region's stage of economic maturity—focusing on education type, capital accumulation, and labor development in a way that fits local realities.

Keywords: Education, Capital, Labor, Inequality, Indonesia, GAM

1 Introduction

Despite decades of steady economic growth and fiscal decentralization, Indonesia still faces persistent gaps in regional income distribution. The central government has rolled out large-scale fiscal transfers and numerous development programs to bridge these divides, yet the differences between provinces—especially between the western and eastern regions—remain striking. These enduring disparities threaten the country's goal of inclusive growth and sustainable development, pointing to deeper structural issues in how investment, education, and labor productivity are spread across regions.

The uneven income distribution isn't just an economic matter—it also carries political and social weight. Provinces with lower income levels often struggle with weaker public services, limited infrastructure, and a narrow industrial base. This combination

M. Nohong et al. (eds.), *Proceedings of the 10th International Conference on Accounting, Management, and Economics (10th ICAME 2025)*, Advances in Economics, Business and Management Research 388,
https://doi.org/10.2991/978-94-6239-709-5_85

can trap them in a cycle of underdevelopment, making it difficult to catch up. While policies such as inter-regional transfers and infrastructure spending have aimed to close the gap, their success has been mixed. Differences in absorptive capacity, workforce skills, and local economic structures continue to limit their impact.

Scholars have long emphasized the importance of human capital, physical capital, and labor in explaining income disparities across regions [1, 2]. Becker's classic human capital theory [2] places education at the center of productivity and income growth, while studies by Barro [1] and Mankiw, Romer, and Weil [3] highlight how both human and physical capital accumulation foster regional convergence. Still, when it comes to Indonesia—where regional diversity is vast—the relationship between these factors and inequality is far from straightforward.

Much of the earlier research on Indonesia relies on linear econometric models, which assume that education, capital, and labor affect inequality in the same way across time and place [4, 5]. That's a big simplification. In reality, the influence of these variables likely shifts as regions evolve. For example, tertiary education might initially widen income gaps—since access is usually concentrated among wealthier groups—but as more people graduate and skills spread, it can begin to narrow inequality [6]. Similarly, capital accumulation may help equalize incomes in less developed provinces but lose its power as economies mature [7].

Labor dynamics tell a similar story. In areas where labor is scarce, higher participation rates can boost productivity and reduce inequality. But in provinces where there's already an oversupply of workers, job creation often can't keep up—leading to rising wage gaps and more informal employment [8]. These complex patterns suggest that a more flexible, nonlinear approach is needed to truly capture the dynamics at play.

To tackle this, the study applies the Generalized Additive Model (GAM) developed by Hastie and Tibshirani [9]. Unlike traditional models, the GAM doesn't assume a fixed relationship between variables—it lets the data shape the curves, revealing thresholds and nonlinear effects that would otherwise go unnoticed. This is particularly useful in Indonesia's context, where economic maturity, education, and capital resources vary so widely across provinces.

The dataset includes 34 provinces from 2004 to 2023, totaling 680 observations. The Williamson Index serves as the measure of interprovincial income inequality [10]. Independent variables include education (broken down into primary, junior high, senior high, and tertiary levels), gross fixed capital formation (PMTB) as a measure of physical capital, and labor force participation. All variables are log-transformed to stabilize variance and allow for elasticity-based interpretation.

This research pursues three main goals. First, it examines how different levels of education affect income inequality in nonlinear ways. Second, it explores threshold effects for capital and labor—pinpointing where they help reduce or inadvertently widen inequality. Finally, it offers policy insights for crafting region-specific development strategies that reflect each province's economic stage.

The main contribution of this paper lies in its method. By applying the GAM, the study blends flexibility with interpretability, uncovering hidden patterns that linear models miss. While earlier research has explored the drivers of inequality in Indonesia, few have captured its nonlinear character so effectively.

The paper proceeds as follows: Section 2 reviews the key theoretical and empirical literature. Section 3 explains the research design and data, including the GAM framework. Section 4 discusses the empirical results and their implications. Finally, Section 5 concludes with recommendations for promoting more balanced and inclusive regional development.

2 Literature Review

The issue of regional income inequality has remained at the forefront of development economics for more than half a century. Theoretical debates have evolved from classical to modern growth perspectives, highlighting the dynamic interaction among education, capital, and labor as key drivers of inequality. The literature demonstrates that these variables influence inequality in nonlinear and context-dependent ways. This section reviews three major strands of relevant scholarship: (1) theories of regional inequality and convergence, (2) empirical evidence on education, capital, and labor in relation to income disparities, and (3) recent methodological advancements in modeling nonlinear relationships.

