The Long Shadow: Childhood Poverty and the Returns to Education

Ade Febriady[*] Agnieszka Postepska Viola Angelini

Version: March 27, 2026

Abstract

This study documents substantial heterogeneity in returns to education by childhood poverty status among Indonesian wage workers aged 15-35. Individuals who grew up poor earn only 1.5 percent per additional year of schooling—less than one-fourth of the 6.8 percent earned by those who were never poor. We estimate these returns using a control-function approach that exploits conditional heteroskedasticity for identification in the absence of exclusion restrictions. The control-function coefficient is three times larger among the poor, indicating markedly stronger positive selection into schooling in this group: only individuals with exceptionally favorable unobserved characteristics attain higher levels of education. We also present descriptive evidence of lower skill accumulation per year of schooling and more limited access to high-paying jobs among disadvantaged individuals, patterns consistent with lower marginal returns. These findings highlight the limited equalizing role of education, measured here by years of schooling.

Keywords: Returns to Education, Childhood Poverty, Control-Function Approach

JEL Classification: I24, I26, I3

[*]Corresponding author: a.febriady@rug.nl. All authors are affiliated with the Department of Economics, Econometrics and Finance, University of Groningen. Agnieszka Postepska and Viola Angelini are also affiliated with the Global Labor Organization (GLO). We thank Rob Alessie and Aukje Nieuwenhuis for helpful discussions. We are also grateful to seminar and conference participants at the University of Groningen, the Causal Inference & Machine Learning Conference 2025 (IMT Lucca), and the Dutch Economists Day 2025 (Leiden University) for valuable comments. Any remaining errors are our own. We acknowledge financial support from LPDP (Indonesia Endowment Fund for Education), which supports Ade's doctoral research.

1 Introduction

Education is widely regarded as an equalizer, yet its labor-market returns may be lower for individuals who grew up poor. Early-life disadvantage generates developmental deficits that constrain skill accumulation over time[1] (see, for example, Heckman, 2006; Cunha and Heckman, 2007; Attanasio et al., 2022) and contribute to persistent and widening socioeconomic status (SES) gaps in academic achievement (von Stumm et al., 2022; Chmielewski, 2019). For instance, two-thirds of the SES gap in math performance among adults aged 25–29 had already emerged by age ten (OECD, 2018). Early disadvantage also extends beyond schooling: even at the same level of educational attainment, individuals from low-SES backgrounds face systematic barriers at the hiring stage (Rivera and Tilcsik, 2016; Belmi et al., 2024). Unequal skill formation and hiring disadvantages together may explain why individuals who grew up poor earn systematically lower returns to education in the labor market.

Yet whether labor-market returns to education actually vary by SES remains an open question. Among early contributions, Papanicolaou and Psacharopoulos (1979) found higher returns for low-SES individuals in the UK, arguing that schooling builds skills more lacking among the poor. By contrast, Armitage and Sabot (1987) documented higher returns for high-SES individuals in Kenya and Tanzania, attributing this to strong complementarities between home and school investments in human capital. In the United States, Cohn and Kiker (1986) found no SES gradient in returns to education, while more recent studies on college returns in the same country report mixed evidence—some find higher returns for low-income individuals (Cheng et al., 2021), others for high-income individuals (Bartik and Hershbein, 2018). These mixed findings likely reflect contextual differences, varying definitions of SES, and different empirical strategies as summarized in Appendix Table A1.

Against this backdrop, we examine whether returns to education among wage workers vary systematically by childhood poverty status, using longitudinal data from the Indone-

[1]Disparities in school quality may also contribute to unequal skill formation; however, data limitations prevent us from examining this channel directly.

sia Family Life Survey (IFLS). We focus on individuals observed during childhood (ages 0–14) and measure their labor-market outcomes as young adults (ages 15–35)[2]. We assess heterogeneity by comparing returns between individuals who experienced childhood poverty and those who did not.

Schooling decisions are endogenous[3]; therefore, OLS estimates of returns lack a causal interpretation. A common solution is to estimate returns using an instrumental variables (IV) approach. However, identifying credible instruments for schooling in this context has proven difficult.[4] We address this challenge by implementing the control-function estimator of Klein and Vella (2010) (KV), following the parametric specification in Farré et al. (2013). This approach achieves identification through conditional heteroskedasticity rather than exclusion restrictions.

Beyond overcoming this limitation, the KV approach offers two additional advantages. First, the control function allows us to characterize how selection into education differs across poverty groups. Second, because identification relies on heteroskedasticity that operates broadly across our samples, it may recover effects closer to the population average rather than effects specific to the subgroup affected by a particular instrument (see discussion in Saniter, 2012).[5]

Our findings reveal that the childhood poor earn returns of just 1.5 percent per additional year of schooling, compared with 6.8 percent for the non-poor—a gap of 5.3 percentage points that nearly doubles the OLS estimate of 2.9 percentage points once endogeneity is corrected. This widening reflects stronger positive selection among the poor: only those with exceptionally favorable unobserved traits—such as ability or motivation—manage to attain higher education. Since these traits are independently rewarded in the

[2]We define childhood over this range to maximize the likelihood of observing early-life conditions given gaps in survey implementation. We exclude individuals who are still enrolled in school. Age 35 is the upper bound that still allows us to observe childhood conditions in at least one survey wave.

[3]Endogeneity arises because unobserved determinants of earnings—such as ability, motivation, or family background—may also affect schooling choices, leading to correlation between schooling and the error term in the earnings equation.

[4]The common IVs are not well-suited here, see Section 3.

[5]The key intuition is that identification through heteroskedasticity operates across all observations rather than through a specific instrument that shifts only a subgroup of compliers; thus, the resulting estimates are not bounded to a local average treatment effect (LATE). Our Breusch–Pagan and White tests confirm significant heteroskedasticity in both the wage and education equations across poverty subsamples (see Table 2).

labor market, controlling for them reveals that the true returns to schooling are substantially lower for the poor than the OLS estimates suggest. To gain insight into the mechanisms, we also document descriptive evidence—differential skill gains per year of schooling and unequal access to high-paying jobs—that are consistent with lower returns, although we do not establish them as causal mechanisms. Taken together, our results suggest that expanding educational access alone may be insufficient to promote equality.

This paper makes several contributions. First, to our knowledge, it is the first to document substantial heterogeneity in returns to education by childhood poverty status in Indonesia. In doing so, it contributes to the broader literature on heterogeneous returns across socioeconomic groups while offering stronger internal validity than is typical in this literature by addressing the endogeneity of schooling decisions.

Second, it contributes to research on both intergenerational mobility and early-life conditions. For the former, we show that childhood poverty constrains economic opportunity not only by limiting educational attainment (see, for example, Bellani and Bia, 2019), but also by lowering the marginal return to an additional year of schooling. For the latter, a literature which has largely examined educational attainment and earnings as outcomes (e.g., Almond, 2006; Almond and Currie, 2011; Black et al., 2007; Maccini and Yang, 2009), we highlight a distinct margin through which early-life disadvantage shapes adult outcomes: the payoff to schooling conditional on attainment.

Third, our findings offer an alternative perspective on Indonesia's "paradox of progress"—educational expansion coinciding with rising inequality (Bourguignon et al., 2005). Beyond the standard explanations of convex returns and unequal access to higher education, we show that individuals from poorer backgrounds systematically earn lower returns even at the same attainment level.[6] Finally, our results speak to the literature on SES gaps in enrollment: Boneva and Rauh (2017) show that low-SES students perceive lower returns to education, and our estimates imply that realized returns are indeed lower, suggesting that policies focused solely on easing liquidity constraints may be insufficient if returns

[6] Alatas and Bourguignon (2005) documented variation in returns across urban/rural and formal/informal sectors during 1980–1996 but did not examine differences by childhood socioeconomic status.

themselves remain unequal.

The remainder of the paper proceeds as follows. Section 2 introduces the Indonesian context. Section 3 describes the empirical strategy and data. Section 4 presents the main results and explores descriptive patterns associated with the returns gap. Section 5 concludes.

2 Indonesia's Context

As the world's largest archipelagic state—spanning more than 5,000 kilometers and comprising 17,508 islands according to official figures, or 13,558 based on satellite-based census estimates (Andréfouët et al., 2022)—Indonesia faces structural challenges to equitable development (Figure 6 in the Appendix). Massive school construction in the 1970s and subsequent expansion of secondary and tertiary education have largely resolved the challenge of access.[7] With access largely achieved, the challenge that remains is one of quality: ensuring that children receive an education of comparable standard regardless of where they live. Geographic fragmentation complicates this, as physical mobility constraints limit students' ability to reach better-resourced schools, while institutional complexity—a split between two ministries and decentralized management by districts of widely varying capacity—produces a patchwork of public, private, and religious schools of uneven quality.

Critically, these spatial disparities in school quality map onto socioeconomic disadvantage. Yarrow et al. (2020) find that teachers in urban areas score significantly higher on mathematics assessments than those in rural areas, with further gaps between public and religious schools. At the district level, Pradhan and de Ree (2014) found substantial heterogeneity in teacher qualifications: in the lowest-performing 10 percent of districts (based on standardized elementary test scores), only 15 percent of teachers held a bachelor's degree, compared to over 70 percent at the top. Recipients of a needs-based schol-

[7] The Net Enrolment Rate (NER) for primary education was already 92% in 1994 and reached 97% by 2025 (Badan Pusat Statistik, 2025). Lower secondary NER rose from 50% to 80%, and upper secondary NER roughly doubled from 33% to 66% over the same period, with particularly sharp growth between 2013 and 2016.

arship for poor households (PIP) are substantially more concentrated in precisely these lower-quality rural schools, and rural students are also less likely to attend preschool—potentially compounding later learning gaps.

These quality gaps are reflected in student outcomes: PIP recipients score about 6 percentage points lower in mathematics, even after controlling for student, teacher, and school characteristics (Yarrow et al., 2020). The spatial clustering of disadvantage and systematic gaps in service delivery provide context for heterogeneous returns to schooling. While unequal school quality is one plausible contributor, we do not observe it directly in our data; accordingly, we examine more directly observable correlates—differences in skill accumulation and labor-market access—in Section 4.2.

3 Methodology and Data

3.1 Empirical Strategy

We estimate returns to education using a standard Mincerian model (Mincer, 1974), regressing log hourly wages (w) on years of schooling ($educ$) and controls (x), as in Equation 1. The central concern is endogeneity: more able individuals select into higher education, inducing a correlation between the wage equation error u_i and the education equation error v_i ($\text{Cov}(u, v) \neq 0$), so that OLS estimates lack a causal interpretation.

$$w_i = \beta x_i + \gamma \text{educ}_i + u_i, \quad i = 1, \ldots, n \tag{1}$$

$$\text{educ}_i = \delta x_i + v_i \tag{2}$$

To motivate our identification strategy, we reformulate the model in a control function (CF) framework. Decomposing the structural error as $u_i = \lambda v_i + e_i$, where $e_i \perp v_i$, and substituting into Equation 1 yields Equation 3. Including v_i as a control term absorbs the endogenous component, so that e_i is uncorrelated with $educ$ and γ is consistently estimated.

$$w_i = \beta x_i + \gamma \text{educ}_i + \lambda v_i + e_i, \quad i = 1, \ldots, n \tag{3}$$

The problem in estimating Equation 3 is that v_i is a perfect linear combination of *educ* and x_i. A standard solution uses instrumental variables excluded from the wage equation to break this collinearity—the resulting CF estimator is numerically equivalent to IV (Wooldridge, 2015).

