The Impact of Birth Weight on Cognitive and Non-Cognitive Skills in Children: Evidence from Ethiopia, India, Peru, and Vietnam

Karina Gabitova

University of South Florida

Abstract

This paper investigates the causal effect of birth weight on cognitive and non-cognitive outcomes in children aged eight, utilizing longitudinal data from the Young Lives survey in Ethiopia, India, Peru, and Vietnam. To address potential endogeneity in birth weight and reduce bias in Ordinary Least Squares (OLS) estimates, we employ an Instrumental Variables (IV) approach, allowing for a more precise estimation of birth weight's impact on child development. Our IV results reveal that a 1% increase in birth weight—a proxy for improved health at birth—is associated with a 0.71% increase in Peabody Picture Vocabulary Test (PPVT) scores, emphasizing the critical role of early health in shaping cognitive skills. In contrast, we find limited evidence that birth weight significantly influences non-cognitive outcomes, such as self-esteem, suggesting these traits may be predominantly shaped by socio-cultural, environmental, and familial factors. Additionally, the analysis underscores the substantial role of socio-economic variables, particularly parental education and household wealth, in determining cognitive outcomes, often surpassing the direct impact of birth weight. These findings have important policy implications for public health and education sectors in developing countries, indicating that efforts to enhance early-life health should be complemented by broader socio-economic interventions to foster holistic child development.

Introduction

Low birth weight has long been recognized as a significant determinant of increased mortality risks among newborns and poorer developmental outcomes across various stages of life, including adverse health, educational, and labor market outcomes, as well as cognitive skills such as IQ. Approximately 20% of newborns worldwide are born with a birth weight below the threshold of 2,500 grams, which presents a substantial challenge for resource-limited developing countries (WHO, 2014). As a measure of poor health at birth, low birth weight imposes constraints on children's learning capacities, often resulting in poorer educational and economic outcomes in comparison to children born with a healthy weight.

While extensive research has investigated the long-term consequences of poor health at birth and its effects on adult outcomes, there is a relatively limited understanding of its impact on developmental outcomes during mid-childhood. Existing studies suggest that children born with higher birth weights generally perform better on cognitive tests during mid-childhood, with these effects being particularly pronounced among vulnerable groups, such as girls in rural areas or those born to less-educated mothers. However, the influence of low birth weight on non-cognitive skills, such as self-esteem, during mid-childhood and adolescence remains largely underexplored within the economics literature. The period between the ages of 5 and 8 is particularly important, as it represents a developmental window where policy interventions aimed at enhancing child development can yield lasting effects (Almond, Currie, & Duque, 2018). Consequently, understanding how birth weight influences both cognitive and non-cognitive skills is essential for formulating policy strategies that can effectively improve children's developmental trajectories.

This study seeks to fill this gap in the literature by examining the effects of low birth weight on both non-cognitive skills, such as self-esteem, and cognitive skills, including performance on vocabulary and mathematics tests, among 8-year-old children in four developing countries: Ethiopia, India, Peru, and Vietnam. The study will also explore potential heterogeneous effects by gender, age, household income, and parental education.

The data for this analysis is drawn from the Young Lives longitudinal survey, which follows two cohorts of children: one aged approximately 1 year at the start of the survey, and another around 8 years old. This survey offers a rich set of information on early childhood health, including birth weight, gestational age, and the health status of newborns, along with later academic performance (measured by the Peabody Picture Vocabulary Test and numeracy tests) and responses to self-esteem related questions.

The causal relationship between birth weight and various developmental outcomes may be subject to bias due to the endogeneity of birth weight, sample selection, and unobserved heterogeneity when estimated using ordinary least squares (OLS) regression. This study addresses these potential biases by employing an instrumental variable (IV) approach, utilizing maternal height and preterm birth status as instruments for birth weight. The endogeneity of birth weight has been extensively documented, particularly in developed countries, where twin or sibling fixed effects models are commonly used to control for such biases. However, the application of similar methods in low- and middle-income countries remains limited (Currie & Vogl, 2013; Nandi et al., 2017). In these settings, the effects of birth weight may vary due to contextual factors such as restricted access to quality education, inadequate infrastructure, gender biases, and socio-economic disparities, which can impede the conversion of early health advantages into human capital accumulation and long-term economic outcomes. Additionally, parental investment in children may be influenced by the child's health status, with healthier children likely receiving more resources and attention, which can further confound the estimates (Almond & Mazumder, 2013; Mani, 2012; Anand et al., 2018).

The OLS regression results suggest a positive association between birth weight and cognitive ability, with a 1% increase in birth weight corresponding to a 0.62% increase in the Peabody Picture Vocabulary Test (PPVT) score. However, this estimate may suffer from downward bias due to unobserved factors, such as parental health behaviors, that could simultaneously affect both birth weight and cognitive outcomes. By utilizing IV estimation, this bias is reduced, and the results indicate a stronger relationship, with a 1% increase in birth weight leading to a 0.71% increase in the PPVT score. This comparison highlights the importance of employing IV methods to obtain a more accurate estimate of the causal effect of birth weight on cognitive outcomes.

In contrast, no statistically significant relationship was found between birth weight and non-cognitive outcomes, such as self-esteem. This lack of significance may stem from the inherent difficulty in defining and measuring self-esteem, which, unlike cognitive abilities, is more subjective and influenced by a wide range of factors, such as family dynamics, societal expectations, and cultural norms, that are challenging to capture in the data. Consequently, these findings suggest that the influence of birth weight on non-cognitive outcomes may be more indirect or weaker, reflecting the complex nature of non-cognitive skill development.