Classical development economics provides an early foundation for understanding regional disparities. Kuznets [3] proposed the famous inverted U-shaped relationship between income inequality and economic development, arguing that inequality initially rises as countries industrialize and urbanize but eventually declines once broader segments of the population gain access to education and employment opportunities. This hypothesis has shaped the analysis of inequality trends in developing economies, including Indonesia. However, later studies have noted that the Kuznets curve does not always hold empirically, especially in nations where institutional quality and structural transformation are uneven [4, 5].

Neoclassical growth theory expanded the analysis by linking inequality to factor accumulation and convergence processes. According to Solow's model and its subsequent extensions, regions with lower initial capital stocks should experience faster growth, leading to income convergence over time. Yet, this outcome assumes similar savings rates, technology levels, and human capital endowments—conditions that rarely hold in reality [1, 3]. Differences in educational attainment, infrastructure, and labor productivity often prevent poorer regions from catching up. Thus, inequality persists not merely because of initial differences but due to structural and institutional constraints that shape how capital and labor are utilized.

Human capital theory, rooted in the works of Becker [2] and Schultz [6], underscores education as the central mechanism for improving productivity and promoting income mobility. Investment in education enhances individual skills, technological adoption, and overall economic competitiveness. However, the relationship between education and inequality is complex. At early stages of development, access to education tends to be concentrated among wealthier households, leading to higher inequality. Over time, as public education expands and more individuals acquire skills, inequality tends to decline [7]. This nonlinear pattern implies that the equalizing effect of education depends heavily on the inclusiveness and quality of educational systems.

Empirical studies on Indonesia and other developing countries have revealed mixed results regarding the impact of education on income distribution. Akita and Lukman [8] found that improvements in education reduce inequality across Indonesian provinces, but the effect is stronger in provinces with higher secondary education completion rates. Conversely, Tansel and Daoud [9] reported that tertiary education could initially widen inequality due to a scarcity of skilled jobs and high wage premiums in urban centers. These findings suggest that the inequality-education nexus is stage-dependent, warranting models that can accommodate varying marginal effects.

Physical capital also plays a central role in shaping regional inequality. In the neoclassical tradition, capital accumulation promotes economic growth by enhancing productivity and enabling industrialization. Yet, the benefits of capital investment are often spatially concentrated, favoring provinces with existing infrastructure and business ecosystems. Empirical research has shown that the marginal productivity of investment is higher in underdeveloped regions, implying diminishing returns to capital in richer areas [10]. However, the unequal distribution of investment opportunities often prevents convergence. In Indonesia, capital-intensive projects have historically been concentrated in Java and other western provinces, while eastern provinces remain resource-dependent and infrastructure-poor [11]. Such disparities perpetuate inequality despite aggregate growth.

Labor dynamics further interact with education and capital to determine income distribution. According to dual-sector and structural transformation theories, economic growth is typically accompanied by labor migration from low-productivity agricultural sectors to higher-productivity industrial and service sectors [12]. When this transition proceeds smoothly, inequality declines. However, when labor absorption in the modern sector lags behind the expansion of the labor force, urban unemployment and informal employment rise, leading to higher inequality [13]. In Indonesia, this structural imbalance remains pronounced. Provinces with large informal sectors often experience slower wage growth and weaker links between education and productivity, reinforcing inequality across regions [14].

Recent studies highlight the importance of incorporating nonlinearities and spatial heterogeneity into inequality analysis. Linear econometric models, such as Ordinary Least Squares (OLS) and Fixed Effects (FE) estimations, assume constant marginal effects and thus fail to capture threshold-dependent relationships. For instance, the impact of education on inequality may change once a province achieves a certain level of human capital saturation, while the effect of investment may vary depending on infrastructure maturity [15]. These complex dynamics necessitate more flexible modeling approaches that allow variables to exert effects that evolve with the data.

In response to these limitations, scholars have increasingly adopted nonparametric and semiparametric methods. The Generalized Additive Model (GAM), developed by Hastie and Tibshirani [16], has emerged as one of the most effective frameworks for exploring nonlinear relationships while preserving interpretability. Unlike machine learning models that often function as black boxes, the GAM estimates smooth functions for each explanatory variable, providing intuitive graphical interpretations. Wood
[17] further advanced this method by introducing penalized splines and robust inference

procedures. In economics, GAMs have been successfully applied to topics such as productivity growth, environmental efficiency, and income inequality [18, 19].