However, finding credible instruments is often difficult, and Indonesia is no exception. The school construction instrument of Duflo (2001) does not apply to our much younger sample, and Roodman (2025) raises concerns about its identification. Compulsory schooling reforms provide limited leverage: the 1994 reform had little effect on attainment (Lewis and Nguyen, 2020), and the 2015 reform post-dates our data. Family background instruments are particularly likely to violate the exclusion restriction (see discussion in Card, 1999). Given these limitations, we adopt the CF estimator of Klein and Vella (2010)—henceforth KV—which achieves identification through conditional heteroskedasticity rather than exclusion restrictions. KV show that that if in at least one of the equations the error is heteroskedastic and the correlation between u and v is constant (ρ), an appropriate CF can be derived. To see how, express λ in Equation 3 as:

$$\lambda = \frac{\text{cov}(u,v)}{\text{var}(v)} = \frac{\text{cov}(u,v)}{\sigma_v^2} = \frac{\text{cov}(u,v)}{\sigma_v \sigma_v} \cdot \frac{\sigma_u}{\sigma_u} = \frac{\text{cov}(u,v)}{\sigma_u \sigma_v} \cdot \frac{\sigma_u}{\sigma_v} = \rho \cdot \frac{\sigma_u}{\sigma_v}$$

where $\rho = \text{Cov}(u,v)/(\sigma_u \sigma_v)$. When the error variances are heteroskedastic—i.e., σ_j varies with observables via functions $S_j(x_j)$ for $j = u, v$—the endogeneity parameter becomes:

$$\lambda(x_u, x_v) = \rho \cdot \left[\frac{S_u(x_{ui})}{S_v(x_{vi})}\right]$$

Substituting this parameter into Equation 3 yields the final CF estimation equation:

$$w_i = \beta x_i + \gamma \text{educ} + \rho \frac{S_u(x_{ui})}{S_v(x_{vi})} v_i + e_i \quad i = 1, \ldots, n \tag{4}$$

Estimation of Equation 4 proceeds in two steps. First, we recover a consistent estimate of v from Equation 2, denoted $\hat{v}$. From this, we estimate the heteroskedasticity functions S_v and then S_u. Klein and Vella (2010, 2009) estimate these semi-parametrically, but Farré et al. (2013) show that a parametric specification performs comparably. We follow this parametric approach, specifying $S_{ji}^2 = \exp(z_{ji}\theta_j)$ for $j = u, v$, where z denotes the predictors in the variance function, which need not equal x.[8] Details of the estimation procedure are provided in Appendix A.6.

Two identification conditions underpin the KV estimator. First, the constant correlation condition (CCC) requires $E[u_i v_i \mid x_i] = \rho$: conditional on observables, the correlation between the structural errors is constant. Appendix A.5 shows that a multiplicative error structure supports this condition. Although untestable, CCC is implicitly embedded in standard IV strategies (Saniter, 2012)—as the λ decomposition above makes clear—and has been applied in related wage-education settings (Klein and Vella, 2009; Farré et al., 2013; Saniter, 2012; Souza and Zylberstajn, 2020).

Second, the ratio $S_u(z_i)/S_v(z_i)$ must not be constant across individuals, requiring heteroskedasticity in at least one equation. Birth region, for example, affects both the mean and variance of educational attainment. We confirm this requirement by rejecting homoskedasticity using the Breusch–Pagan and White tests (Table A8).

Table 1 presents the control variables. Two choices merit discussion. First, we control for age rather than potential experience ($age - schooling - entry\ age$), which is a deterministic function of schooling that complicates endogeneity corrections (Buscha and Dickson, 2023). Second, we exclude occupation and post-schooling cognitive skills from the wage equation because they are themselves outcomes of schooling (Angrist and Pischke, 2009); instead, we use them descriptively when exploring patterns associated with the returns gap.

Our preferred specification uses the same controls in both mean equations and in both variance functions, except that birth region is excluded from the wage variance (S_u). The rationale is that birth region plausibly affects the variance of educational attainment—through heterogeneous school quality and regional opportunity costs—but

[8]Any excluded predictors in z do not directly identify γ as in IV, but they provide identifying power by explaining heteroskedasticity in the error terms (Farré et al., 2013).

Table 1: Control Variables by Equation in the Baseline Model

Equation	**List of Control Variables** x, z
Conditional Mean Equations	
Wage (Equation 1)	Age, age^2, gender, ethnicity, birth region
Education (Equation 2)	Age, age^2, gender, ethnicity, birth region
Conditional Variance Equations $(S_j(z), \quad j = u, v)$	
Education (S_v)	Age, age^2, gender, ethnicity, birth region
Wage (S_u)	Age, age^2, gender, ethnicity

is less relevant for wage variance conditional on education, since wages are determined primarily in current labor markets.[9] Breusch–Pagan and White tests confirm that birth region contributes meaningfully to educational variance (Table A8), and including it in both variance functions yields consistent results (Table A14).

To examine heterogeneous returns, we split the sample by childhood poverty status, estimating Equation 4 separately for each group. This allows group-specific correction terms and sidesteps the problem of including a binary endogenous regressor, for which a linear CF might be inappropriate due to inherent heteroskedasticity. As a robustness check, we replace the binary indicator with continuous childhood consumption expenditure and implement a secondary CF—an extension that is straightforward given that the KV method readily accommodates multiple continuous endogenous terms (see, for example, Postepska, 2019)—as described in Section 4.3.

[9]Indonesia's 2020 Population Census documents substantial internal mobility: around 27 million (10 percent) individuals reside in a province different from their province of birth (Badan Pusat Statistik, 2023).

3.2 Data

We use data from the Indonesia Family Life Survey (IFLS), a longitudinal household survey initiated in 1993 that represents approximately 83 percent of Indonesia's population at baseline (Strauss et al., 2016). Our analysis draws on all five waves (1993, 1997, 2000, 2007, 2014), exploiting the survey's high recontact rates to track individuals from childhood into young adulthood.

We build the analytical sample from individuals observed in IFLS5 (2014) who are aged 15 or older and no longer enrolled in school, ensuring a focus on labor-market participants. Key variables are wages, educational attainment, and childhood welfare—measured as per capita consumption expenditure (PCE) during ages 0–14.

Childhood poverty is measured from earlier waves: an individual is classified as childhood poor if ever observed below the World Bank's $2.15/day poverty line (2017 PPP) during ages 0–14. We use the full 0–14 age range rather than restricting the sample to early childhood to maximize the sample size in light of the spacing between survey waves. For example, limiting the window to ages 0–5 would mechanically exclude the oldest individuals in our sample (aged 27–35), who were not surveyed during their earliest years. In addition, individuals would be observed in only one survey wave within that narrower window, increasing the risk of misclassifying those who experienced poverty during unobserved periods. Detailed variable definitions appear in Appendix A6.

This "ever poor" classification may introduce asymmetric misclassification: individuals observed as non-poor could have experienced poverty in unobserved waves, contaminating the non-poor group and attenuating the estimated gap toward zero.[10] If anything, this misclassification biases our estimates of the returns gap toward zero. Beyond measurement error, childhood poverty may itself be endogenous; we address this concern in Section 4.3 by replacing the binary indicator with continuous childhood consumption

[10]Our sample's ever-poor rate of 54 percent (3,231 of 5,939) is plausible given that the World Bank's poverty headcount at the same $2.15/day threshold was 61% in 1993, 44% in 2000, and 22% in 2007—the survey years from which we observe childhood consumption—with a spike to 68% during the 1997–98 crisis (World Bank, 2026). The rate falls below the early-1990s cross-sectional headcount, as expected: an "ever poor" measure based on intermittent survey waves will undercount true cumulative exposure, and conditioning on wage employment further excludes individuals in deeper poverty who are less likely to hold formal jobs.

expenditure and implementing a secondary control function.

The resulting sample spans ages 15–35, ensuring that at least part of each individual's childhood was recorded in earlier waves. Because we condition on wage employment, our estimates do not generalize to the self-employed or those outside the labor force. However, descriptive evidence suggests that this conditioning does not introduce ability-based selection: Appendix Table A16 shows that cognitive ability (Raven's scores and numeracy) is virtually identical between wage workers and self-employed, with normalized differences below 0.05 in both poverty groups.[11] Table A4 summarizes sample construction; Table A5 shows no meaningful attrition bias.

The final analytical sample consists of 5,939 individuals, split into two groups: those who experienced childhood poverty (Poor, $n = 3{,}231$) and those who did not (Non-poor, $n = 2{,}708$). Full summary statistics are available in Appendix A7.

3.2.1 Descriptive Statistics

Table A7 summarizes the two groups. The childhood poor earn approximately 31 percent lower hourly wages (log: 8.72 vs. 9.10), attain 1.93 fewer years of schooling (9.91 vs. 11.83), and are substantially more likely to come from rural areas (Java–village: +8.6 pp; Off-Java–village: +9.4 pp), suggesting a spatial dimension to poverty that intersects with limited access to quality education and formal employment. Figure 1 shows that while wages increase with education for both groups, the childhood poor earn less at almost every level. The gap widens with attainment, becoming statistically significant at elementary completion (6 years) and junior secondary completion (9 years), and growing more pronounced and consistent beyond senior secondary completion (12 years).

[11] If less-able poor individuals sorted into self-employment, the wage-worker sample among the poor would be positively selected on ability, potentially understating the returns gap. The similar cognitive profiles across employment types, for both the poor and non-poor, provide evidence against this concern. Wage workers do have about one more year of schooling than the self-employed, but this gap is nearly identical for the poor (0.33) and non-poor (0.35), so it does not create differential selection by poverty status.

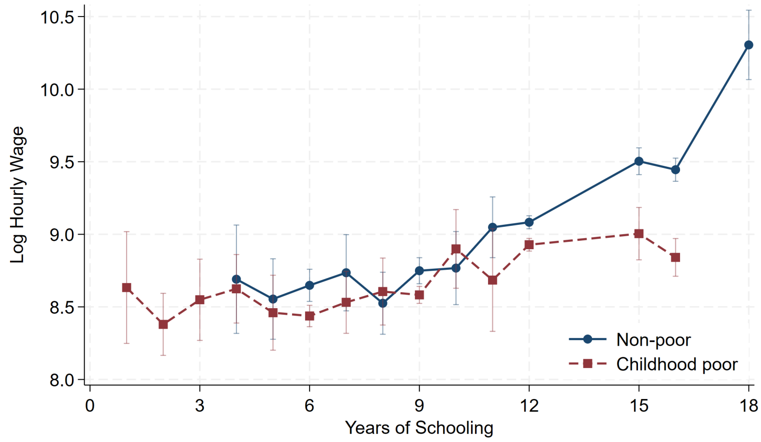

Figure 1: Mean Log Hourly Wage by Years of Schooling and Childhood Poverty Status

Notes: Figure displays mean log hourly wages by years of schooling with 95% confidence intervals. Sample: wage workers aged 15–35 from IFLS waves 1–5. Only cells with $n \geq 20$ observations are shown.

4 Results and Discussion

Table 2 reports OLS and Klein–Vella (KV) estimates of returns to schooling for the full sample and by childhood poverty status. OLS serves as a benchmark but yields biased estimates when schooling is endogenous. The KV estimator addresses this through a control function correction term, $(S_u/S_v) \cdot \hat{v}$, whose coefficient captures the direction and magnitude of selection into schooling. Comparing this coefficient across subgroups reveals whether selection differs systematically between individuals who grew up poor and those who did not. All specifications include the same set of covariates; full results appear in Appendix Table A17.

