Child Labour, Human Capital and Beliefs

Devis Decet Kaman Lyu

June 30, 2024*

Abstract

In contexts where child labour is pervasive, household decisions about allocating children's time between school and work involve a trade-off: current returns from child labour versus future returns from education. This paper tests for the existence of a third factor: future returns from child labour, as parents view farm work as an investment in agricultural skills. We provide evidence on each component of this trade-off in the context of rural Ghana by leveraging four waves of survey data on 5,000 households, exogenous shocks to agricultural productivity, and a vignette survey design to elicit parental beliefs.

JEL-Classification: O12, I2, J13, J2.

1 Introduction

Child labour is pervasive in developing countries. In fact, the share of children working is rising in some sub-Saharan African countries and is predominantly concentrated in agriculture (ILO, 2020). Household decisions on how to allocate children's time between school and work

*Devis Decet, Northwestern University (email: ddecet@u.northwestern.edu). Kaman Lyu, Northwestern University (email: kamanlyu2025@u.northwestern.edu). We would like to thank Christopher Udry, Lori Beaman, Seema Jayachandran, Dean Karlan, and Nancy Qian for their guidance. We are thankful to the researchers and the staff of the Global Poverty Research Lab and the Institute of Statistical, Social and Economic Research of the University of Ghana who collected the data; we are especially thankful to Albert Wilson, Samuel Amedzro, Bernice Karbu, Enoch Vukey, Joana Kwawu, Lydia Boahen, Kaitlin Obeng-Konadu, Tracy Commodore, Samuel Ampaw, Francis Dompae, Fred Dzanku, and Andew Nickow for their efforts in the data collection and management process. For helpful comments, we are grateful to Marie-Louise Decamps, Andrew Dillon, Matthew Honnor, Edoardo Teso, Evgeni Rachkovski, David Hope, Elizabeth Jaramillo, Eugenia Frezza, Adeline Delavande and Ryu Matsuura, as well as to the participants of the Development Advising Group, Applied Micro Lunch, and the Development Lunch Seminar, Development Seminar at Northwestern University, Public Policy Research Network Graduate Workshop and Novafrica 2024.

can have negative implications on human capital formation and a child's future earning ability as an adult (Baland and Robinson, 2000; Beegle, Dehejia, and Gatti, 2009). Economists predominantly analyse child labour through the lens of human capital theory (Schultz, 1960): parents weigh the future returns to schooling against the current opportunity cost of schooling, which includes the foregone economic contribution of children. However, little is known about the existence of a third component of this trade-off: parents may perceive that there are future returns to child labour.

This paper focuses on children working on family farms and tests for the existence of this third component. Parents may believe that working on farms is an investment in agricultural skills which can have higher returns than schooling if the child were to become a farmer. In this sense, child labour on family farms is akin to an apprenticeship (Mokyr, 2019). Parents are allocating children to work instead of school because they view child labour as an investment in farming-specific human capital. Human capital is multi-dimensional (Lise and Postel-Vinay, 2020; Cunha and Heckman, 2007) and the returns to each dimension differ among occupations. Indeed, formal schooling may have low returns for farming (Barker, 2022). In fact, work on the farm alongside parents may be important for transmitting agricultural skills (Bass, 2004; Rosenzweig and Wolpin, 1985).

Testing this hypothesis is challenging as it requires measuring children's time allocation between school and work on the farm, as well as parental beliefs. In this paper, we address this challenge using a combination of rich data and empirical approaches. We leverage four waves of survey data on 5,000 households in rural Ghana, plausibly exogenous shocks to harvest value, and a vignette survey design to elicit parental beliefs. This allows us to (i) describe the relationship between child labour supply and schooling, and (ii) shed light on the role of current returns to child labour, future returns to schooling, and future returns to child labour in shaping such allocations.

We first present descriptive facts on child labour prevalence and intra-household inequality in time allocation and human capital. We develop an accurate measurement of child labour in which children *self*-report economic activity, and compare it with parents-reported child labour supply. 15% of children aged 5 to 11 and 35% of children aged 12 to 18 reported having worked on the farm in the week prior to the interview. We also provide information on when they work relative to school, hours worked, regular tasks, and age they first started work. While most children combine school and working, there is a clear sorting pattern: children who have higher cognitive ability tend to have better educational attainment and to work less on the farm, after taking household fixed effects and controlling for age and sex. Moreover, consistent with the existing literature (Jayachandran and Pande, 2017), firstborns exhibit better test scores, are physically larger, work less, and attend school more often than their younger siblings. Additionally, we show that parents diversify occupations across children. Despite being extremely aspirational in general, they believe that children with higher cognitive ability are more likely to become an employee for the government or a private firm, and less likely to become a farmer.