In the Indonesian context, the application of GAM to inequality analysis remains limited. Most regional inequality studies focus on spatial decomposition or linear regressions, leaving potential nonlinear patterns unexplored. For example, Akita [20] examined inequality decomposition across sectors and provinces, while Resosudarmo and Vidyattama [21] analyzed spatial spillover effects using panel data models. However, neither approach addressed the possibility that the influence of education, capital, and labor varies nonlinearly across development stages. This gap provides the rationale for employing GAM in the present study.

Based on the reviewed literature, three conclusions emerge. First, the relationship between education and inequality follows a nonlinear trajectory that depends on accessibility and inclusiveness. Second, the effects of physical capital and labor are also nonlinear and region-specific, often displaying diminishing or threshold-based patterns. Third, the methodological evolution toward nonparametric approaches such as GAM offers a promising avenue to better capture these complex dynamics. Addressing these dimensions is critical for producing policy-relevant insights into Indonesia's persistent regional inequality.

3 Research Methodology

This study uses a quantitative approach with panel data analysis to explore how education, capital formation, and labor interact in shaping regional income inequality across Indonesia's 34 provinces from 2004 to 2023. The framework is designed to uncover nonlinear patterns — the kinds of stage-dependent and threshold effects that are often hidden when researchers rely on conventional linear models. This section explains the data sources, key variables, model specification, estimation method, and the main hypotheses.

3.1 Data and Sources

The research draws on secondary data compiled from several well-established and publicly available sources. Most of the data come from *Badan Pusat Statistik* (BPS — Statistics Indonesia), supplemented by figures from the Ministry of Finance and the World Bank. Together, these form a balanced panel dataset of 680 observations (34 provinces over 20 years, from 2004 to 2023).

The dependent variable in this study is the Williamson Index (WI) — a widely used measure of income inequality between provinces. It is calculated as the populationweighted coefficient of variation in per capita Gross Regional Domestic Product (GRDP) [10].

Complementing the dependent variable, the explanatory variables are categorised into three core dimensions of production inputs. Firstly, human capital accumulation is measured through educational attainment, stratified into four distinct levels: primary, junior secondary, senior secondary, and tertiary education. These data are derived from

the National Socio-Economic Survey (Susenas) to capture the stock of human capital at various stages of schooling. Secondly, physical capital is proxied by Gross Fixed Capital Formation (PMTB) at constant prices. Sourced from Statistics Indonesia (BPS), this indicator reflects the real investment in productive assets and infrastructure at the provincial level. Thirdly, the labour factor is defined as the aggregate labour force, encompassing both employed and unemployed individuals aged 15 years and older, utilising data from the National Labour Force Survey (Sakernas).

To ensure robust estimation and comparability across heterogeneous provinces, all independent variables underwent natural logarithmic transformation. This methodological step serves to linearise exponential growth trends and attenuate potential data skewness. Consequently, the coefficients derived from the linear reference model can be interpreted directly as elasticity estimates, providing precise insights into the sensitivity of regional inequality to changes in these fundamental production factors.

3.2 Variable Construction

Table 1 summarizes the variables used in the analysis, their symbols, units of measurement, and data sources.

Table 1. Definition of Variables

Variable	Symbol	Unit	Source	Description
Williamson Index	WI	Ratio	BPS	Measure of interprovincial inequality
Primary School Graduates	SD	Persons	BPS–Susenas	Proxy for primary education attainment
Junior High School Graduates	SMP	Persons	BPS–Susenas	Proxy for junior education attainment
Senior High School Graduates	SMA	Persons	BPS–Susenas	Proxy for senior education attainment
Tertiary Graduates	PT	Persons	BPS–Susenas	Proxy for higher education attainment
Gross Fixed Capital Formation	PMTB	Billion IDR	BPS	Proxy for physical capital investment
Labor Force	LAB	Persons	BPS–Sakernas	Total labor force aged ≥15 years

All monetary figures were adjusted using Indonesia's GDP deflator (base year 2010 = 100) to remove the effects of inflation and ensure that comparisons over time reflect real values rather than price changes. Special attention was given to handling missing data, particularly for North Kalimantan, which became a province in 2012. Data entries for the years before its establishment were coded as zero and later excluded from the model estimation. This approach helps maintain both the balance and reliability of the dataset.

3.3 Empirical Framework

This study builds its empirical framework on the Generalized Additive Model (GAM), a semiparametric extension of the traditional linear regression model developed by Hastie and Tibshirani [16]. Unlike standard linear models that assume a constant relationship between variables, the GAM allows each explanatory factor to influence the dependent variable through a smooth, flexible, and nonlinear function. This flexibility makes it especially useful for capturing the complex economic dynamics found across regions—where the marginal effects of education, capital, or labor may shift depending on a province's stage of development.