The results reveal substantial heterogeneity in returns to education based on childhood poverty status. OLS estimates indicate that each additional year of schooling increases wages by 8.0 percent in the full sample, with returns of 8.8 percent for the non-poor and 5.8 percent for the poor, a gap of 2.9 percentage points. When applying the KV control function to address endogeneity, the overall returns drop to 4.2 percent for the full sample, but the correction operates very differently across groups: returns decrease to 6.8 percent for the non-poor but fall sharply to just 1.5 percent for the poor. This widens the gap to 5.3 percentage points or 78 percent lower returns for those who faced childhood poverty.

Table 2: Returns to Education by Childhood Poverty Status

	OLS				Klein-Vella			
	All	Non-poor	Poor	Diff	All	Non-poor	Poor	Diff
Years of schooling	0.080***	0.088***	0.058***	0.029***	0.042***	0.068***	0.015**	0.053***
	(0.003)	(0.005)	(0.005)	(0.007)	(0.008)	(0.007)	(0.007)	(0.010)
$(S_u/S_v) \cdot \hat{v}$					0.183***	0.077**	0.230***	-0.153***
					(0.037)	(0.034)	(0.036)	(0.049)
Observations	5,939	2,708	3,231		5,939	2,708	3,231	
Heteroskedasticity Tests (p-values):								
BP test (wage)					0.000	0.000	0.000	
BP test (educ)					0.000	0.000	0.000	
White test (wage)					0.000	0.024	0.000	
White test (educ)					0.000	0.000	0.000	

Notes: Dependent variable: log hourly wage. Bootstrapped standard errors (500 replications) in parentheses. Diff = Non-poor minus Poor; SE computed as $\sqrt{SE_{NP}^2 + SE_P^2}$. All regressions include controls for age, age squared, gender, ethnicity, and birth region. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Correcting for endogeneity lowers the estimated returns, which may appear atypical given that IV estimates in the returns-to-education literature often exceed OLS. However, as Card (2001) explains, supply-side instruments—such as school construction or compulsory schooling laws—identify a LATE for compliers whose marginal returns are plausibly high. IV estimates can therefore exceed the population average even when OLS is upward biased by ability. Previous Indonesian IV estimates of 6.8–10.6 percent (Duflo, 2001) and 12–17 percent from recent replication (Roodman, 2025) illustrate this pattern. Both reflect LATEs from supply-side interventions that target contexts with high marginal returns, and both focus exclusively on men, who tend to earn higher returns.[12]

Moreover, previous applications of the KV estimator that report higher returns than OLS do not interpret this pattern as evidence of no ability bias. Instead, these studies propose a plausible error structure in which a predictable positive component (ability) interacts with an unpredictable negative component (the overeducation penalty); see the discussion in (Farré et al., 2013; Klein and Vella, 2009). In this framework, a higher KV estimate captures the net effect of these offsetting forces rather than the absence of selection on ability. In addition, these studies analyze returns to education in different institutional settings, such as the United States and Australia, where labor market conditions and educational systems differ.

4.1 Understanding the OLS-KV Gap: Differential Selection

The downward adjustment from OLS to KV reflects positive selection: individuals with advantageous unobserved characteristics—higher ability, motivation, or supportive family backgrounds—are more likely to obtain additional schooling and to earn higher wages, as confirmed by the positive control function coefficient (ρ).

This positive selection is markedly stronger among the childhood poor: the correction term for the poor (0.230) is roughly three times that of the non-poor (0.077), indicating that among the poor only individuals with exceptionally favorable unobserved traits attain higher education. Once this selection is accounted for, the estimated causal return

[12]More generally, instruments that reduce the cost of schooling tend to affect individuals for whom the marginal return is relatively high, given diminishing returns; see Card (2001, 1999).

for the poor falls sharply.

Figure 2 offers descriptive support. Each panel plots Raven's Progressive Matrices z-scores—a measure of nonverbal reasoning less influenced by schooling—against bins of the education residual ($\hat{v}$), separately by poverty status and education tier. The positive slopes indicate that individuals with higher $\hat{v}$ tend to score higher on Raven's tests, consistent with interpreting $\hat{v}$ as capturing ability-related traits. Mean Raven scores also rise with schooling level, aligning with positive selection on ability. At the tertiary level, the relationship flattens, especially for the non-poor suggesting that other factors—such as motivation or family background—may play a larger role.

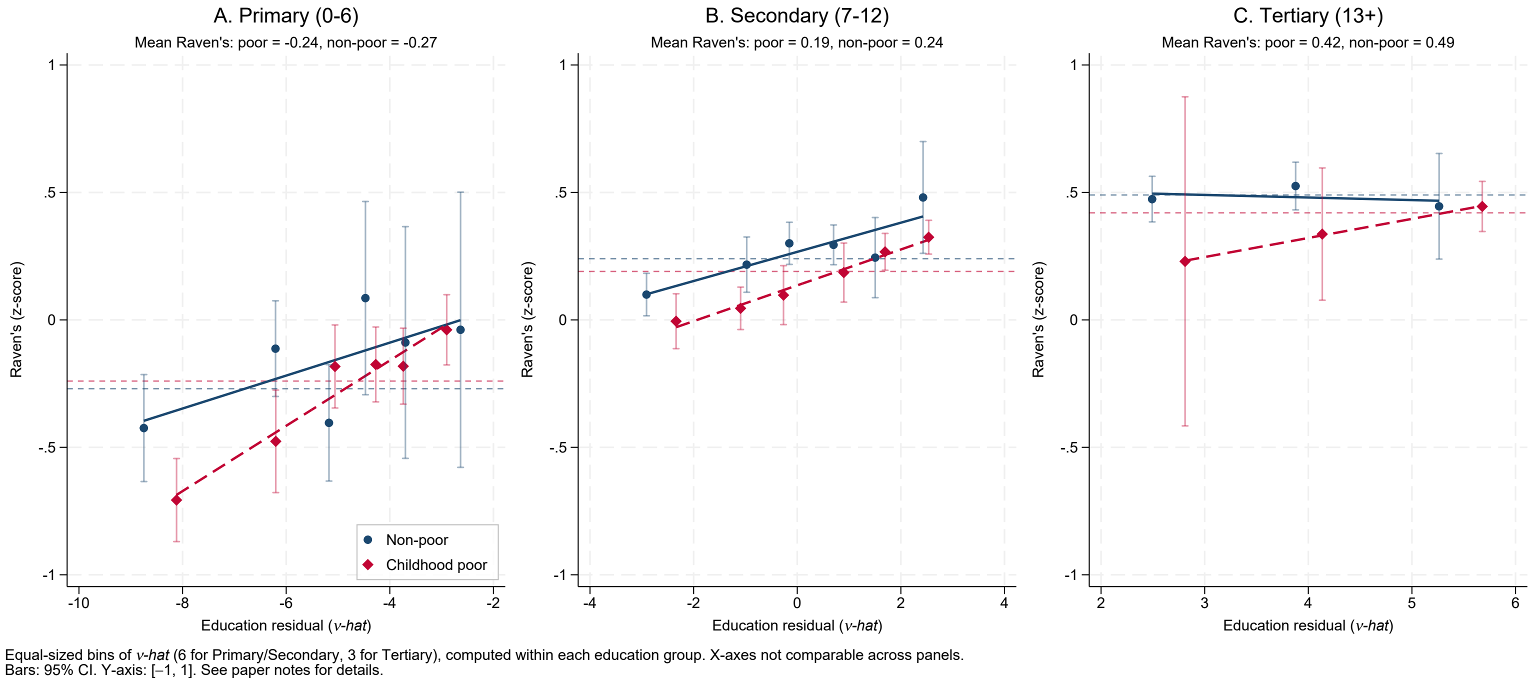

Figure 2: Education Equation Residuals ($\hat{v}$) and Raven's Cognitive Ability by Poverty Status and Education Tier

Notes: Each panel plots mean Raven's Progressive Matrices z-scores against bins of the education equation residual $\hat{v}$ (from Equation 2), separately for non-poor (navy circles, solid fit line) and childhood poor (cranberry diamonds, dashed fit line), within three education tiers: primary (0–6 years), secondary (7–12 years), and tertiary (13+ years). Dots show mean Raven's z-score within 6 equal-sized bins of $\hat{v}$, computed within each education group; x-axes are not comparable across panels. Lines show linear fit. Horizontal dashed lines mark group means for each poverty status. Bars show 95% confidence intervals. Y-axis fixed at $[-1, 1]$.

Since $\hat{v}$ likely captures multiple unobserved factors beyond cognitive ability—including motivation, family networks, and health endowments—so these patterns are suggestive. Nevertheless, they clarify why the returns gap widens after accounting for selection: OLS conflates returns to education with returns to unobserved traits, understating the true gap.

4.2 Descriptive Patterns Associated with Lower Returns

Having established the full extent of the gap, a natural question is what lies behind it. An obvious candidate is school quality. As documented in Section 2, school quality in Indonesia is far from randomly distributed across socioeconomic groups: children from poorer households are more likely to be in lower-quality schools with less qualified teachers and fewer resources, particularly in rural and remote districts. If the poor systematically attend worse schools, then an additional year of schooling may simply deliver less human capital, translating into lower wage returns even with the same number of years completed.

Since a direct test is not feasible, we examine differences in skill accumulation per year of schooling—which partly reflect cumulative school quality—as the closest proxy available in our data. We also explore differential access to high-paying jobs, which speaks to a distinct channel through which childhood poverty may shape labour market returns. These patterns are correlational rather than causal, but together they help characterize why returns may differ by childhood poverty status.

4.2.1 Differential Skill Gains per Year of Schooling

Lower returns among the childhood poor could reflect smaller cognitive gains per additional year of schooling. To examine this, we regress numeracy z-scores on years of schooling, childhood poverty, and their interaction, controlling for gender, ethnicity, and birth region. Numeracy is measured at the same survey wave as wages, by which point all respondents have completed schooling, so scores reflect cumulative learning. Figure 3 visualizes the results. Panel A plots predicted numeracy z-scores against years of education separately for the never-poor and childhood-poor groups, showing that the two groups start at similar levels but diverge as education increases—the poor gain less from each additional year of schooling. Panel B plots the difference in predicted scores (poor minus non-poor) at each education level, making the widening gap explicit. The difference becomes statistically significant at approximately 12 or more years of schooling—corresponding to high school completion—and aligns with the wage gap documented in Figure 1.