Literature Review

A substantial body of research examines the impact of birth weight on economic and cognitive outcomes, particularly in high-income countries where longitudinal data on twins or siblings are available. These studies generally find a positive association between birth weight and later-life outcomes. For example, Figlio et al. (2014) used administrative data from Florida to show that a 10% increase in log of birth weight was associated with a 0.044 standard deviation increase in test scores among students aged 9 to 14, independent of household characteristics and school quality. Similarly, Behrman and Rosenzweig (2004) employed data from 804 monozygotic twin pairs in Minnesota and found that a 1 lb. increase in birth weight correlated with 0.7 additional years of schooling, a 7% increase in earnings, and a 0.6-inch height increase by mid-adulthood. These results emphasize the long-term economic and educational benefits associated with higher birth weight.

Internationally, studies reveal similar patterns. In Norway, Black et al. (2007) found that a 10% increase in birth weight corresponded to a 0.06-point rise in IQ, a 1% increase in earnings, and a nearly 1% higher likelihood of completing high school. In the Philippines, Baguet and Dumas (2019) reported that a 100-gram increase in birth weight was associated with an increase in school grade completion by 0.019 standard deviations. In Chile, Bharadwaj et al. (2018) found that a 10% increase in birth weight improved math test scores by 0.02-0.04 standard deviations among children aged 6 to 18. Torche and Echevarria (2011) used twin fixed-effects models with Chilean data and found that a 400-gram increase in birth weight yielded a 15% increase in math scores among fourth graders. Additionally, Nakamuro et al. (2013), using Japanese data and a twin fixed-effects approach, found positive effects of birth weight on academic performance at age 15, though there was no significant effect on earnings or years of schooling.

The literature also suggests that twin fixed-effects models, while effective in some cases, may lead to underestimations of birth weight effects due to the frequent low birth weights and premature births among twins. Kumar et al. (2022) highlighted this limitation, finding that twins tend to have lower birth weights, which could lead to underestimations of the impact of birth weight on cognitive and economic outcomes for singletons. Some studies have sought alternative methods, such as instrumental variable (IV) approaches, to address birth weight endogeneity. For instance, Lin and Liu (2009) used the number of doctors and public health budgets as instruments for birth weight in Taiwan, finding that higher birth weights were positively associated with school performance, though effects were stronger for children born to younger or less-educated mothers. Conversely, Lin et al. (2017) proposed using genetic markers (SNPs) and twin status as instruments for birth weight and found no significant impact on educational attainment among Chinese children in Hong Kong. These results highlight the need for further methodological refinement, especially in developing countries.

Research on non-cognitive outcomes, particularly self-esteem, has predominantly come from pediatrics and psychology rather than economics. For instance, Levy-Shaft et al. (1994) conducted a study in Israel and found that low-birth-weight children exhibited higher levels of anxiety, depression, and aggression, alongside lower self-esteem as measured by the Tennessee Self-Concept Scale. Similarly, Saigal et al. (2002) examined self-esteem in low-birth-weight adolescents using the Harter Self-Perception Profile, finding significant deficits in athletic performance but no substantial differences in other domains of self-perception. Nair et al. (2009) studied low-birth-weight adolescents in India and found that they had significantly lower self-esteem and intelligence scores in comparison to normal-birthweight peers. These studies indicate a correlation between low birth weight and compromised non-cognitive outcomes; however, they often lack methodologies that fully account for endogeneity, such as IV approaches or sibling fixed effects.

To date, few studies have applied the instrumental variable approach or twin fixed-effects models in estimating the effects of low birth weight on non-cognitive skills, particularly within economic frameworks. Further research utilizing these methodologies is needed to better understand the broader effects of birth weight on both cognitive and non-cognitive development, especially in developing-country contexts where resources and early-life conditions differ substantially from high-income settings.

Methodology

The effect of birth weight on cognitive and non-cognitive outcomes can be analyzed using an Ordinary Least Squares (OLS) model, where each observation represents a child i in household j within sentinel (or cluster) s in country c. The model is as follows:

$$Y_{ijsc} = \beta_0 + \beta_1 \mathrm{BWijsc} + \beta_2 \mathrm{Cijsc} + \beta_3 \mathrm{H}_{\mathrm{js}} + \theta_{\mathrm{s}} + \mu_{\mathrm{ijsc}} \quad (1)$$

where $\mathrm{Y}_{\mathrm{ijsc}}$ denotes the dependent variable, representing either cognitive skills—measured as the log of the PPVT (Peabody Picture Vocabulary Test) scores—or non-cognitive skills, measured by self-esteem. The main independent variable, BWijsc, denotes the log of birth weight. Cijsc represents child characteristics, H_{js} captures household characteristics, θ_{s} includes sentinel (cluster) fixed effects, and μ_{ijsc}represents idiosyncratic error terms.

The model incorporates a range of child-level and household-level control variables to account for individual and environmental factors influencing developmental outcomes. Child-level controls include age and gender, providing basic demographic context for each observation. Household-level variables include years of parental education as an indicator of parental human capital and its potential influence on child development. Household wealth is calculated as the composite average of three indices representing housing quality, access to essential services, and ownership of consumer durables, offering a nuanced measure of socio-economic status. Additional household-level controls encompass religion, ethnicity, and locality type (rural or urban), which are included to capture socio-cultural and regional differences that may impact child outcomes.