The general specification of the model can be written as:

$$WI_{it} = \alpha + f_1(SD_{it}) + f_2(SMP_{it}) + f_3(SMA_{it}) + f_4(PT_{it}) + f_5(PMTB_{it}) + f_6(LAB_{it}) + \varepsilon_{it} \quad (1)$$

where:

- WI_{it}= Williamson Index for province *i* in year *t*;
- $f_j(\cdot)$= smooth, nonparametric function for variable *j*;
- α= intercept term;
- ε_{it}= error term, assumed to be independently and identically distributed with zero mean and constant variance.

This study builds its empirical framework on the Generalized Additive Model (GAM), a semiparametric extension of the traditional linear regression model developed by Hastie and Tibshirani [16]. Unlike standard linear models that assume a constant relationship between variables, the GAM allows each explanatory factor to influence the dependent variable through a smooth, flexible, and nonlinear function. This flexibility makes it especially useful for capturing the complex economic dynamics found across regions—where the marginal effects of education, capital, or labor may shift depending on a province's stage of development.

3.4 Model Estimation

The empirical estimation was conducted using the Nonparametric Regression module in EViews 13 within a Generalized Additive Model (GAM) framework. To facilitate this analysis, the dataset was organised as a balanced panel structure, with 'PROVINSI' serving as the cross-sectional unit and 'TAHUN' defining the temporal dimension. The model specification incorporated spline functions for each explanatory variable, a methodological choice that allows the estimated effects to vary smoothly across observations rather than being constrained to a fixed linear coefficient. Crucially, the optimal degree of smoothing for each functional relationship was determined endogenously through the Generalized Cross-Validation (GCV) method to minimise overfitting. Following the estimation process, partial effect plots were generated to visualise the trajectory of these non-linear relationships and to assess their statistical significance. To validate the efficacy of this approach, the performance of the GAM was rigorously benchmarked against a conventional linear Fixed Effects (FE) model utilising standard

goodness-of-fit metrics, specifically the Adjusted R^2, Akaike Information Criterion (AIC), Schwarz Criterion (SC), and the percentage of Deviance Explained.

This procedure ensured that the nonlinear effects captured by the GAM were robust, interpretable, and comparable to more traditional linear estimation approaches.

The GAM specification in EViews is implemented as follows:

$$WI_i = \alpha + s_1(SD_i) + s_2(SMP_i) + s_3(SMA_i) + s_4(PT_i) + s_5(PMTB_i) + s_6(LAB_i) + \sum_{j=2}^{K} \delta_j PROV_{ji} + \varepsilon_i \quad (2)$$

The estimated smooth functions are illustrated through graphical plots, where the horizontal axis represents each predictor variable and the vertical axis indicates its estimated partial contribution to the Williamson Index. The shaded areas around each curve depict the 95% confidence intervals, providing a visual sense of the statistical uncertainty surrounding the estimated relationships.

3.5 Model Validation

To evaluate the validity of the model, a series of diagnostic checks and comparative assessments were carried out. First, the explanatory power of the GAM was compared with that of a standard linear Fixed Effects (FE) model. The results show that the GAM explained 82.4% of the deviance, while the FE model accounted for only 68.1%, suggesting that the GAM captures the underlying dynamics far more effectively.

Next, model fit was assessed using the Akaike Information Criterion (AIC) and the Schwarz Criterion (SC). Both values were notably lower for the GAM (1250.3 and 1310.7, respectively) than for the FE model (1420.7 and 1481.2), confirming that the GAM provides a superior fit despite its greater flexibility and complexity.

Residual diagnostics were also performed to ensure that key assumptions were met — including normality, homoscedasticity, and independence. The residuals followed an approximately normal distribution, with no signs of serial correlation or heteroscedasticity, indicating that the model specification is sound. Finally, the partial effect plots were examined to visualize threshold points and nonlinear patterns across variables, offering deeper insights into how education, capital, and labor interact to shape inequality.

3.6 Research Hypotheses

Based on the theoretical framework and literature review, the study tests the following hypotheses:

- H_1: The relationship between education (SD, SMP, SMA, PT) and regional income inequality is nonlinear and level-dependent.
- H_2: Physical capital (PMTB) has a diminishing equalizing effect on inequality as the level of capital accumulation increases.
- H_3: Labor (LAB) exhibits a U-shaped relationship with inequality, reducing it up to a certain threshold but increasing it beyond that level.

These hypotheses are evaluated empirically using the GAM specification defined in Equation (1).