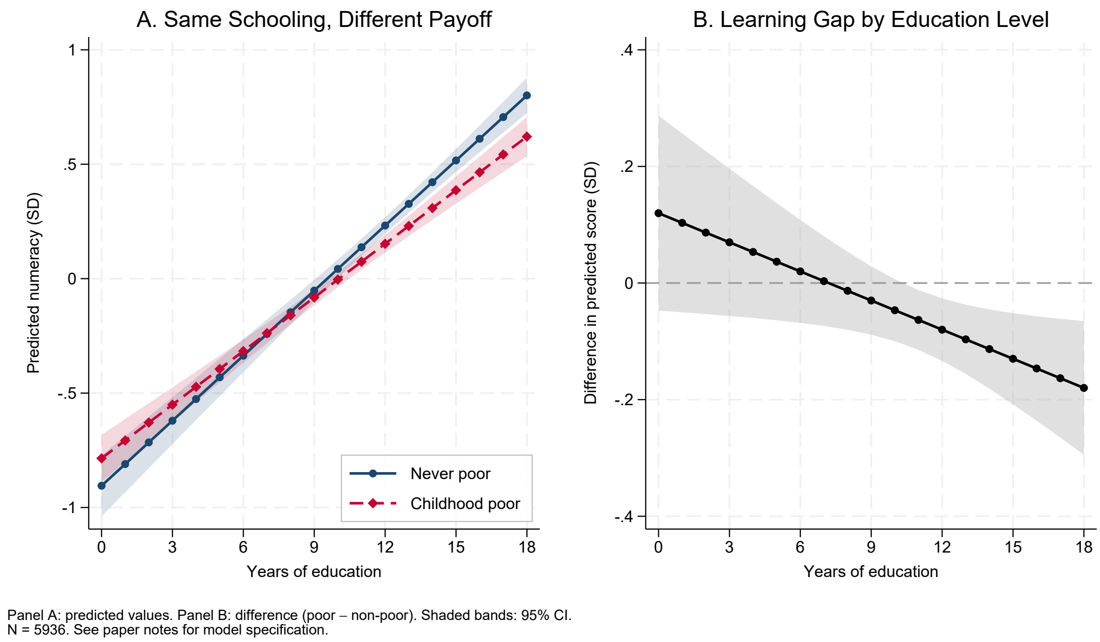

Figure 3: Unequal Learning from the Same Education

Notes: OLS: $y_i = \beta_1\, educ_i + \beta_2\, cpoor_i + \beta_3(educ_i \times cpoor_i) + \mathbf{X}_i\boldsymbol{\gamma} + \varepsilon_i$, where y is the numeracy z-score. Controls ($\mathbf{X}$): gender, ethnicity, birth region. $N = 5{,}936$. $\hat{\beta}_3 = -0.017$ ($p = 0.025$). Panel A: predicted numeracy z-scores by years of schooling for never-poor (navy, solid) and childhood-poor (cranberry, dashed). Panel B: difference in predicted scores (poor $-$ non-poor) at each education level. Shaded bands: 95% CI. Education range trimmed to observed support.

This differential learning gradient is consistent with differences in early-life foundations that may shape the capacity to benefit from schooling. Figure 4 documents these differences using horizontal coefficient plots showing the mean difference between childhood-poor and non-poor individuals from bivariate regressions. Panel A displays childhood conditions in percentage points: the childhood poor are substantially less likely to have attended kindergarten and more likely to have experienced hunger and school absence (albeit not statistically significant). Panel B displays family background in raw units: the poor have parents with considerably lower education levels and more siblings, potentially diluting parental time and resources per child.

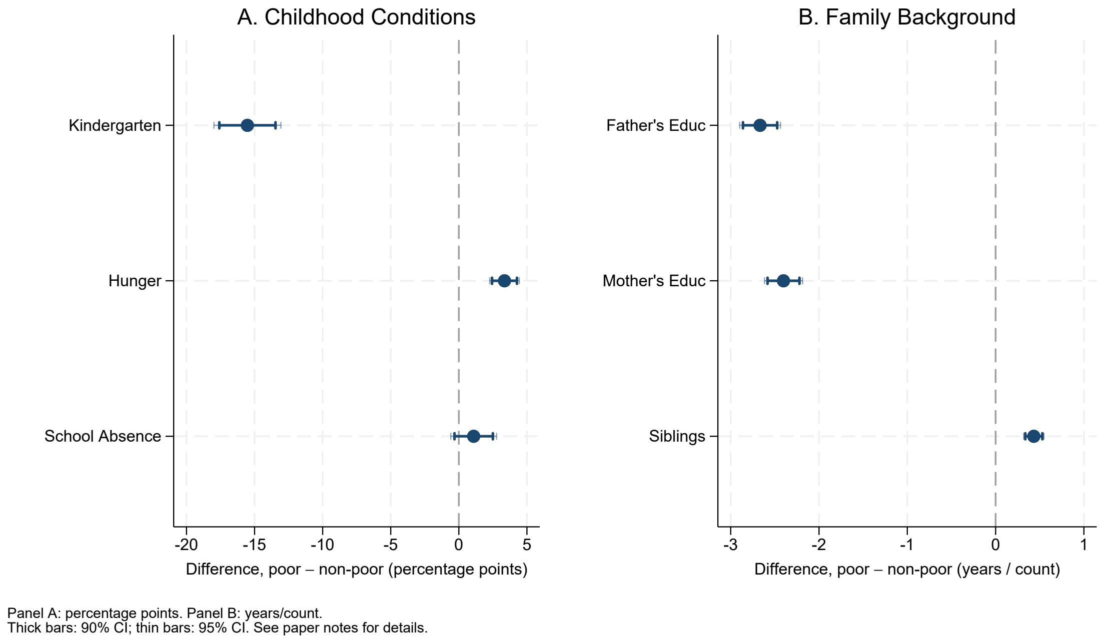

Figure 4: Early-Life Foundations: Poverty Gap

Notes: Each point shows the mean difference between childhood-poor and non-poor individuals from a bivariate OLS regression on childhood poverty status. Panel A: childhood conditions (kindergarten attendance, hunger, school absence) in percentage points. Panel B: family background (father's education, mother's education, siblings) in years or count. Vertical dashed line marks zero. Thick bars: 90% CI; thin bars: 95% CI. N: kindergarten = 5,917; hunger = 5,550; school absence = 5,543; father's educ = 4,882; mother's educ = 5,205; siblings = 5,939.

One interpretation draws on the self-productivity framework of Cunha and Heckman (2007): children who start school with weaker foundations fall further behind over time, as each year's learning builds on the previous year's. Even with equal years of schooling, cumulative skills end up lower among the poor. We note, however, that differences in school quality, peer effects, or other poverty-correlated factors could also contribute.

4.2.2 Labor-Market Barriers

Even if the poor accumulated skills at the same rate, they could earn lower returns if they face barriers to high-paying jobs. Randomized audit studies show that signaling low-SES origins (Rivera and Tilcsik, 2016) or first-generation college student status (Belmi et al., 2024) reduces employer callbacks. To probe whether analogous patterns exist in our data, we examine labor-market outcomes by education level within numeracy terciles.

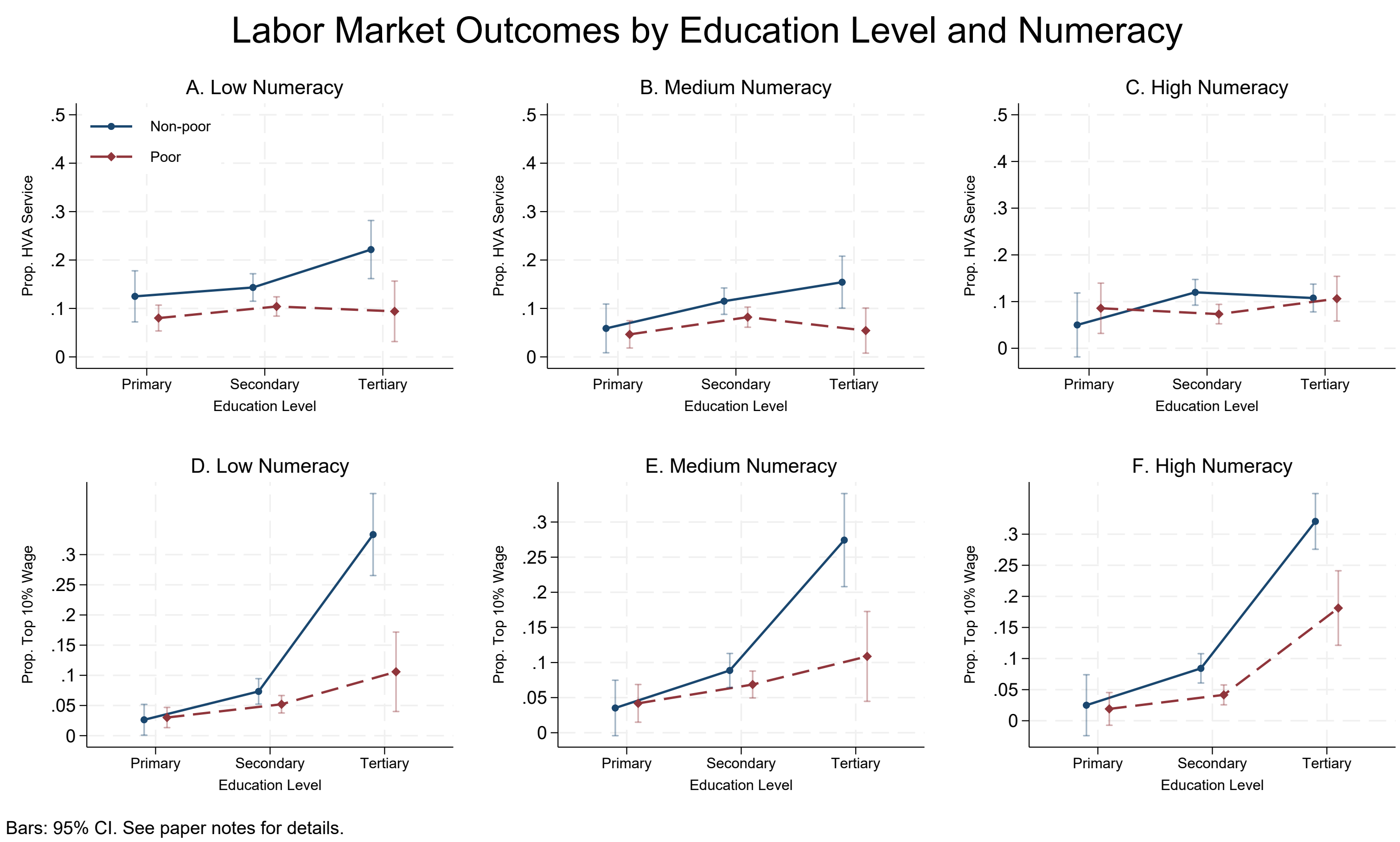

Figure 5: Labor Market Outcomes by Education Level and Numeracy Tercile

Notes: Row 1 (Panels A–C) shows the proportion employed in high value-added service sectors (finance, professional services, education, healthcare). Row 2 (Panels D–F) shows the proportion in the top 10% of the hourly wage distribution. Each column represents a numeracy tercile (Low, Medium, High). X-axis shows education groups (Primary, Secondary, Tertiary). Lines compare childhood poor vs non-poor. Bars show 95% confidence intervals.

Figure 5 reveals persistent gaps. Conditional on the same numeracy level, the childhood poor are less likely to work in high-value-added service sectors (top row) and less likely to reach the top decile of wages (bottom row). These gaps generally appear across all numeracy terciles. Because numeracy is itself influenced by education, conditioning on it should therefore be read as descriptive associations.

Possible explanations include disparities in networks and job information, geographic barriers, or employer discrimination based on observable markers of socioeconomic background. However, we cannot definitively identify which mechanisms are operative with our data.

4.3 Robustness Checks

We conduct several analyses to assess the robustness of the returns gap. First, we examine whether the disparity is driven by differences in educational attainment. Since the non-poor complete more schooling on average, the estimated returns might be upwardly biased if the education–earnings relationship is nonlinear. We therefore include a squared term for years of schooling to evaluate marginal returns across attainment levels.[13] Differential returns persist across all education levels and widen at higher levels (Table A11), reinforcing that the gap is not an artifact of attainment differences.