The primary challenge in identifying the causal effect of birth weight on cognitive and non-cognitive skills arises from unobserved heterogeneity, specifically genetic and environmental factors that are potentially correlated with both birth weight and developmental outcomes. Such correlations can bias estimates. The medical literature provides substantial evidence that maternal health behaviors during pregnancy—such as cigarette smoking and alcohol consumption—are linked to reduced birth weight and other adverse physical outcomes in newborns. These behaviors are more common among individuals of lower socioeconomic status and education levels, who are also generally less likely to invest in their children's education.

Parental health behaviors have also been shown to influence non-cognitive outcomes in children; for instance, maternal smoking during pregnancy is associated with heightened symptoms of depression and anxiety in offspring (Moylan et al., 2015). Consequently, if certain parental health behaviors positively influence both birth weight and cognitive or non-cognitive outcomes, then the OLS estimate of β_1 in equation (1) may overestimate the causal effect of birth weight on test scores and self-esteem. This issue has often been addressed in the economics literature using twin or sibling fixed effects, as in Figlio et al. (2014); however, this approach is unfeasible with the YL survey data, which samples only one child per household.

An alternative solution is to employ an instrumental variable (IV) approach using a two-stage least squares (2SLS) model:

$$\text{BWijsc} = \alpha_0 + \alpha_1 \text{Zijsc} + \alpha_2 C_{ijsc} + \alpha_3 H_{jsc} + \theta_s + \phi_c + \eta_{ijsc} \quad (2)$$

$$Y_{ijsc} = \beta_4 + \beta_5 \text{BWijsc} + \beta_6 \text{Cijsc} + \beta_7 H_{jsc} + \theta_s + \phi_c + \nu_{ijsc} \quad (3)$$

In the first-stage equation (2), the log of birth weight is regressed on instrumental and exogenous variables. Following Kumar et al. (2022), the mother's height and an indicator for premature birth are used as instruments for birth weight. These instruments are assumed to be exogenous, affecting math and vocabulary scores and self-esteem solely through their impact on birth weight.

In the second-stage equation, the outcome variables are regressed on the predicted birth weight obtained from the first stage, along with other exogenous covariates. To satisfy the conditions for a valid instrumental variable approach: (1) the relevance condition requires that the instruments are correlated with the endogenous variable (birth weight is correlated with maternal height and premature birth status); (2) the exclusion restriction requires that the instruments affect the outcomes only through birth weight (i.e., mother's height and premature birth status influence cognitive and non-cognitive outcomes only via birth weight); and (3) the independence assumption ensures that the instruments are uncorrelated with other explanatory variables, such as child age, gender, birth order, household wealth, parental education and religion, and locality type, establishing that no other causal pathways exist between the instruments and the outcome variables beyond birth weight.

In this study, the Heckman correction is further applied to address potential selection bias associated with the availability of birth weight data. This approach begins with a first-stage probit model that estimates the probability of observing birth weight data in the dataset. Key predictors in this probit model include variables capturing the type of birth (e.g., hospital versus home) and the presence of official birth weight documentation, as these factors are indicative of socioeconomic resources and healthcare access. The outcome of this estimation is the inverse Mills ratio derived as the ratio of the probability density function to the cumulative density function associated with the probit model, which quantifies the likelihood of birth weight data being observed due to unobserved factors.

In the second stage, the inverse Mills ratio derived from the probit model is integrated into the original instrumental variables (IV) regression model. This inclusion effectively adjusts for self-selection bias by controlling for factors that may influence both the recording of birth weight and subsequent investments in a child's human capital. By including this correction term, the model accounts for the possibility that households with recorded birth weights may have systematically greater resources or motivation to support a child's cognitive and non-cognitive development, thus addressing the risk of overestimating the effect of birth weight on later-life outcomes.

This methodological refinement enhances the validity of the IV estimates, ensuring a more accurate assessment of the causal impact of birth weight on developmental outcomes, particularly in contexts where selection bias related to data availability could otherwise confound results.

Data

The analysis for this paper utilizes data from the Young Lives longitudinal survey, which offers a rich, multi-dimensional dataset of two cohorts of children in four low- and middle-income countries: Ethiopia, India, Peru, and Vietnam. Initiated in 2002, Young Lives has conducted follow-up interviews approximately every three years to capture the dynamic impacts of poverty on child development, health, education, and socio-economic conditions. The survey design, which samples children across 20 sentinel sites in each country, ensures representation across diverse regional and socio-economic contexts, allowing for greater insight into how environmental and regional differences influence child outcomes.

The dataset consists of two cohorts in each country: a younger cohort of approximately 2,000 children aged 6 to 18 months at recruitment in 2002, and an older cohort of about 1,000 children aged 7 to 8 years old. For the purposes of this study, I restrict the sample to the younger cohort, as only this group includes data on birth events, specifically birth weight and prematurity indicators, which are crucial for examining the role of early health indicators on later developmental outcomes. Data on these birth characteristics were collected in the initial 2002 survey round, providing a robust baseline for studying the effect of early health conditions on outcomes in middle childhood.