3.7 Summary of Methodological Approach

In summary, this study's methodological framework brings together a balanced provincial panel dataset, a semiparametric estimation approach using the Generalized Additive Model (GAM), and a series of rigorous validation tests. The flexibility of the GAM makes it possible to move beyond the restrictive assumptions of linearity and constant marginal effects, revealing nonlinear relationships that are essential for understanding the complexities of Indonesia's regional income inequality. By blending theoretical depth, empirical precision, and methodological innovation, this research establishes a strong foundation for the following sections, which present and discuss the key findings in detail.

4 Results

This section presents and interprets the empirical findings from the Generalized Additive Model (GAM), which examines how education, physical capital, and labor influence regional income inequality across Indonesia's 34 provinces between 2004 and 2023. The discussion unfolds in five parts: an assessment of model fit and performance, an exploration of the nonlinear effects of education, an analysis of capital formation, a review of labor's role, and finally, a comparison of the GAM results with those from traditional linear models.

4.1 Model Fit and Comparative Performance

Table 2 presents a comparison of the Generalized Additive Model (GAM) and the benchmark Linear Fixed Effects (FE) model. The results show that the GAM performs markedly better across all key metrics. It achieves a higher adjusted R^2, lower values for both the Akaike Information Criterion (AIC) and the Schwarz Criterion (SC), and explains a larger share of total deviance. Together, these indicators confirm that the GAM provides a more accurate and comprehensive representation of the factors driving regional income inequality in Indonesia.

Table 2. Comparison of Model Performance

Statistic	GAM (Proposed Model)	Linear FE (Benchmark Model)	Interpretation
Adjusted R^2	0.824	0.681	GAM explains 82.4% of variance, compared to 68.1% in FE
AIC	1250.3	1420.7	Lower AIC indicates better model fit
SC	1310.7	1481.2	GAM retains better fit despite higher complexity

Statistic	GAM (Proposed Model)	Linear FE (Benchmark Model)	Interpretation
Deviance Explained	82.4%	–	Confirms GAM's superior explanatory capacity

The notable improvement in model fit indicates that the nonlinear relationships captured by the GAM are both statistically significant and economically meaningful. The lower AIC and SC values further suggest that introducing smooth functional forms enhances model accuracy without leading to overfitting. With 82.4% of the total deviance explained, the GAM successfully accounts for more than four-fifths of the variation in regional income inequality, underscoring its robustness and suitability for analyzing complex regional dynamics.

5 Discussion

5.1 Nonlinear Effects of Education

The smooth functions for the four education variables—Primary School (SD), Junior High School (SMP), Senior High School (SMA), and Tertiary Education (PT)—reveal clear nonlinear patterns. Each level of education contributes differently to income inequality across Indonesia's provinces, reflecting the diverse stages of regional development.

The partial effect plots show that primary (SD) and junior high (SMP) education have a consistently negative relationship with the Williamson Index. As the number of graduates increases, the index steadily declines, suggesting that expanding access to basic education plays a strong equalizing role. These findings are consistent with theories linking early education to broad-based productivity gains and reduced skill disparities [6, 8]. The 95% confidence intervals for both SD and SMP remain entirely below zero, confirming the statistical significance of their inequality-reducing effects.

For senior high education (SMA), the pattern is nonlinear but still predominantly negative. At lower completion levels, its effect on inequality is minimal, indicating that secondary education has limited impact in early stages of development. However, once graduation rates reach mid-level thresholds, the curve slopes downward more sharply—showing that secondary education becomes increasingly effective in narrowing income gaps as more individuals complete this level. This finding supports the idea of "threshold externalities," where the benefits of education expand substantially once a critical mass of participants is achieved [7].

The relationship for tertiary education (PT) is notably different, forming an inverted U-shaped curve. At early stages—when the number of university graduates is relatively small—the effect is positive, meaning inequality initially increases. This aligns with the "elite formation" hypothesis, where early access to higher education is concentrated among wealthier households, amplifying wage differentials [9, 10]. However, once the number of tertiary graduates surpasses roughly 600,000 800,000 per province, the curve bends downward and eventually crosses below zero. Beyond this point, higher

education starts to reduce inequality by boosting overall human capital, supporting technological diffusion, and expanding access to better employment opportunities across wider segments of society.

Overall, these results highlight that the impact of education on income inequality is highly stage-dependent. Basic education consistently reduces inequality, secondary education grows more effective as participation rises, and tertiary education shifts from an inequality-enhancing to an equalizing force as access broadens. Consequently, policies promoting educational equity should be tailored to both the level of education and the regional stage of development, ensuring that investments in education yield the strongest and most inclusive outcomes.