Additionally, childhood poverty itself may be endogenous. The KV framework can accommodate additional control functions, but a binary poverty indicator is problematic because linear probability models are inherently heteroskedastic. We therefore treat average childhood per capita consumption expenditure (PCE)[14] as an additional endogenous regressor, capturing heterogeneity through an education$\times$PCE interaction rather than sample splits. The results reveal a positive socioeconomic gradient: each additional dollar in childhood PCE raises returns by approximately 0.2 percentage points (Table A12). These findings corroborate our main result—even when treating SES as an additional

[13] As Wooldridge (2015) shows, the control function approach accommodates nonlinear functions of the endogenous variable—including squared terms and higher-order polynomials—using only the first-stage residuals, so no additional correction term is required for the squared term.

[14] We compute the average of observed childhood PCE across survey waves (with price adjustment). In Indonesia, PCE forms the basis for official poverty calculations.

endogenous regressor, the poor earn lower returns.

We also demonstrate that our results are not sensitive to outlier treatment or the specification of heteroskedasticity functions. Table A13 in the Appendix presents robustness checks using alternative dependent variable treatments. Our baseline model trims the top one percent of hourly wages; alternatives include no trimming, winsorizing at the 99th percentile, and recomputing hourly wages from unwinsorized working hours (our baseline winsorizes weekly hours at the 1st and 99th percentiles before computing hourly wages). All specifications yield similar results, with the gap ranging from 4.2 to 6.1 percentage points. Additionally, our baseline specification excludes birth region from the wage variance function (S_u). Table A14 in the Appendix shows that including birth region in both variance functions yields a larger gap of 8.1 percentage points, driven primarily by higher estimated returns for the non-poor (9.7 percent versus 6.8 percent in the baseline), while returns for the poor remain similarly low (1.5 percent).

We also verify that our findings are robust to the definition of the poverty line. Our baseline uses the World Bank's international poverty line of $2.15 per day (2017 PPP). Table A15 in the Appendix presents results using alternative thresholds: a stricter definition at 0.8×PL ($1.72/day), which classifies fewer individuals as poor, and a more relaxed definition at 1.2×PL ($2.58/day), which classifies more individuals as poor. Across all specifications, the childhood poor consistently earn significantly lower returns than the non-poor, with the gap ranging from 3.8 to 5.7 percentage points. This robustness across poverty definitions reinforces our main finding that those who grew up poor have substantially lower returns to education.

5 Conclusion

This study documents that wage-workers who experienced childhood poverty earn substantially lower returns to education in Indonesia. Using a control function estimator identified through conditional heteroskedasticity, we find that the childhood poor earn returns of just 1.5 percent per additional year of schooling, compared with 6.8 percent for the non-poor. This gap widens from 2.9 percentage points under OLS to 5.3 percentage points once endogeneity is corrected. The widening reflects stronger positive selection among the poor: individuals from disadvantaged backgrounds who obtain more schooling are more positively selected on favorable unobservables, such as ability, so that OLS attributes to education what partly reflects their exceptional characteristics. We also find descriptive patterns consistent with differential skill gains and labor-market barriers, though we do not establish them as causal mechanisms.

Several limitations warrant acknowledgment. First, our estimates are conditional on wage employment and do not capture returns among the self-employed or those outside the labor force. Second, the sample includes wage-workers aged 15–35, mostly reflecting early-career outcomes. Since returns to education evolve over the life cycle (Buscha and Dickson, 2023), the 1.5 percent return observed among individuals from poor backgrounds may reflect slower initial wage growth rather than permanently lower returns. If disadvantaged workers face barriers to high-quality first jobs but converge over time, the gap may narrow; if they remain in low-mobility occupations, it may persist or widen. Future research using longer panels could help explore these dynamics. Third, the KV estimator relies on heteroskedasticity-based identification, which—though partly testable—imposes structural assumptions. However, it is reassuring that the OLS estimates still indicate a statistically significant gap of 2.9 percentage points, which may represent a lower bound if, as our results suggest, positive selection is stronger among the poor. Finally, while we document patterns consistent with differential skill gains and labor-market barriers, this analysis is descriptive. Establishing the causal mechanisms underlying these effects should be central to future research.

Our findings carry broader implications. Substantially lower returns to education among the poor mean that childhood poverty not only limits educational attainment (Bellani and Bia, 2019) but also the economic benefit of each year of schooling. This offers a complementary perspective on Indonesia's "paradox of progress"—the coexistence of educational expansion and rising inequality (Bourguignon et al., 2005; Alatas and Bourguignon, 2005): even at the same attainment level, individuals from poorer backgrounds earn systematically lower returns. Our results also speak to persistently low enrollment among disadvantaged students. Boneva and Rauh (2017) show that low-SES students perceive lower returns to education; our estimates suggest these perceptions may be well-founded. Policies that expand enrollment without addressing the conditions that shape what students gain from school may do little to narrow the gap.

References

Alatas, Vivi and François Bourguignon, "The Evolution of Income Distribution during Indonesia's Fast Growth, 1980–96," in François Bourguignon, Francisco H. G. Ferreira, and Nora Lustig, eds., *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*, Washington, DC: World Bank and Oxford University Press, 2005, pp. 175–218.

Almond, Douglas, "Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population," *Journal of Political Economy*, 2006, *114* (4), 672–712.

_ **and Janet Currie**, "Human Capital Development before Age Five," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4, Elsevier, 2011, chapter 15, pp. 1315–1486.

Andréfouët, Serge, Mégane Paul, and A. Riza Farhan, "Indonesia's 13558 Islands: A New Census from Space and a First Step towards a One Map for Small Islands Policy," *Marine Policy*, 2022, *135*, 104848.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

Armitage, Jane and Richard Sabot, "Socioeconomic Background and the Returns to Schooling in Two Low-Income Economies," *Economica*, 1987, *54*, 103–108.

Attanasio, Orazio, Sarah Cattan, and Costas Meghir, "Early Childhood Development, Human Capital, and Poverty," *Annual Review of Economics*, 2022, *14*, 853–892.

Badan Pusat Statistik, *Statistics of Migration Indonesia: Results of the 2020 Population Census*, Jakarta: BPS—Statistics Indonesia, 2023.

— , "Education Indicator, 1994–2025," https://www.bps.go.id/en/statistics-table/1/MTUyNSMx/education-indicator--1994-2024.html 2025. Last updated: December 15, 2025. Accessed: February 24, 2026.

Bartik, Timothy J. and Brad J. Hershbein, "Degrees of Poverty: The Relationship between Family Income Background and the Returns to Education," Technical Report, W.E. Upjohn Institute 3 2018.

Bellani, Luna and Michela Bia, "The Long-Run Effect of Childhood Poverty and the Mediating Role of Education," *Journal of the Royal Statistical Society Series A: Statistics in Society*, 1 2019, *182* (1), 37–68.

Belmi, Peter, Kelly Raz, Margaret A. Neale, and Melissa C. Thomas-Hunt, "The Consequences of Revealing First-Generation Status," *Organization Science*, 2024, *35* (2), 667–697.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes," *The Quarterly Journal of Economics*, 2007, *122* (1), 409–439.

Boneva, Teodora and Christopher Rauh, "Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns," Working Paper 6756, CESifo Working Paper Series December 2017.

Bourguignon, François, Francisco H.G. Ferreira, and Nora Lustig, *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*, Washington, DC: World Bank and Oxford University Press, 2005.

Buscha, Franz and Matt Dickson, "Returns to Education: Individuals," in "Handbook of Labor, Human Resources and Population Economics," Springer, 2023, pp. 1–39.

Card, David, "The Causal Effect of Education on Earnings," *Handbook of Labor Economics*, 1999, *3*, 1801–1863.

— , "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 2001, *69* (5), 1127–1160.

Cheng, Siwei, Jennie E Brand, Xiang Zhou, Yu Xie, and Michael Hout, "Heterogeneous Returns to College over the Life Course," *Science Advances*, 2021, *7* (51), eabg7641.

Chmielewski, Anna K, "The Global Increase in the Socioeconomic Achievement Gap, 1964 to 2015," *American Sociological Review*, 2019, *84* (3), 517–544.

Cohn, Elchanan and B. F. Kiker, "Socioeconomic Background, Schooling, Experience and Monetary Rewards in the United States," *Economica*, 1986, *53* (212), 497–503.

Cunha, Flavio and James J Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

Duflo, Esther, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 2001, *91* (4), 795–813.

Farré, Lídia, Roger Klein, and Francis Vella, "A Parametric Control Function Approach to Estimating the Returns to Schooling in the Absence of Exclusion Restrictions: An Application to the NLSY," *Empirical Economics*, 2 2013, *44* (1), 111–133.

Heckman, James J, "Skill Formation and the Economics of Investing in Disadvantaged Children," *Science*, 2006, *312* (5782), 1900–1902.

Imbens, Guido W and Donald B Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, New York: Cambridge University Press, 2015.

Klein, Roger and Francis Vella, "Estimating the Return to Endogenous Schooling Decisions via Conditional Second Moments," *Journal of Human Resources*, 2009, *44* (4), 1047–1065.

— **and** — , "Estimating a Class of Triangular Simultaneous Equations Models without Exclusion Restrictions," *Journal of Econometrics*, 2 2010, *154* (2), 154–164.

Lewis, Blane D. and Hieu T.M. Nguyen, "Assessing the Causal Impact of Compulsory Schooling Policy in Indonesia," *International Journal of Educational Research*, 1 2020, *104.*

Maccini, Sharon and Dean Yang, "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, 2009, *99* (3), 1006–1026.

Mincer, Jacob, *Schooling, Experience, and Earnings*, New York: Columbia University Press, 1974.

OECD, *Equity in Education: Breaking Down Barriers to Social Mobility* PISA, Paris: OECD Publishing, 2018.

Papanicolaou, John and George Psacharopoulos, "Socioeconomic Background, Schooling and Monetary Rewards in the United Kingdom," *Economica*, 1979, *46* (184), 435–439.

Postepska, Agnieszka, "Ethnic Capital and Intergenerational Transmission of Educational Attainment," *Journal of Applied Econometrics*, 2019, *34* (4), 606–611.

Pradhan, Menno and Joppe de Ree, "District Governance and Student Learning in Indonesia," Technical Report 397, Asian Development Bank June 2014.

Rivera, Lauren A. and András Tilcsik, "Class Advantage, Commitment Penalty: The Gendered Effect of Social Class Signals in an Elite Labor Market," *American Sociological Review*, 2016, *81* (6), 1097–1131.

Roodman, David, "Schooling and Labor Market Consequences of School Construction in Indonesia: Comment," 2025.

Saniter, Nils, "Estimating Heterogeneous Returns to Education in Germany via Conditional Heteroskedasticity," IZA Discussion Paper 6813, Institute for the Study of Labor (IZA) 2012.

Souza, André Portela and Eduardo Zylberstajn, "Estimating the Returns to Education Using a Parametric Control Function Approach: Evidences for a Developing Country," *Brazilian Review of Econometrics*, March 2020, *39* (2).

Strauss, John, Firman Witoelar, and Bondan Sikoki, "The Fifth Wave of the Indonesia Family Life Survey (IFLS5): Overview and Field Report," Technical Report WR-1143/1-NIA/NICHD, RAND Corporation, Santa Monica, CA March 2016.

von Stumm, Sophie, Sophie Nicole Cave, and Paul Wakeling, "Persistent Association between Family Socioeconomic Status and Primary School Performance in Britain over 95 Years," *npj Science of Learning*, 2022, *7* (1), 4.