Young Lives collects a wide range of information across multiple dimensions, including household socio-economic status, parental education, and household demographics, along with detailed child-specific data on nutrition, health behaviors, education, cognitive skills, physical development, and subjective well-being. This study focuses on two primary developmental outcomes measured when the younger cohort reached approximately eight years of age in the third survey round: cognitive outcomes, assessed through the Peabody Picture Vocabulary Test (PPVT) scores, and non-cognitive outcomes, particularly self-esteem measures (see appendix).

Maternal height, an important control for prenatal health and nutritional status, is only recorded in the fifth survey round; however, as adult height is a fixed characteristic, this timing does not present issues for the analysis. Additional control variables, including child-specific and household characteristics, are taken from the initial 2002 survey round, allowing for a comprehensive analysis that accounts for early household conditions and demographic factors.

The design of Young Lives, with its longitudinal approach and inclusion of critical health, educational, and economic variables across different developmental stages, provides a robust dataset for analyzing the long-term impacts of birth weight on both cognitive and non-cognitive outcomes. This study aims to leverage this rich dataset to contribute to our understanding of how early-life health factors interact with broader socio-economic conditions to shape developmental trajectories in middle childhood within low- and middle-income country contexts.

Results

Table 1 presents descriptive statistics for key variables in the analysis, focusing on child development outcomes and background characteristics. These include cognitive and non-cognitive abilities, child demographics, parental education, and family socioeconomic status. The first column shows that, on average, children in the sample exhibit moderate variation in cognitive abilities, with the mean PPVT (Peabody Picture Vocabulary Test) score indicating a diverse range of cognitive capabilities. The PPVT scores, which measure reading comprehension, range from 0 to 5.28 within the sample. The self-esteem score, with a mean of 2.994 and a standard deviation of 0.32, suggests limited variation in self-esteem levels among the children. Approximately 14.4% of the sample were born prematurely, a key variable in the analysis of birth outcomes.

Table 2 reports results from a linear regression analysis examining the relationship between cognitive ability and independent variables such as birth weight, gender, ethnicity, parental education, wealth index, and region. Cognitive ability is measured by the log of the PPVT score. While socio-economic predictors have a substantial effect, the residual variance indicates that other unobserved factors significantly influence cognitive outcomes, suggesting that factors beyond those captured in the model also contribute to variations in cognitive ability (Torche & Echevarría, 2011).

Table 3 presents the results of a two-stage least squares (2SLS) regression analysis, addressing potential endogeneity concerns. This analysis examines the relationship between birth weight and cognitive ability, using maternal height and preterm birth as instrumental variables (IVs). Preterm birth is strongly negatively correlated with birth weight, confirming its validity as an instrument in the first stage. The model is robust, with an F-test value of 18.77, well above the traditional threshold of 10, indicating that the instruments are sufficiently strong. Additionally, the first stage passes the under-identification test with a Kleibergen-Paap rk LM statistic of 16.622, which is significantly higher than the critical value, further validating the strength of the instruments. In the second stage, the coefficient for birth weight is statistically significant at the 10% level, suggesting that a 1% increase in birth weight leads to a 0.71% increase in cognitive ability, as measured by the PPVT score. The presence of endogeneity is confirmed by a highly significant p-value, and the Sargan test indicates that the instruments are not over-identified, confirming their validity.

Table 4 expands upon the two-stage least squares (2SLS) regression by incorporating the Inverse Mills Ratio (IMR) to correct for potential selection bias in the availability of birth weight data. The IMR is derived from a first-stage probit model, where the probability of observing birth weight data is regressed on the availability of birth weight documentation and the type of birth (hospital vs. home). The analysis uses maternal height and preterm birth as instrumental variables. While the inclusion of the IMR does not significantly alter the results, the model confirms that birth weight has a significant effect on cognitive ability. The first-stage model's intercept is highly significant, and the coefficients for maternal height and preterm birth are robust, with preterm birth showing a strong negative correlation with birth weight. The F-test results indicate that the instruments are relevant, and the Sargan test again supports the validity of the instruments. The analysis confirms that birth weight has a positive and significant impact on cognitive ability.

Table 5 examines the relationship between independent variables and self-esteem, focusing on the effect of birth weight. The coefficient for birth weight is not statistically significant, suggesting that birth weight does not have a meaningful impact on self-esteem. This finding is important as it implies that factors other than birth health, such as socio-cultural and environmental influences, are more likely to shape self-esteem. Notably, the negative coefficient for region suggests that the area of residence may be associated with lower self-esteem, potentially due to varying social norms, access to resources, or environmental factors. Other variables, such as parental education and schooling, have little effect on self-esteem, though socio-cultural factors such as religion significantly influence it. This result aligns with existing literature (Nakamuro et al., 2013; Baguet & Dumas, 2019), which suggests that self-esteem is more strongly influenced by socio-cultural factors than by early health indicators such as birth weight.

Table 6 presents a two-stage least squares (2SLS) regression analysis with self-esteem as the outcome variable, using instrumental variables to address potential endogeneity. The first-stage results indicate that the intercept is highly significant, and both maternal height and preterm birth serve as valid instruments for birth weight. Preterm birth is particularly strong, with a significant negative correlation with birth weight. However, the F-test value of 7.2 is slightly below the ideal threshold of 10, indicating room for improvement in instrument relevance. The second-stage results show that the coefficient for birth weight is not statistically significant, suggesting that birth weight does not significantly affect non-cognitive abilities such as self-esteem. This supports the idea that non-cognitive outcomes are shaped more by environmental and socio-cultural factors than by early physical health.