Having shown the relevance of intra-household inequality in child labour and schooling, in the second part of the paper we turn to investigating which factors drive parental decisions on human capital investments. We organize our analysis around the trade-off between child labour and schooling. We shed light on each of the three components of this trade-off: current returns to child labour, future returns to schooling, and future returns to child labour. We start by providing evidence on the role current returns to child labour. Specifically, we study the impact of positive, transitory shocks to household agricultural productivity. These shocks increase the short-term returns of child labour but at the same time reduce the marginal utility of the additional income. We are able to isolate the exogenous component of these shocks relying on an instrumental variable for the value of harvest produced by a household in a year. Specifically, we build the instrument by using Lasso to select among a large set of plot-level land characteristics, weather conditions, and their interactions (Belloni, Chen, et al., 2012; Belloni and Chernozhukov, 2013). We find that a 10% increase in the value of harvest raise the probability that children work on the farm by 1.3%, and reduce the likelihood that they attend school by 0.37%. This suggests that the substitution effect dominates the income effect. Positive shocks to current returns to child labour induce parents to put their children to work more and spend less time in school. Moreover, we find suggestive evidence that such shocks can have persistent effects. Children whose household experienced a 10% increase in harvest value are 4.3% more likely to be a farm worker eight years later.

We then present evidence on the role of future returns to schooling as proxied by higher cognitive skills. We exploit the panel structure of our data, and study the relationship between each child's characteristics in 2018 and the observed time allocation in 2022. This helps to assuage concerns about reverse causality whereby contemporaneous children cognitive ability are likely to be affected by school attendance. We show that children with higher cognitive skills are indeed less likely to work on the farm and more likely to attend school four years later.

Finally, we investigate whether parents perceive child labour as an investment in human capital to increase their children's future agricultural productivity. Using three hypothetical scenarios via a vignette, we find that parents believe that a child who works on family farms will be 1.6 times more productive as a farmer than if they had attained two more years of schooling with no farming experience. In fact, they believe this child will be *twice* as productive than if they had ran errands in their spare time (instead of farming) and had no additional schooling. Then, we show that parents behave consistently with these beliefs. Children who have parents with high perceived returns to child labour in farming and are designated to become a farmer, work more on the farm and substitute away from school. This effect is concentrated in children aged 12–18; in the context of Ghana, this is the age in which parents decide whether children should invest in formal education and try to enroll in Senior Secondary School. In a placebo exercise, we show that perceived returns to child labour in farming do not predict the time allocation for children who are designated to become a teacher or an employee. Overall, these results suggest that parents do perceive future returns to child labour for farming, and that they behave consistently with these beliefs by choosing the time allocation that maximises the relevant dimension of human capital.

Our findings contribute to several strands of literature. First and foremost, we contribute to a large body of literature that studies the determinants of child labour[1]: poverty (Edmonds

[1]See Edmonds, Schultz, and Strauss, 2008 for a comprehensive review

and Schady, 2012; Canagarajah and Nielsen, 2001; Basu and Van, 1998); parental preferences (Ejrnæs and Pörtner, 2004; Basu, 1999); credit market imperfections (Baland and Robinson, 2000); imperfect labour markets or lack of available adults or skilled labour (Edmonds and Theoharides, 2020; Doepke and Zilibotti, 2005); and imperfect altruism or incomplete contracting problems (Bau et al., 2021; Ashraf et al., 2020; Banerjee, 2004). Similar to Bau et al. (2021), Shah and Steinberg (2017), and Atkin (2016), we find that positive wage shocks can increase child labour supply. While these papers usually employ district-level wage shocks, the richness of our data enables us to show this using *household*-specific agricultural productivity shocks. We are particularly close to the papers which stress how child labour have negative implications on human capital formation and schooling outcomes (Baland and Robinson, 2000; Shah and Steinberg, 2017; Bau et al., 2021). However, in contrast to these papers, we demonstrate that parents may perceive positive implications of child labour on human capital formation. Consistent with Rosenzweig and Wolpin (1985), we highlight how child labor can be a way to acquire agricultural skills. Crucially, we are able to elicit actual parental beliefs and show how they impact the time allocation of their own children. Similar to Sviatschi (2022), we stress the importance of occupation-specific human capital. In our context, the investment decision depends on parental beliefs on a child's future occupation, and not on market conditions attracting children into work and out of school. By providing both parents- and child-reported measurements of child labour, we also speak to a growing body of literature that documents the widespread variation in child labour statistics in large-scale national surveys of developing countries (Dillon et al., 2012; Guarcello et al., 2010) and how parents tend to under-report child labour (Lichand and Wolf, 2023).