Wooldridge, Jeffrey M, "Control Function Methods in Applied Econometrics," *Journal of Human Resources*, 2015, *50* (2), 420–445.

World Bank, "Poverty and Inequality Platform," 2026. Version 20250930_2017_01_02_PROD [data set].

Yarrow, Noah, Rythia Afkar, Eema Masood, and Bernard Gauthier, "Measuring the Quality of MoRA's Education Services," Technical Report, World Bank, Jakarta 2020. World Bank.

A Appendix

Data Availability Statement

The Indonesia Family Life Survey (IFLS) data used in this study are publicly available from the RAND Corporation website. Registration is required for data access. Replication code, the custom Stata package `kvpara`, and all output tables and figures are available at the project's GitHub repository.

A.1 Summary of Literature

Table A1 summarizes empirical studies on whether returns to education vary by socioeconomic status (SES). Findings are mixed, likely reflecting differences in economic setting, SES measures, and estimation strategies. Many studies rely on OLS without strong causal identification, which may also explain the inconsistencies.

Table A1: Findings on the Returns to Education by SES

Study	**SES Measure**	**Method**	**Finding on Returns**
Papanicolaou & Psacharopoulos (1979) *United Kingdom*	Father's occupation	OLS	**Decline** in returns with higher SES
Armitage & Sabot (1986) *Kenya & Tanzania*	Parental education	OLS	**Increase** in returns with higher SES
Cohn & Kiker (1986) *United States*	Father's occupation (with IQ proxy)	OLS	**Similar** returns across SES groups
Bartik & Hershbein (2018) *US – College*	Family income (age 13–17)	OLS	**Increase** in returns with higher SES
Cheng et al. (2021) *US – College*	Parental income, education, ability	Matching + growth models	**Decline** in returns with higher SES

Note: Studies are arranged chronologically. SES refers to socioeconomic status, typically measured through parental characteristics. Returns refer to the marginal effect of education on earnings. OLS = Ordinary Least Squares estimation.

A.2 Levels and Years of Schooling

The formal education system comprises elementary (6 years), junior secondary (3 years), senior secondary (3 years), and tertiary levels; Table A2 in the Appendix details how we construct years of schooling from these levels.

Table A2: Structure of the Indonesian Education System and Implied Years of Schooling

Education Level *(Indonesian term)*	**Duration**	**Cumulative Years**
Primary Education		
Primary (*SD*)	6 years	1–6
Secondary Education		
Jr Secondary (*SMP*)	3 years	7–9
Sr Secondary (*SMA/SMK*)	3 years	10–12
Tertiary Education		
Diploma I/II/III	1–3 years	13–15
Bachelor (*Sarjana*)	4 years	13–16
Master (*Magister*)	2 years	17–18
Doctoral (*Doktor*)	4 years	19–22

Note: Cumulative years reflect the total implied years of schooling, consistent with the mapping in Table A3. Senior secondary comprises academic (*SMU*) and vocational (*SMK*) tracks. Equivalency programmes (Package A, B, and C) are treated as equivalent to primary, junior secondary, and senior secondary levels, respectively. Early childhood education is excluded. Total years of schooling range from 0 (never attended) to 22 (doctoral completion). Compulsory education policies were enacted for 6 years (1984), 9 years (1994), and 12 years (2015), but none provides a viable instrumental variable for our sample (aged 15–35 in 2014): the 6-year policy predates all sample cohorts (born 1979–1999), leaving no untreated comparison group; the 9-year policy was not effectively enforced and had limited impact on educational attainment (Lewis and Nguyen, 2020); and the 12-year policy post-dates the survey period.

Table A3: Construction of Years of Schooling from Survey Responses

Highest Level Attended	**Highest Grade Completed** *at that level*							
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Primary	0	1	2	3	4	5	6	6
Jr Secondary	6	7	8	9	9	9	9	9
Sr Secondary	9	10	11	12	12	12	12	12
Diploma I/II/III	12	13	14	15	15	15	15	15
Bachelor	12	13	14	15	16	16	16	16
Master	16	17	18	18	18	18	18	18
Doctoral	18	19	20	21	22	22	22	22

Note: Years of schooling is constructed from two survey variables: (i) the highest level of education attended and (ii) the highest grade completed at that level. Each cell reports the implied cumulative years of schooling. A value of 0 indicates not finishing first grade, while 7 indicates completion of the respective education level.

A.3 Indonesia's Geographic Scale: A European Comparison

Figure 6: Comparison of Indonesia's geographic scale overlaid on Europe.
Source: Author-generated image using ChatGPT and DALL·E (OpenAI), April 2025.

A.4 Building the Analytical Sample and Attrition Test

Table A4: Sample Construction

Step	Description	N
1	Age 15–35, positive hourly wage & non-missing education	6,375
2	+ Non-missing childhood poverty indicator	6,293
3	+ Trim top 1% of hourly wage (main sample)	6,193
4	+ Baseline controls (age, gender, ethnicity, birth region)	5,939

Notes: Step 5 defines the baseline sample used in the main results. Positive hourly wages exclude individuals with zero or missing wages, removing about 4 percent (283 observations) of the age-restricted sample, likely due to temporary non-employment or reporting errors.

Table A5: Attrition Test: Baseline vs Initial Sample

	Step 1 (Reference)		Step 4 (Baseline)				
Variable	Mean	(SD)	Mean	(SD)	Diff	Norm. Diff.	p-value
Log hourly wage	8.897	(0.888)	8.892	(0.883)	0.004	0.00	0.108
Years of schooling	10.765	(3.507)	10.783	(3.472)	-0.019	-0.01	0.115
Age	27.277	(5.107)	27.194	(5.070)	0.082	0.02	0.000
Female	0.366	(0.482)	0.369	(0.482)	-0.003	-0.01	0.109
Javanese	0.553	(0.497)	0.543	(0.498)	0.010	0.02	0.000
DKI Jakarta	0.063	(0.242)	0.061	(0.240)	0.001	0.00	0.028
Java - Village	0.350	(0.477)	0.353	(0.478)	-0.003	-0.01	0.011
Java - Small town	0.124	(0.329)	0.124	(0.329)	0.000	0.00	0.869
Java - Big city	0.033	(0.180)	0.033	(0.178)	0.001	0.00	0.047
Off-Java - Village	0.277	(0.448)	0.277	(0.448)	-0.000	-0.00	0.795
Off-Java - Small town	0.115	(0.319)	0.115	(0.319)	-0.000	-0.00	0.825
Off-Java - Big city	0.038	(0.190)	0.037	(0.188)	0.001	0.01	0.016
N	6,375		5,939				

Notes: Compares baseline analysis sample (Step 5) to the initial wage-worker sample (Step 2). Norm. Diff. $| > 0.25|$ suggests meaningful imbalance.

A.5 Error Structure

To motivate the plausibility of the *constant correlation condition* (CCC), we assume a multiplicative error structure. Let u^* and v^* denote the unscaled components of the errors, each with constant variance. The functions $S_j(x_j), \quad j = u, v$ capture the heteroskedastic terms, allowing the variance of the errors to depend on observable characteristics x.

$$u = S_u(x) \cdot u^* \tag{A.5}$$

$$v = S_v(x) \cdot v^* \tag{A.6}$$

- u^* and v^* are the **homoskedastic errors**: they are unobserved factors that have constant variance.
- $S_u(x)$ and $S_v(x)$ are **heteroskedasticity functions**: scaling factors that allow the total variance of u and v to depend on x.

Computing the Correlation Between u and v

To compute the correlation coefficient between u and v, we show that:

$$\rho = \frac{\text{Cov}(u, v)}{\sqrt{\text{Var}(u)} \cdot \sqrt{\text{Var}(v)}}$$

Step 1: Substitute the decomposed expressions.

We use the fact that $u = S_u(x) \cdot u^*$ and $v = S_v(x) \cdot v^*$:

$$\text{Cov}(u, v) = \text{Cov}(S_u(x) \cdot u^*, S_v(x) \cdot v^*)$$

Step 2: Factor out the deterministic functions of x.

Since $S_u(x)$ and $S_v(x)$ are just functions of observables and not random variables, we can factor them out:

$$\text{Cov}(u, v) = S_u(x) \cdot S_v(x) \cdot \text{Cov}(u^*, v^*)$$

Step 3: Similarly compute the standard deviations.

We apply the same decomposition to compute:

$$\sqrt{\text{Var}(u)} = S_u(x) \cdot \sqrt{\text{Var}(u^*)}, \quad \sqrt{\text{Var}(v)} = S_v(x) \cdot \sqrt{\text{Var}(v^*)}$$

But since u^* and v^* are homoskedastic, we can just write:

$$\sqrt{\text{Var}(u)} = S_u(x), \quad \sqrt{\text{Var}(v)} = S_v(x)$$

Step 4: Now plug back into the correlation formula.

$$\rho = \frac{S_u(x) \cdot S_v(x) \cdot \text{Cov}(u^*, v^*)}{S_u(x) \cdot S_v(x)} = \text{Cov}(u^*, v^*)$$

A.6 Estimation procedures

Following the parameterization of KV (2010),

$$S_{ji}^2 = \exp(z_{ji}\theta_j), \quad j = u, v, \tag{A.7}$$

where z_{ji} is the vector of variables considered to produce the heteroskedasticity in the respective equations and θ_j is a vector of unknown parameters.

Given the parameterization of S_{ji}^2, the estimation procedure is the following:

(i) Regress educ_i on its predictors to obtain a consistent estimate of the residual which we denote $\hat{v}_i$.

(ii) Estimate θ_v through ordinary least squares (OLS) using $\ln(\hat{v}_i^2)$ as the dependent variable. Note that we can estimate θ linearly via OLS after log-transformation of the residual ($\hat{v}_i^2$). Compute the standard deviation of the error as $\hat{S}_{vi} = \sqrt{\exp(z_{vi}\hat{\theta}_v)}$.

(iii) Using $\hat{v}_i$ and $\hat{S}_{vi}$, we estimate the wage equation parameters as follows:

- Estimate θ_u in $S_u(z_i)$ in a similar manner as is done for the education equation. For a given value of β, say β_c, we define the residual $u_i(\beta_c)$. Using this value of $u_i(\beta_c)$ we regress $\ln(u_i(\beta_c)^2)$ on $z_{ui}\theta_{cu}$ where we also use candidate values for θ_{cu}. From this regression we compute $\hat{S}_{ui}(\beta_c)$ as $\sqrt{\exp(z_{ui}\theta_{cu})}$ and estimate ρ_{0c} as:

$$\min_{\rho_{0c}} \sum_{i=1}^{n} \left(u_i(\beta_c) - \rho_{0c} \frac{\hat{S}_{ui}(\beta_c)}{\hat{S}_{vi}} \hat{v}_i \right)^2. \tag{A.8}$$

The final estimates of β_c, θ_{cu} and ρ_{0c} are those that minimize the above equation and are obtained through a standard iterative procedure.

We implement this estimator in Stata using a custom command `kvpara`, which is available in the replication package at github.com/febriady/Returns-to-Education. The command includes full documentation and allows for flexible specification of the heteroskedasticity functions, multiple endogenous regressors, and bootstrap standard errors.