Post-estimation tests further validate the model. The Sargan over-identification test indicates that the instruments are valid and uncorrelated with the error term. The Durbin-Wu-Hausman test confirms the presence of endogeneity, justifying the use of 2SLS over OLS for unbiased estimation. These results are consistent with previous studies, such as those by Kumar et al. (2023), which found that while birth weight is an important early health indicator, its direct impact on non-cognitive skills like self-esteem is less significant compared to socio-environmental influences.

Table 7 presents results from the 2SLS regression analysis including the Inverse Mills Ratio to account for selection bias in the availability of birth weight data. The first-stage results confirm the validity of the instruments, with maternal height and preterm birth as strong predictors of birth weight. The F-test and Kleibergen-Paap statistics indicate that the instruments are sufficiently strong. The second-stage results reveal that birth weight does not significantly affect non-cognitive outcomes such as self-esteem, further supporting the conclusion that socio-cultural factors play a more substantial role in shaping these outcomes than early physical health.

The findings suggest that while birth weight has a significant impact on cognitive ability, its effect on non-cognitive outcomes such as self-esteem is minimal. This highlights the complex interplay between early health indicators and socio-environmental factors in shaping children's developmental outcomes. The use of instrumental variables and correction for selection bias strengthens the validity of these results, underscoring the need for careful consideration of both health and socio-cultural influences in developmental research.

Conclusion

This study examines the effects of birth weight on cognitive and non-cognitive skills in children aged eight across Ethiopia, India, Peru, and Vietnam, using data from the Young

Lives longitudinal survey. The primary objective was to address a gap in existing literature by investigating the impact of low birth weight on mid-childhood outcomes, with particular emphasis on non-cognitive skills such as self-esteem—a topic that has received relatively less attention in research to date. By employing both Ordinary Least Squares (OLS) and Instrumental Variables (IV) approaches, the study provides a robust analysis of how birth weight influences child development within diverse socio-economic and cultural contexts.

The findings indicate a significant positive relationship between birth weight and cognitive outcomes, as measured by vocabulary and reading comprehension test scores. These results highlight the crucial role of early health indicators, such as birth weight, in shaping a child's cognitive abilities. Children with higher birth weights tend to perform better in cognitive assessments, reinforcing the importance of maternal and prenatal health interventions. However, the study found limited evidence suggesting a strong relationship between birth weight and non-cognitive outcomes, such as self-esteem. This suggests that while birth weight is a key determinant of cognitive development, non-cognitive skills are more influenced by environmental, socio-cultural, and familial factors (Bharadwaj, Lundborg, et al., 2018).

Furthermore, the study reveals that socio-economic factors—such as parental education and household wealth—have a significant influence on cognitive outcomes, often outweighing the impact of birth weight alone. These factors were consistently associated with better performance in cognitive tests, underscoring the complex interplay between biological factors like birth weight and the broader socio-economic environment. The use of instrumental variable analysis further supports this finding, demonstrating that while birth weight is an important determinant of cognitive skills, it operates in conjunction with other critical socio-economic and familial influences to shape a child's developmental trajectory (Anand et al., 2021).

The implications of these findings are wide-reaching and have important policy and economic considerations, particularly in developing countries. First, the study underscores the importance of maternal and child health interventions to improve birth weight, thereby potentially enhancing cognitive development and educational outcomes for children. By prioritizing maternal nutrition, access to prenatal care, and early childhood health programs, policymakers can address low birth weight and improve long-term educational and economic outcomes for children, contributing to the reduction of intergenerational poverty. These interventions can create a foundational basis for better educational achievements and a more skilled workforce, ultimately supporting economic growth and social mobility.

Moreover, the study's findings on the limited impact of birth weight on non-cognitive skills like self-esteem point to the complex relationship between biological and environmental factors in child development. While health interventions that focus on early physical health are essential, they must be complemented by broader efforts that address socio-cultural and familial factors, such as family dynamics, educational support, and community environments. This suggests that policies aimed at improving non-cognitive skills should focus not only on health-related aspects but also on creating positive, supportive environments that promote self-esteem, resilience, and social skills in children (Torche & Echevarría, 2011; Moylan et al., 2015).

From an economic perspective, the importance of both cognitive and non-cognitive skills cannot be overstated, as these abilities directly influence an individual's future productivity and ability to contribute to society. Interventions that enhance cognitive and non-cognitive outcomes are likely to have far-reaching economic benefits, particularly in terms of reducing social inequality and improving human capital. Investing in child development, therefore, is not only a matter of equity and social justice but also an economic imperative.

However, the study has several limitations that warrant consideration. One of the primary limitations is the reliance on self-reported data for certain variables, which may introduce potential bias. Additionally, the study's focus on children at age eight may not fully capture the entire developmental trajectory of cognitive and non-cognitive skills, as these abilities continue to evolve throughout childhood and adolescence.

Future research should consider incorporating genetic factors and exploring how these interact with environmental influences on both cognitive and non-cognitive development.

Integrating genetic data into future studies could provide valuable insights into the hereditary traits that contribute to variations in child development, particularly in children with low birth weight. Furthermore, examining the impact of postnatal interventions, such as early childhood education programs and family support initiatives, could offer important perspectives on how these interventions mitigate the potential disadvantages associated with low birth weight.