5.2 Nonlinear Impact of Physical Capital

The partial effect function for Gross Fixed Capital Formation (PMTB) reveals a concave relationship with the Williamson Index. At lower levels of capital accumulation, increases in PMTB substantially reduce income inequality—reflecting the powerful role of investment in stimulating productivity, improving infrastructure, and expanding regional employment opportunities [11]. However, as PMTB levels continue to rise, the curve gradually flattens and eventually turns slightly positive, suggesting that the equalizing benefits of capital investment diminish over time.

This pattern captures a familiar development dynamic: as regional economies mature, new investment tends to concentrate in already developed provinces, where higher productivity, market access, and agglomeration economies offer greater returns. Over time, this concentration can inadvertently reinforce spatial disparities rather than mitigate them [12, 13].

From a policy standpoint, these results highlight the importance of strategic and regionally differentiated capital allocation. Public and private investments should be directed toward capital-scarce regions, where the marginal returns to investment are highest and the potential for reducing inequality is greatest. Equally important are complementary policies that strengthen governance, infrastructure, and institutional capacity—factors that enhance a region's ability to absorb and effectively utilize new capital.

5.3 Labor and Inequality Dynamics

The estimated smooth function for the labor variable (LAB) reveals a clear U-shaped relationship with the Williamson Index. At lower levels of labor participation, an expanding workforce helps reduce inequality by broadening employment opportunities and improving income distribution. However, beyond a certain threshold—when labor supply begins to outpace job creation—the curve turns upward, suggesting that an excess labor force can actually intensify inequality.

This finding aligns with dual-sector development models [12, 13], which describe how early stages of structural transformation—from agriculture to industry—often promote convergence, while later stages can generate wage polarization if the industrial sector's capacity to absorb workers stagnates. In Indonesia, this pattern is visible in

several provinces where formal job creation has not kept pace with rapid labor force growth, resulting in persistent informality and widening wage gaps.

The U-shaped relationship underscores that labor's influence on inequality depends critically on employment elasticity. When the labor market is dynamic and inclusive, additional labor contributes to greater equality. But when growth becomes jobless or heavily capital-intensive, expanding labor supply leads to underemployment and rising inequality.

To address this, labor policies must be designed to complement education and investment strategies—focusing on labor-intensive industries, skills development, and productivity-driven job creation that ensure inclusive participation in economic growth.

5.4 Comparative Interpretation and Policy Insights

A comparative analysis between the nonlinear Generalized Additive Model (GAM) and the conventional Linear Fixed Effects (FE) model underscores the value of moving beyond linear assumptions. In the FE model, the coefficients for education, capital, and labor are fixed across all provinces and years—masking the threshold effects that naturally emerge as regions develop at different paces. By contrast, the GAM reveals that these relationships are far more nuanced and dynamic, varying meaningfully across provinces and stages of economic maturity.

For instance, the FE model suggests that tertiary education uniformly increases inequality, producing a positive coefficient. However, the GAM shows that this effect only holds in early stages of expansion—when access to higher education is limited—and reverses once participation broadens. Similarly, while the FE model treats investment as an inherently equalizing factor, the GAM demonstrates that its impact diminishes as regions become more developed. These contrasts highlight how the nonlinear framework captures the real-world complexity that linear models tend to oversimplify, resulting in insights that are far more relevant for policy design.

Overall, the findings suggest that Indonesia's regional inequality is driven by interacting nonlinear forces. The effects of education, capital, and labor differ across provinces depending on their development stage. In early-stage regions, basic education and capital accumulation yield the strongest equalizing effects. In contrast, more advanced provinces benefit most from tertiary education expansion and labor market reforms that improve inclusiveness and employment quality.

Based on these insights, three strategic policy directions emerge:

- Differentiated Education Policy – Expand secondary and tertiary education access in lagging regions, while prioritizing quality improvement in advanced ones.
- Balanced Capital Allocation – Direct infrastructure and investment toward capitalscarce provinces to maximize equalizing returns.
- Inclusive Labor Policy – Encourage labor-intensive industrialization and vocational training programs to align labor growth with productive job creation.

Together, these strategies can help reshape Indonesia's inequality landscape by addressing the nonlinear, region-specific drivers of economic disparity.

6 Conclusion and Policy Recommendations

This study explored how education, physical capital, and labor interact in nonlinear ways to shape regional income inequality across 34 Indonesian provinces between 2004 and 2023. Using the Generalized Additive Model (GAM), the research uncovered stage-dependent and threshold-based dynamics that traditional linear approaches tend to overlook. The findings offer fresh theoretical and policy insights into how inequality evolves and how regional development strategies should be tailored to local conditions.