A.7 Variable Definitions and Construction Notes

Table A6: Variable Definitions and Construction Notes

Variable	Definition	Construction/Additional Notes
Outcome and treatment variables		
Log hourly wage	Natural logarithm of hourly earnings from primary job	Computed from weekly wage (monthly wage/4) divided by weekly working hours. Weekly working hours are winsorized at the 1st and 99th percentiles to address extreme values before calculating hourly wages. Top 1% of hourly wages trimmed in baseline specification. Missing monthly wages are imputed using mid-points of wage ranges when available.
Years of schooling	Total years of formal education completed	Based on highest education level and grade completed (see Table A2).
Sample stratification variable		

Table A6 – continued from previous page

Variable	Definition	Construction/Additional Notes
Childhood poverty	Binary indicator: 1 if ever experienced extreme poverty during childhood (ages 0–14), 0 otherwise	Defined using the World Bank's international poverty line of $2.15 (2017 PPP) per capita per day. Daily per capita consumption derived from monthly household consumption using: monthly PCE $\times$ 12 $\div$ 365. Individuals classified as "ever poor" if observed in poverty in at least one IFLS wave during childhood.
Control variables		
Age	Individual's age in years	As recorded in IFLS 2014.
Age squared	Age squared	Included to capture non-linear age-earnings profile.
Female	Binary indicator: 1 if female, 0 otherwise	Self-reported in household roster.
Javanese	Binary indicator: 1 if Javanese ethnicity, 0 otherwise	Self-reported ethnicity. Javanese is the majority ethnic group in Indonesia.
Birth region indicators (omitted: DKI Jakarta)		

Table A6 – continued from previous page

Variable	Definition	Construction/Additional Notes
Birth region	Set of binary indicators for birthplace region and urbanicity	Based on self-reported recall of birthplace type (village, small town, or big city) and birth province from the IFLS migration module. Classification reflects respondents' subjective characterization of their birthplace rather than official statistical definitions. Categories: Java–Village, Java–Small town, Java–Big city, Off-Java–Village, Off-Java–Small town, Off-Java–Big city. Reference category: DKI Jakarta.
Other Variables		
Raven's z-score	Standardized cognitive ability score (by age-group)	Constructed from Raven's Colored Progressive Matrices (a closer measure of innate ability) and numeracy tests (more dependent on acquired skills). Classified by eight age groups: 15–19, 20–24, 25–29, 30–34, 35–39, 40–49, 50–59, 60+. Measured in the latest wave; since our sample is no longer in school, it serves as a post-schooling measurement.

Table A6 – continued from previous page

Variable	Definition	Construction/Additional Notes
Numeracy z-score	Standardized numeracy test (by age-group)	A component of the cognitive skills measurement focusing on numeracy. See above. As it is more dependent on acquired skills, we use it to proxy for learning quality when exploring what explains the differential returns to schooling.
Father's education	Father's total years of formal education	Calculated as *years of schooling* above.
Mother's education	Mother's total years of formal education	Same as above.
Number of siblings	Total number of siblings	Sum of siblings living in and outside the household.
Kindergarten	Binary indicator: 1 if attended kindergarten, 0 otherwise	Self-reported recall of kindergarten attendance.
Hunger	Binary indicator: 1 if experienced hunger during childhood, 0 otherwise	From IFLS early-life health module (eh08).
School absence	Binary indicator: 1 if missed school due to poor health during childhood, 0 otherwise	From IFLS early-life health module (eh02).

Table A6 – continued from previous page

Variable	Definition	Construction/Additional Notes
High value-added services	Binary indicator: 1 if employed in finance, professional services, education, or healthcare sectors	Constructed from 9-category sector classification.

A.8 Full Summary Statistics

Table A7: Summary Statistics: Baseline Variables by Childhood Poverty Status

	Non-poor		Poor				
Variable	Mean	(SD)	Mean	(SD)	Diff	Norm. Diff.	p-value
Log hourly wage	9.095	(0.867)	8.723	(0.859)	0.372	0.43	0.000
Years of schooling	11.831	(3.314)	9.906	(3.356)	1.925	0.58	0.000
Age	27.739	(5.091)	26.738	(5.008)	1.002	0.20	0.000
Female	0.400	(0.490)	0.343	(0.475)	0.057	0.12	0.000
Ethnicity (non-Javanese)	0.566	(0.496)	0.524	(0.500)	0.042	0.09	0.001
DKI Jakarta	0.103	(0.304)	0.027	(0.161)	0.076	0.31	0.000
Java - Village	0.306	(0.461)	0.392	(0.488)	-0.086	-0.18	0.000
Java - Small town	0.134	(0.340)	0.115	(0.320)	0.018	0.06	0.034
Java - Big city	0.041	(0.199)	0.025	(0.157)	0.016	0.09	0.001
Off-Java - Village	0.226	(0.419)	0.320	(0.467)	-0.094	-0.21	0.000
Off-Java - Small town	0.136	(0.343)	0.098	(0.298)	0.038	0.12	0.000
Off-Java - Big city	0.054	(0.226)	0.022	(0.148)	0.032	0.17	0.000
Observations	2,708		3,231			Total: 5,939	

Notes: Diff = Non-poor minus Poor. Norm. Diff. = Normalized difference, calculated as $(\bar{X}_{NP} - \bar{X}_P)/\sqrt{(s^2_{NP} + s^2_P)/2}$; values $| > 0.25|$ suggest meaningful imbalance (Imbens and Rubin, 2015). Source: IFLS waves 1–5.

A.9 Estimation of Education Equation 2

Table A8: Schooling Equation (First Stage)

	All	Non-poor	Poor
Constant	-3.315***	-6.266***	-3.655***
	(0.492)	(0.525)	(0.528)
Female	0.978***	1.039***	0.725***
	(0.079)	(0.081)	(0.081)
Ethnicity (Javanese)	0.006	0.084	-0.136
	(0.082)	(0.083)	(0.084)
Age	1.076***	1.293***	1.078***
	(0.032)	(0.033)	(0.034)
Age squared	-0.020***	-0.023***	-0.021***
	(0.000)	(0.001)	(0.001)
Java - Village	-1.025***	-0.493***	-0.280**
	(0.129)	(0.131)	(0.136)
Java - Small town	-0.458***	0.097	-0.023
	(0.163)	(0.152)	(0.166)
Java - Big city	0.197	0.737***	0.295
	(0.221)	(0.235)	(0.229)
Off-Java - Village	-0.799***	-0.185	0.012
	(0.131)	(0.135)	(0.138)
Off-Java - Small town	0.148	0.804***	0.307*
	(0.162)	(0.151)	(0.160)
Off-Java - Big city	0.670***	0.708***	1.046***
	(0.203)	(0.208)	(0.206)
Observations	5,939	2,708	3,231
R^2	0.064	0.089	0.041
F-statistic	40.27	26.21	13.77
White test p	0.000	0.000	0.000
Breusch-Pagan p	0.000	0.000	0.000

Notes: Dependent variable is years of schooling. The first stage regresses education on baseline controls and birth region fixed effects (reference: Jakarta). Bootstrapped standard errors in parentheses. White and Breusch-Pagan tests report p-values for heteroskedasticity in residuals; rejection supports identification via heteroskedasticity. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Heteroskedasticity Function: Education Equation (S_v)

	All	Non-poor	Poor
Constant	-0.716*	0.866**	-6.305***
	(0.418)	(0.418)	(0.433)
Female	-0.153**	0.378***	0.219***
	(0.061)	(0.065)	(0.069)
Ethnicity (Javanese)	0.242***	0.124**	0.411***
	(0.056)	(0.054)	(0.056)
Age	0.017	-0.163***	0.485***
	(0.023)	(0.023)	(0.024)
Age squared	0.001**	0.004***	-0.008***
	(0.000)	(0.000)	(0.000)
Java - Village	1.126***	1.360***	0.636***
	(0.237)	(0.225)	(0.233)
Java - Small town	0.630**	0.809***	0.834***
	(0.259)	(0.243)	(0.260)
Java - Big city	0.296	0.833**	0.861**
	(0.392)	(0.410)	(0.417)
Off-Java - Village	1.202***	1.495***	0.780***
	(0.238)	(0.226)	(0.232)
Off-Java - Small town	0.595**	1.208***	0.886***
	(0.279)	(0.254)	(0.270)
Off-Java - Big city	-0.306	0.759**	-0.006
	(0.387)	(0.356)	(0.348)
Observations	5,939	2,708	3,231

Notes: Estimates from the heteroskedasticity function $\ln(\hat{v}^2) = z'\theta_v$ where $\hat{v}$ is the first-stage residual from the education equation. $S_v = \sqrt{\exp(z'\hat{\theta}_v)}$ generates the predicted standard deviation used in the control function. Reference category for birth region is Jakarta. Bootstrapped standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Heteroskedasticity Function: Wage Equation (S_u)

	All	Non-poor	Poor
Constant	-1.200***	0.483	-2.435***
	(0.331)	(0.343)	(0.359)
Female	0.412***	0.307***	0.424***
	(0.052)	(0.049)	(0.054)
Ethnicity (Javanese)	0.063	-0.088*	0.134***
	(0.050)	(0.046)	(0.048)
Age	-0.068***	-0.179***	0.027
	(0.021)	(0.021)	(0.022)
Age squared	0.001***	0.003***	-0.001
	(0.000)	(0.000)	(0.000)
Observations	5,939	2,708	3,231

Notes: Estimates from the heteroskedasticity function $\ln(\hat{u}^2) = x'\theta_u$ where $\hat{u}$ is the residual from the wage equation. $S_u = \sqrt{\exp(x'\hat{\theta}_u)}$ generates the predicted standard deviation used in the control function. Birth region is excluded from S_u in the main specification (region excluded from wage equation variance). Bootstrapped standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.10 Robustness: Nonlinear Returns

Table A11: Nonlinear Returns to Education

	Log hourly wage		
	Non-poor	Poor	Diff (ME only)
Panel A: Coefficient Estimates			
Years of schooling	-0.025**	0.016	-0.041***
	(0.012)	(0.011)	(0.016)
Years of schooling^2	0.004***	-0.000	0.004***
	(0.001)	(0.001)	(0.001)
$(S_u/S_v) \cdot \hat{v}$	0.086**	0.227***	
	(0.036)	(0.036)	
Panel B: Marginal Effects at Specified Education Levels			
Observations	2,708	3,231	
Joint p-value (educ, educ^2)	0.000	0.356	
ME at 6 years (SE)	0.025 (0.013)	0.016 (0.013)	0.010 (0.018)
ME at 9 years (SE)	0.051 (0.015)	0.015 (0.014)	0.035* (0.021)
ME at 12 years (SE)	0.076 (0.017)	0.015 (0.017)	0.061** (0.024)
ME at 16 years (SE)	0.110 (0.021)	0.015 (0.020)	0.095*** (0.029)

Notes: Panel A reports Klein-Vella estimates with education and education squared. Panel B reports marginal effects at 6, 9, 12, and 16 years of schooling. Standard errors for marginal effects computed via delta method, ignoring covariance. Bootstrapped standard errors (500 replications) for KV estimation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.11 Robustness: Controlling for Two Endogenous Regressors

The KV framework accommodates multiple endogenous regressors provided they are continuous. We replace the binary poverty indicator with average childhood per capita consumption expenditure (PCE, in $2017 PPP per day) and construct a secondary control function for it. Each additional dollar in childhood PCE raises the returns to schooling by 0.2 percentage points, and marginal effects confirm that returns are lower for those below the poverty line ($2.15), reinforcing our core finding.