Another key avenue for future research is the exploration of regional and cultural differences in the impact of birth weight on child development. While this study focused on four countries, extending the research to include a broader range of countries with varying economic conditions, healthcare infrastructures, and cultural contexts could enhance our understanding of how birth weight influences developmental outcomes in different settings. Such an expansion would also help identify context-specific interventions that might be more effective in certain environments. Additionally, understanding the subjective experiences of children with low birth weight—particularly in terms of their self-esteem and non-cognitive skills—could provide valuable insights to inform more targeted and effective policy interventions.

In conclusion, this study provides important evidence on the role of birth weight in shaping cognitive and non-cognitive outcomes in children, particularly in the context of developing countries. The findings have significant policy implications for maternal and child health, as well as for educational and socio-economic planning, with the potential to improve long-term developmental outcomes and reduce poverty across generations. Policymakers should recognize the complex interplay between biological, socio-economic, and cultural factors in shaping children's developmental trajectories and design policies that target both early health outcomes and the broader socio-cultural environment.

Table 1. Descriptive Statistics

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
Cognitive skills	Log of PPVT raw score at age 8	7447	4.168	0.50	0.00	5.28
Non-cognitive skills	Self-esteem measure	7381	2.994	0.32	1.00	4.00
Child’s birth weight (BW)	Log of BW	4766	8.016	0.19	6.91	8.70
Maternal height	Height in cm	2050	152.185	5.95	112.9	182
Prematurity	Binary variable (1=yes, 0=no)	8062	0.144	0.35	0	1
Rural residence	Binary variable (1=yes, 0=no)	8062	0.626	0.48	0	1
Female	Binary variable (1=yes, 0=no)	8062	0.481	0.50	0	1
Child's age	Age in months	8062	11.663	3.45	4.00	22.00
Father education	Father's level of education	7211	5.473	4.91	0.00	29.00
Mother education	Mother's level of education	7897	4.58	4.72	0.00	28.00
Wealth index	Wealth index	8029	0.376	0.23	0.00	0.96
Inverse Mills Ratio	Inverse Mills Ratio	7242	0.608	0.58	0.00	1.92

Table 2. OLS Results (Cognitive)

Log(PPVT)	Coef.	St.Err.	t-value	p-value	Sig
Log(BW)	0.12	0.034	3.57	0.0000	***
Region	0.002	0.001	3.6	0.0000	***
Cluster	-0.009	0.001	-8.13	0.0000	***
Rural	0.018	0.016	1.13	0.2600	
Female	-0.05	0.012	-4.05	0.0000	***
Ethnicity	-0.001	0	-1.88	0.0610	*
Religion	0.034	0.002	16.43	0.0000	***
Child's age	0.016	0.002	9.12	0.0000	***
Father educated	0.007	0.002	3.99	0.0000	***
Mother educated	0.01	0.002	5.22	0.0000	***
Wealth index	0.567	0.039	14.72	0.0000	***
Constant	2.538	0.277	9.17	0.0000	***
R-squared	0.333				
F-test	176.421				
Prob > F	0.000				
Number of obs	3902				

Table 3. 2SLS Results (Cognitive)

	Model 01	Model 02
	1st Stage	2nd Stage
Constant	8.043***	-1.813
	(0.0611)	(3.069)
Log(BW)		0.709*
		(0.382)
All variables in main specification	Yes	Yes
Observations	3902	3902
Instrumental variables	Coff.	t-stats
Maternal height	-6.59e-06*	1.756
	(3.70e-06)	
Prematurity	-0.119***	4.594
	(0.0259)	
R-squared		0.151
Observations	825	825
Post-estimations test for instrumental variables:		
Predictive power partial R2		
Robust F-test	18.77	15.31
p-value	0.000	0.000
underindentification test		
Kleibergen–Paap rk LM statistic	19.482	32.449
p-value	0.0001	0.000
Weakindentification test		
Kleibergen–Paap Wald rk F statistic	16.622	16.622
10% maximal IV size	11.24	19.93
Overidentification test		
Sargan statistic		1.964
Chi-sq.(3) p-value		0.1611
Endogeneity test		
Durbin–Wu–Hausman tests		
Chi-sq.(1) p-value	0.002340	

Table 4. 2SLS with IMR Results (Cognitive)

	Model 01	Model 02
	First Stage	2nd Stage
Constant	8.064***	-1.802
	(0.0639)	(3.064)
Log(BW)		0.716*
		(0.380)
All variables in main specification	Yes	Yes
Observations	3902	3902
Instrumental variables	Coff.	t-stats
Maternal height	-6.36e-06*	1.702
	(3.71e-06)	
Prematurity	-0.119***	4.576
	(0.0260)	
R-squared		0.124
Observations	722	722
Post-estimations test for instrumental variables:		
Predictive power partial R2		
Robust F-test	17.31	14.21
p-value	0.000	0.000
underindentification test		
Kleibergen–Paap rk LM statistic	19.534	32.7
p-value	0.000	0.000
Weakindentification test		
Kleibergen–Paap Wald rk F statistic	16.736	16.736
10% maximal IV size	19.9	19.93
Overidentification test		
Sargan statistic		1.731
Chi-sq.(3) p-value		0.1882
Endogeneity test		
Durbin–Wu–Hausman tests		
Chi-sq.(1) p-value	0.0223	