The empirical results reveal that education's influence on inequality varies by level of attainment. Primary (SD) and junior high (SMP) education consistently reduce inequality, confirming that broad access to basic education remains a powerful equalizer. Senior high (SMA) education also contributes to equality—but only once participation surpasses a certain threshold, indicating that mass secondary education is essential for building an inclusive human capital base. In contrast, tertiary education (PT) follows an inverted U-shaped pattern: at early stages, limited access allows high-income groups to capture most of the benefits, widening inequality; yet as access expands, higher education begins to close income gaps by spreading skills and productivity more broadly. This pattern supports the education threshold hypothesis, showing that education's impact depends on both scale and inclusiveness.

Physical capital formation—proxied by Gross Fixed Capital Formation (PMTB)—displays a concave relationship with inequality. In less developed provinces, investment helps narrow income gaps by boosting productivity and infrastructure. However, as regions become wealthier, the marginal equalizing effect of capital declines. At high investment levels, new capital increasingly flows to already developed areas, reinforcing agglomeration advantages and widening spatial disparities. These findings suggest that investment's equalizing potential hinges on regional absorptive capacity and balanced capital distribution.

Labor participation exhibits a U-shaped relationship with inequality. In early development phases, a growing labor force promotes equality by expanding employment opportunities. But beyond a certain point—when job creation lags behind labor supply or growth is concentrated in informal sectors—inequality rises again. This underscores the need for policies that align labor supply with productive, high-quality job creation. Provinces with rapid labor growth but limited industrial diversification face heightened inequality risks, highlighting the importance of labor market flexibility, skills development, and industrial upgrading.

A comparison between the GAM and the Linear Fixed Effects (FE) model further demonstrates the value of capturing nonlinear dynamics. The GAM achieved an adjusted R^2 of 0.824, significantly outperforming the FE model's 0.681, while also producing lower AIC and SC values. The GAM's smooth function plots vividly illustrate how nonlinear and threshold effects operate across provinces—evidence that would be obscured in a purely linear framework. These results confirm that inequality is shaped by interactions that evolve with regional development, and that forcing linear assumptions can lead to oversimplified or misleading policy conclusions.

Theoretically, this study contributes by integrating nonlinear modeling into the analysis of regional inequality determinants, extending classic theories such as the Kuznets

and human capital hypotheses. It demonstrates that the effects of education, capital, and labor are not static, but evolve alongside structural transformation. Methodologically, it shows that the GAM offers a valuable balance between flexibility and interpretability, allowing researchers to visualize nonlinear effects transparently without relying on opaque machine-learning models. Empirically, this research provides one of the most comprehensive analyses to date of Indonesia's inequality dynamics—spanning two decades of provincial data—and offers actionable insights for more inclusive and regionally balanced development.

6.1 Policy Recommendations

Building on the empirical findings, this study proposes three strategic policy directions to reduce regional income inequality through more inclusive and differentiated development approaches.

Level-Specific Human Capital Policy. Policies should focus on expanding and equalizing access to secondary and tertiary education, while maintaining universal coverage at the primary level. Investments in teacher quality, digital learning infrastructure, and vocational training are essential to ensure that education translates into tangible productivity and income gains.

In less developed provinces, policy efforts should prioritize basic and secondary education to establish a strong foundation of human capital. In contrast, more advanced regions should emphasize tertiary and technical education to promote innovation and skill upgrading. Additionally, a more equitable geographic distribution of higher education institutions is critical to prevent the concentration of knowledge and income opportunities in major urban centers.

Spatially Balanced Capital Investment. Both public and private investments must be strategically directed to provinces where marginal returns to capital are highest and where capital scarcity constrains growth. Decentralizing infrastructure development—particularly in transportation, energy, and logistics—can help lagging regions attract new industries and investment.

Improving institutional quality and governance at the provincial level is equally important for ensuring efficient capital use and preventing resource misallocation. Targeted fiscal incentives, coupled with well-structured public–private partnerships (PPPs), can further stimulate productive investment in underdeveloped regions, fostering long-term convergence and spatial equity.

Inclusive and Productivity-Driven Labor Market Policies. Labor market reforms should aim to enhance job creation in labor-intensive and value-added industries. Vocational training programs that align with local economic potential—such as manufacturing, agro-industry, and services—are key to boosting employability and productivity.

Policies that improve labor mobility across provinces through better information systems, transport infrastructure, and social protection mechanisms can reduce wage disparities and promote balanced growth. Furthermore, formalizing informal employment and strengthening small and medium enterprises (SMEs) will ensure that labor participation contributes to reducing inequality rather than perpetuating it.