Table A12: Robustness: Controlling for Two Endogenous Regressors

	OLS	Klein-Vella
Years of schooling	0.072^{***}	0.033^{***}
	(0.005)	(0.011)
Childhood PCE	-0.009	-0.046^{***}
	(0.013)	(0.017)
Years of schooling × PCE	0.002	0.002^{*}
	(0.001)	(0.001)
$(S_u/S_v) \cdot \hat{v}_1$		0.180^{***}
		(0.050)
$(S_u/S_v) \cdot \hat{v}_2$		0.141^{***}
		(0.039)
Observations	5,939	5,939
Marginal effects at PCE levels:		
At $1		0.035^{***} (0.011)
At $2.15 (poverty line)		0.036^{***} (0.011)
At $4		0.040^{***} (0.011)
At $8		0.046^{***} (0.013)

Notes: Dependent variable is log hourly wage. PCE = daily per capita consumption expenditure in $2017 PPP during childhood (ages 0–14). The Klein-Vella column uses control function estimator with two endogenous regressors (education and PCE). ρ_{educ} and ρ_{pce} are control function coefficients for education and PCE, respectively. Marginal effects calculated as $\beta_{\text{educ}} + \beta_{\text{interaction}} \times$ PCE. Standard errors for marginal effects computed via delta method, ignoring covariance. Bootstrapped standard errors (500 replications) for Klein-Vella estimation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.12 Robustness Check Tables

Table A13: Robustness: Dependent Variable Treatment

	Main (trimmed)			No trim			Winsorized			Raw hours		
	Non-poor	Poor	Diff	Non-poor	Poor	Diff	Non-poor	Poor	Diff	Non-poor	Poor	Diff
Years of schooling	0.068***	0.015**	0.053***	0.069***	0.011	0.058***	0.072***	0.011	0.061***	0.056***	0.014**	0.042***
	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.011)	(0.007)	(0.007)	(0.010)	(0.007)	(0.007)	(0.010)
$(S_u/S_v)\cdot\hat{v}$	0.077**	0.230***	-0.153***	0.093**	0.246***	-0.153***	0.073**	0.246***	-0.173***	0.107***	0.237***	-0.130**
	(0.034)	(0.036)	(0.049)	(0.038)	(0.038)	(0.054)	(0.035)	(0.036)	(0.050)	(0.037)	(0.035)	(0.051)
Observations	2,708	3,231		2,734	3,238		2,734	3,238		2,659	3,212	

Notes: All columns use Klein-Vella estimator. Main = trimmed at top 1%; No trim = no treatment; Winsorized = top 1% winsorized; Raw hours = hourly wages recomputed from unwinsorized weekly working hours (baseline winsorizes hours at 1st and 99th percentiles). NP = Non-poor; P = Poor; Diff = NP minus P. Bootstrapped standard errors (500 replications) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Robustness: Heteroskedasticity Function Specification

	Log hourly wage							
	Main (region excl. from S_u)				Alternative ($x = z$)			
	All	Non-poor	Poor	Diff	All	Non-poor	Poor	Diff
Years of schooling	0.042***	0.068***	0.015**	0.053***	0.047***	0.097***	0.015*	0.081***
	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.008)	(0.012)
$(S_u/S_v) \cdot \hat{v}$	0.183***	0.077**	0.230***	-0.153***	0.159***	-0.035	0.226***	
	(0.037)	(0.034)	(0.036)	(0.049)	(0.043)	(0.040)	(0.043)	
Observations	5,939	2,708	3,231		5,939	2,708	3,231	
Specification:								
Region in S_u	No	No	No		Yes	Yes	Yes	
Region in S_v	Yes	Yes	Yes		Yes	Yes	Yes	
CF Components:								
Mean S_u, [SD]	0.393, [0.041]	0.384, [0.037]	0.394, [0.043]		0.393, [0.049]	0.389, [0.061]	0.396, [0.048]	
Mean S_v, [SD]	1.939, [0.474]	1.474, [0.367]	2.065, [0.412]		1.939, [0.474]	1.474, [0.367]	2.065, [0.412]	
Mean $\hat{v}$, [SD]	-0.000, [3.359]	-0.000, [3.164]	0.000, [3.286]		-0.000, [3.359]	-0.000, [3.164]	0.000, [3.286]	
Mean (S_u/S_v), [SD]	0.218, [0.071]	0.279, [0.085]	0.198, [0.043]		0.215, [0.060]	0.275, [0.054]	0.198, [0.042]	
Mean $(S_u/S_v) \cdot \hat{v}$, [SD]	-0.001, [0.721]	-0.000, [0.847]	-0.000, [0.633]		-0.001, [0.712]	-0.000, [0.834]	-0.000, [0.636]	
Het. Tests:								
BP test p (wage)	0.000	0.000	0.000		0.000	0.000	0.000	
BP test p (educ)	0.000	0.000	0.000		0.000	0.000	0.000	
White test p (wage)	0.000	0.024	0.000		0.000	0.024	0.000	
White test p (educ)	0.000	0.000	0.000		0.000	0.000	0.000	

Notes: Main specification excludes birth region from wage variance function (S_u); Alternative ($x = z$) includes birth region in both S_u and S_v. Diff = Non-poor minus Poor. Bootstrapped standard errors (500 replications) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: Robustness: Alternative Poverty Line Definitions

	Main (\$2.15)			0.8×PL (\$1.72)			1.2×PL (\$2.58)		
	Non-poor	Poor	Diff	Non-poor	Poor	Diff	Non-poor	Poor	Diff
Years of schooling	0.068***	0.015**	0.053***	0.050***	0.012	0.038***	0.087***	0.031***	0.057***
	(0.007)	(0.007)	(0.010)	(0.007)	(0.008)	(0.010)	(0.007)	(0.007)	(0.010)
$(S_u/S_v) \cdot \hat{v}$	0.077**	0.230***	-0.153***	0.132***	0.231***	-0.099*	0.017	0.173***	-0.156***
	(0.034)	(0.036)	(0.049)	(0.034)	(0.038)	(0.051)	(0.036)	(0.036)	(0.051)
Observations	2,708	3,231		3,537	2,402		2,040	3,899	

Notes: All columns use Klein-Vella estimator. Main = \$2.15/day poverty line (2017 PPP); 0.8×PL = \$1.72/day (stricter, fewer classified as poor); 1.2×PL = \$2.58/day (more inclusive, more classified as poor). NP = Non-poor; P = Poor; Diff = NP minus P. Bootstrapped standard errors (500 replications) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A16: Selection into Wage Work: Ability by Employment Type and Poverty Status

	Non-poor			Poor		
	Wage worker	Self-empl.	Norm. diff.	Wage worker	Self-empl.	Norm. diff.
Years of schooling	11.77	10.58	0.35***	9.83	8.70	0.33***
	(3.39)	(3.43)		(3.46)	(3.40)	
Raven's z-score	0.21	0.17	0.04	0.08	0.05	0.04
	(0.94)	(0.89)		(0.92)	(0.87)	
Numeracy z-score	0.18	0.15	0.04	-0.04	-0.07	0.04
	(1.06)	(1.01)		(0.95)	(0.91)	
Observations	3,011	1,144		3,605	1,594	
N (cognitive)	3,009	1,143		3,604	1,593	

Notes: Sample includes all workers aged 15–35 (wage workers and self-employed) with non-missing childhood poverty status, not currently enrolled in school. The wage-worker counts exceed the main regression sample because this table does not require non-missing wages or education. Self-employed includes own-account workers and employers. Wage workers include government employees, private employees, and casual workers. Raven's z-scores and numeracy z-scores are standardized within age groups. Norm. diff. = Normalized difference, calculated as $(\bar{X}_W - \bar{X}_S)/\sqrt{(s_W^2 + s_S^2)/2}$; values $| > 0.25|$ suggest meaningful imbalance (Imbens and Rubin, 2015). Stars on normalized differences denote significance from two-sample t-tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N (cognitive) reports the number of individuals with non-missing Raven's test scores. Source: IFLS waves 1–5.

A.13 Full Estimation Results

Table A17: Full Estimation Results: Returns to Education by Childhood Poverty Status

	OLS			Klein-Vella		
	All	Non-poor	Poor	All	Non-poor	Poor
Years of schooling	0.080^{***}	0.088^{***}	0.058^{***}	0.042^{***}	0.068^{***}	0.015^{**}
	(0.003)	(0.005)	(0.005)	(0.008)	(0.007)	(0.007)
$(S_u/S_v) \cdot \hat{v}$				0.183^{***}	0.077^{**}	0.230^{***}
				(0.037)	(0.034)	(0.036)
Age	0.104^{***}	0.147^{***}	0.094^{***}	0.141^{***}	0.172^{***}	0.139^{***}
	(0.023)	(0.035)	(0.030)	(0.008)	(0.008)	(0.008)
Age squared	-0.002^{***}	-0.002^{***}	-0.001^{**}	-0.002^{***}	-0.003^{***}	-0.002^{***}
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Female	-0.302^{***}	-0.232^{***}	-0.377^{***}	-0.265^{***}	-0.211^{***}	-0.346^{***}
	(0.023)	(0.032)	(0.032)	(0.019)	(0.017)	(0.018)
Javanese	0.067^{***}	0.064^{*}	0.063^{*}	0.066^{***}	0.066^{***}	0.057^{***}
	(0.025)	(0.036)	(0.034)	(0.020)	(0.020)	(0.020)
Java - Village	-0.373^{***}	-0.321^{***}	-0.362^{***}	-0.414^{***}	-0.331^{***}	-0.370^{***}
	(0.044)	(0.053)	(0.091)	(0.040)	(0.038)	(0.037)
Java - Small town	-0.291^{***}	-0.155^{***}	-0.378^{***}	-0.310^{***}	-0.153^{***}	-0.375^{***}
	(0.049)	(0.059)	(0.098)	(0.043)	(0.040)	(0.041)
Java - Big city	-0.223^{***}	-0.169^{**}	-0.258^{**}	-0.210^{***}	-0.153^{***}	-0.236^{***}
	(0.067)	(0.085)	(0.118)	(0.054)	(0.054)	(0.049)
Off-Java - Village	-0.397^{***}	-0.323^{***}	-0.390^{***}	-0.429^{***}	-0.326^{***}	-0.384^{***}
	(0.043)	(0.053)	(0.089)	(0.038)	(0.036)	(0.036)
Off-Java - Small town	-0.325^{***}	-0.178^{***}	-0.458^{***}	-0.320^{***}	-0.161^{***}	-0.438^{***}
	(0.049)	(0.055)	(0.099)	(0.042)	(0.040)	(0.041)
Off-Java - Big city	-0.264^{***}	-0.187^{***}	-0.393^{***}	-0.236^{***}	-0.171^{***}	-0.341^{***}
	(0.066)	(0.071)	(0.142)	(0.055)	(0.049)	(0.053)
Observations	5,939	2,708	3,231	5,939	2,708	3,231
Het. Tests (p-values):						
BP test (wage)				0.000	0.000	0.000
BP test (educ)				0.000	0.000	0.000
White test (wage)				0.000	0.024	0.000
White test (educ)				0.000	0.000	0.000

Notes: Dependent variable: log hourly wage. Standard errors in parentheses. OLS columns use robust standard errors. KV columns show bootstrapped standard errors (500 replications). Klein-Vella specification excludes birth region from the wage variance function (S_u). Reference category for birth region is DKI Jakarta. Het. Tests report p-values from Breusch-Pagan and White tests for heteroskedasticity. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.