Table 5. OLS Results (Non-cognitive)

Self-Esteem	Coef.	St.Err.	t-value	p-value	Sig
Log(BW)	-0.018	0.027	-0.67	0.501	
Region	-0.005	0.001	-9.21	0	***
Cluster	0.002	0.001	2.59	0.01	***
Rural	-0.025	0.013	-1.93	0.053	*
Female	0.006	0.01	0.65	0.518	
Ethnicity	0	0	1.62	0.105	
Religion	-0.006	0.002	-3.98	0	***
Child’s age	0.002	0.001	1.25	0.211	
Father educated	0.002	0.001	1.13	0.26	
Mother educated	0.003	0.001	1.82	0.069	*
Wealth Index	0.042	0.03	1.39	0.164	
Constant	3.272	0.218	15.01	0	***
R-squared	0.107				
F-test	42.548				
Prob > F	0.000				
Number of obs	3907				

Table 6. 2SLS Results (Non-cognitive)

	Model 01	Model 02
	1st Stage	2nd Stage
Constant	8.054***	3.595*
	(0.0590)	(2.123)
Log(BW)		-0.0684
		(0.264)
All variables in main specification	Yes	Yes
Observations	3907	3907
Instrumental variables	Coff.	t-stats
Maternal height	-6.42E-06*	1.743
	(3.69e-06)	
Prematurity	-0.12***	4.633
	(0.0259)	
R-squared		0.151
Observations	832	832
Post-estimations test for instrumental variables:		
Predictive power partial R2		
Robust F-test	7.2	6.32
p-value	0.000	0.000
underindentification test		
Kleibergen–Paap rk LM statistic	23.31	34.273
p-value	0.0001	0.000
Weakindentification test		
Kleibergen–Paap Wald rk F statistic	17.593	20.593
10% maximal IV size	11.473	19.93
Overidentification test		
Sargan statistic		2.392
Chi-sq.(3) p-value		0.1219
Endogeneity test		
Durbin–Wu–Hausman tests		
Chi-sq.(1) p-value	0.0012	

Table 7. 2SLS with IMR Results (Non-cognitive)

	Model 01	Model 02
	1st Stage	2nd Stage
Constant	8.079***	3.528*
	(0.0608)	(2.116)
Log(BW)		-0.0504
		(0.262)
All variables in main specification	Yes	Yes
Observations	744	744
Instrumental variables	Coff.	t-stats
height2	-6.16E-06*	1.660
	(3.70e-06)	
preterm	-0.121***	4.670
	(0.0259)	
R-squared		0.151
Observations	3907	3907
Post-estimations test for instrumental variables:		
Predictive power partial R2		
Robust F-test	6.85	6.22
p-value	0.000	0.000
underindentification test		
Kleibergen–Paap rk LM statistic	19.859	34.556
p-value	0.0001	0.000
Weakindentification test		
Kleibergen–Paap Wald rk F statistic	17.724	17.724
10% maximal IV size	11.472	15.93
Overidentification test		
Sargan statistic		2.923
Chi-sq.(3) p-value		0.0873
Endogeneity test		
Durbin–Wu–Hausman tests		
Chi-sq.(1) p-value	0.0341	

References

Almond, D., Currie, J., 2011. Killing Me Softly: The Fetal Origins Hypothesis. Journal of Economic Perspective. 25(3), 153-72.

Almond, D., Currie, J., Duque, V., 2018. Childhood circumstances and adult outcomes: Act II. Journal of Economic Literature. 56(4), 1360-1446.

Almond. D., Mazumder, B., 2013. Fetal origins and Parental Responses. Annual Review of Economics, 5(1), 37-56

Anand, P., Behrman, J.R., Dang, Hai-Anh H., Jones, S., 2018. Varied Patterns of CatchUp in Child Growth: Evidence From Young Lives. Social Science and Medicine. 214, 206-13.

Baguet, M, Dumas, C., 2019. How does birth weight affect health and human capital? A short- and long-term evaluation. Health Economics 28(5), 597-617.

Behrman, J.R., Rosenzweig, M., 2004. Returns to birthweight. Review of Economics and Statistics. 86(2), 586-601.

Bharadwaj, P., Eberhard, J., Neilson, C. 2018. Health at Birth, Parental Investments and Academic Outcomes. Journal of labor Economics. 36(2), 349-394

Black, S.E., Devereux, P.J., Salvanes, K.G., 2007. From the Cradle to the Labor Market? The Effect of Birth Weight on Adult 2XWFRPHV´ Quarterly Journal of Economics. 122(1), 409-439.

Campbell, J., 1998. Book review: Peabody picture vocabulary test. Journal of Psychoeducational Assessment 16(4), 334±338.

Clifford, Angela, Linda Lang, and Ruoling Chen. "Effects of maternal cigarette smoking during pregnancy on cognitive parameters of children and young adults: a literature review." *Neurotoxicology and teratology* 34.6 (2012): 560-570.

Crookston, B. T., Schott, W., Cueto, S., Dearden, K. A., Engle, P., Georgiadis, A., Lundeen, E. A., Penny, M. E., Stein, A. D. & Behrman, J. R., 2013. Postinfancy growth, schooling, and cognitive achievement: Young Lives. The American Journal of Clinical Nutrition, 98(6), 1555±1563.