Broader Implications and Future Research Directions. The findings of this study carry significant implications for both policy and academic inquiry. From a policy standpoint, they highlight that regional inequality is inherently multidimensional and nonlinear, demanding adaptive strategies rather than uniform solutions. Policymakers must adopt stage-sensitive approaches that reflect each province's level of development and institutional capacity. Expanding tertiary education or boosting investment without complementary labor reforms may unintentionally exacerbate inequality—underscoring the need for policy coordination across sectors.

From a research perspective, this study opens several promising avenues for further exploration. Future work could integrate spatial dependence and spillover effects through models such as spatial GAMs or geographically weighted regression. Incorporating institutional quality and governance indicators could also deepen understanding of how political economy dynamics mediate the effects of education, capital, and labor on inequality. Finally, future research should explore the gender and digital dimensions of human capital formation, as these factors are increasingly shaping Indonesia's evolving labor market and inequality landscape.

Concluding Remarks. In conclusion, this study demonstrates that Indonesia's regional income inequality is driven by complex, nonlinear interactions among education, physical capital, and labor. By applying the Generalized Additive Model (GAM), the research uncovered important threshold effects and diminishing returns that linear models often overlook. The evidence clearly shows that the forces reducing inequality at one stage of development may amplify it at another.

Recognizing these nonlinearities allows policymakers to craft more effective, regionally tailored development strategies—aligning education, investment, and labor policies with the specific conditions of each province. Sustained commitment to equitable human capital development, spatially balanced investment, and inclusive labor market reform will be vital in achieving Indonesia's long-term goal of reducing regional disparities and promoting inclusive, sustainable growth.

References

1. Barro, R.J.: Economic Growth and Convergence Across the World. American Economic Review 106(5), 407–443 (1991).
2. Becker, G.S.: Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. University of Chicago Press (1993).
3. Mankiw, N.G., Romer, D., and Weil, D.N.: A Contribution to the Empirics of Economic Growth. Quarterly Journal of Economics 107(2), 407–437 (1992).

4. Kuznets, S.: Economic Growth and Income Inequality. American Economic Review 45(1), 1–28 (1955).
5. Piketty, T.: Capital in the Twenty-First Century. Harvard University Press (2014).
6. Schultz, T.W.: Investment in Human Capital. American Economic Review 51(1), 1–17 (1961).
7. Barro, R.J., and Lee, J.W.: Education Matters: Global Schooling and Economic Growth. Oxford University Press (2013).
8. Akita, T., and Lukman, R.A.: Spatial Patterns of Inequality in Indonesia: 1993–2018. Bulletin of Indonesian Economic Studies 56(2), 187–212 (2020).
9. Tansel, A., and Daoud, Y.: Returns to Education and Regional Inequality in Developing Countries. Economic Modelling 76, 200–214 (2019).
10. Williamson, J.G.: Regional Inequality and the Process of National Development. Economic Development and Cultural Change 13(4), 3–45 (1965).
11. World Bank: Indonesia: Equity and Development Report. World Bank, Washington, D.C. (2022).
12. Lewis, W.A.: Economic Development with Unlimited Supplies of Labour. Manchester School 22(2), 139–191 (1954).
13. Todaro, M.P., and Smith, S.C.: Economic Development. 13th ed. Pearson Education (2020).
14. McCulloch, N., and Sjahrir, B.S.: Endowments, Location or Luck? Evaluating Regional Economic Performance in Indonesia. Asian Economic Journal 32(3), 249–273 (2018).
15. Balisacan, A.M., and Fuwa, N.: Growth, Inequality, and Politics in Developing Countries: Insights from Southeast Asia. Asian Economic Policy Review 12(1), 68–95 (2017).
16. Hastie, T., and Tibshirani, R.: Generalized Additive Models. Chapman & Hall/CRC, London (1990).
17. Wood, S.N.: Generalized Additive Models: An Introduction with R. 2nd ed. CRC Press, Boca Raton (2017).
18. Kumbhakar, S.C., and Lovell, C.A.K.: Stochastic Frontier Analysis. Cambridge University Press (2000).
19. Mitra, A., and Yadav, R.: Nonlinear Dynamics of Regional Growth and Inequality: Evidence from South Asia. Empirical Economics 65(4), 1623–1651 (2023).
20. Akita, T.: Decomposing Regional Income Inequality in Indonesia: A Sectoral Perspective. Annals of Regional Science 44(2), 421–439 (2010).
21. Resosudarmo, B.P., and Vidyattama, Y.: Regional Income Disparities in Indonesia: A Panel Data Analysis. ASEAN Economic Bulletin 27(3), 260–277 (2010).