Currie, J., Vogl, T., 2013. Early-Life Health and Adult Circumstance in Developing Countries. Annual Review of Economics. 5, 1±36.

Figlio, David, et al. "The effects of poor neonatal health on children's cognitive development." *American Economic Review* 104.12 (2014): 3921-3955.

Figlio, D., Guryan, J., Karbownik, K., Roth, J. 2014. The Effects of Poor Neonatal Health on &KLOGUHQ¶V Cognitive Development. American Economic Review. 104(12), 3921- 3955

Kumar, S. 2016. The effect of birth order on schooling in India. Applied Economics Letters. 23, 1325-1328.

Kumar, Santosh, et al. "Birth Weight and Cognitive Development during Childhood: Evidence from India." *Economic Papers: A journal of applied economics and policy* 41.2 (2022): 155-175.

Levy-Shiff, Rachel, et al. "Emotional and behavioral adjustment in children born prematurely." *Journal of Clinical Child Psychology* 23.3 (1994): 323-333.

Lin, M.J., Liu, J.T., 2009. Do lower birth weight babies have lower grades? Twin fixed effect and instrumental variable method evidence from Taiwan. Social Science Medicine. 68(10), 1780-1787.

Lin, S.L., Leung, G.M., Schooling, C.M., 2017. The Effect of Birth Weight on Academic Performance: Instrumental Variable Analysis. American Journal of Epidemiology, 185(9), 853-859.

Mani, S., 2012. Is there Complete, Partial, or No Recovery from Childhood Malnutrition? - Empirical Evidence from Indonesia. Oxford Bulletin of Economics and Statistics, 74(5), 691-715.

Moylan, S., Gustavson, K., Øverland, S. *et al.* The impact of maternal smoking during pregnancy on depressive and anxiety behaviors in children: the Norwegian Mother and Child Cohort Study. *BMC Med* **13**, 24 (2015). https://doi.org/10.1186/s12916-014-0257-4

Nair, M. K., et al. "Low birthweight babies-outcome at 13 years." *Indian Pediatr* 46.Suppl (2009): s71-s74.

Nakamuro, M., Uzuki, Y., Inui, T., 2013. The effects of birth weight: does fetal origin really matter for long-run outcomes? Economics Letters. 121(1), 53-58.

Nandi, A., Lutter, R., & Laxminarayan, R., 2017. Breastfeeding Duration and Adolescent Educational Outcomes: Longitudinal Evidence From India. Food and Nutrition Bulletin. 38(4), 528-541.

Oreopoulos, P., Stabile, M., Walld, R., Roos, L.L., 2008. Short-, medium-, and long-term consequences of poor infant health: an analysis using siblings and twins. Journal of Human Resources. 43(1), 88-138.

Paxson, C. & Schady, N., 2007. Cognitive development among young children in Ecuador the roles of wealth, health, and parenting. Journal of Human Resources 42(1), 49±84.

Saigal, Saroj, et al. "Self-esteem of adolescents who were born prematurely." *Pediatrics* 109.3 (2002): 429-433.

Torche, F., Echevarria, G., 2011. The effect of birthweight on childhood cognitive development in a middle-income country. International Journal of Epidemiology. 40(4), 1008-18.

WHO, 2014. Global Nutrition Targets 2025: Low Birth Weight Policy Brief. World Health Organization.

Appendix

The survey incorporates 34 items that probe the emotional states of children across various dimensions, notably interpersonal interactions, particularly familial and peer relationships, as well as their problem-solving capabilities, motivation, ambition, and self-perception. These items are structured as statements and employ a 4-point Likert scale ranging from "strongly disagree" to "strongly agree," deliberately omitting a neutral midpoint to elicit clear positional responses. To specifically analyze feelings associated with self-esteem as distinct from the quality of social relationships, 17 questions (statements) have been selectively utilized:

1) If someone opposes me, I can find the means and the ways to get what I want.
2) I am as good as most other people.
3) When I am confronted with a problem, I can usually find several solutions.
4) Overall, I have a lot to be proud of.
5) If I am in trouble, I can usually think of a solution.
6) I am confident that I could deal efficiently with unexpected events.
7) I can do things as well as most people.
8) I can always manage to solve difficult problems if I try hard enough.
9) It is easy for me to stick to my aims and accomplish my goals.
10) If I have children of my own, I want to bring them up like my parents raised me.
11) I can remain calm when facing difficulties because I can rely on my coping abilities.
12) A lot of things about me are good.
13) I can usually handle whatever comes my way.
14) Thanks to my resourcefulness, I know how to handle unforeseen situations.

15) When I do something, I do it well.

16) I can solve most problems if I invest the necessary effort.

17) In general, I like being the way I am.

For the construction of a variable quantifying feelings directly linked to self-confidence, the aggregate of responses to these 17 questions is computed. Each response is assigned a numerical value—1 for 'strongly disagree', 2 for 'disagree', 3 for 'agree', and 4 for 'strongly agree'. The total of these values is then averaged by the number of responses to account for any non-responses.

This resultant variable, a composite self-esteem score, functions as the dependent variable in our analysis. Elevated scores on this metric signify higher self-esteem, denoting a positive self-assessment and enhanced confidence. This measurement is pivotal in assessing how environmental factors influence adolescent self-esteem across diverse settings, thereby offering insights into their psychological resilience and overall mental health in the context of adverse conditions. The detailed enumeration of the survey questions can be found in the appendix for further reference.