We also contribute to the body of work which studies intra-household inequality (Chiappori and Meghir, 2015; Berman, Rotunno, and Ziparo, 2020; Carneiro, Rasul, and Salvati, 2023) and to the literature focusing on the role of subjective beliefs as a determinant of parents' human capital investment (Cunha, Elo, and Culhane, 2013; Attanasio, Boneva, and Rauh, 2022; Giannola, 2024). We bridge these two literatures by stressing how parental beliefs on future occupations and on the production of occupation-specific human capital can shape intra-household inequality.

Finally, we speak to the large literature on education policy in developing countries (Mbiti et al., 2019; Das et al., 2013; Duflo, Dupas, and Kremer, 2021). Our results are complementary to Allen IV (2022), who shows how a greater overlap between school and farming calendars reduces the time available for both schooling and farm-based child labour. In light of our findings, a smaller overlap may allow parents to invest in farming-specific human capital without harming a child's formal education. More generally, as parental beliefs on returns to farm work seem rooted in cultural norms, this paper also joins the growing body of work focusing on how policies should be adapted to context-specific culture (Bau, 2021; Ashraf et al., 2020; Moscona and Seck, 2021).

The rest of the paper is as follows. Section 2 characterises the setting of child labour in Ghana, describes the data, and details our measurement of child labour. Section 3 documents descriptive facts on intra-household allocations on chilren's school and work. Section 4 tests for determinants of child labour, and presents evidence on the role of current returns to child labour, future returns to schooling, and future returns to child labour. Finally, Section 5 concludes.

2 Context & Data

2.1 Setting

There are currently 265 million working children worldwide—this is almost 17% of the global child population (ILO, 2015). Whilst most of the world have been facing a downward trend in recent years, the rate of child labour remains persistently high in Africa (Ortiz-Ospina and Roser, 2016).[2] In fact, some sub-Saharan African countries have seen an *increase* in the absolute number and share of children working since 2012 (ILO, 2020). To highlight the lack of progress pertinent to this region, ILO (2020) states that "there are now more children in child labour in sub-Saharan Africa than in the rest of the world combined". The high rate of child

[2] See Figures B.1 and A.2 for the share of children in employment across continents and countries over time.

labour in sub-Saharan Africa may be due to its dependence on an agrarian economy. Indeed, 85% of child labour is in agriculture (ILO, 2016). More specifically, 72% of all child labour (and 83% of all working children aged 5 to 11) occur on family farms and businesses (ILO, 2020). This form of family-oriented child labour is the focus of this paper.

In Ghana, which is the setting of this paper, 28.5% of all children participate in some form of economic activity (GSS, 2014). In particular, 26.7% of children aged 5 to 14 and 48.1% of children aged 15 to 17 are working—mostly as unpaid workers on family farms or enterprises. Crucially, 82% of working children combine school with work. It is not the case that parents are taking children out of school completely to focus solely on labour—it is far more common that children are perhaps missing some classes to work or predominantly working after school, on the weekend or during school holidays.

Education in Ghana is compulsory for Primary and Junior Secondary School (ages 4 to 15) but there is low enforcement. The primary enrolment rate is 87%, although one can expect their actual enrolment rate to be far lower (UNESCO, 2016). There is a large body of evidence that education matters (e.g., Jensen, 2010; Duflo, 2001; Krueger and Lindahl, 2001). Moreover, parents in our sample *do* value education and have high perceived returns to completing Senior Secondary School (SSS) to not only their child's own future earnings, but also on marital and fertility outcomes. Figure B.2 shows that almost 90% of all parents desire their child to attend SSS. They believe that attendance would not only double their child's own future monthly earnings, but also allow them to match with a future spouse with earnings that are twice as high. Moreover, they believe that attending SSS would delay the age of marriage by six years and lead to having three fewer children. These parental beliefs therefore suggests that it is not the case that the presence of child labour is due to parents under-valuing the importance of education.

Child labour laws are in existence in Ghana. The minimum working age is 15, and for light work, it is 13. However, enforcement is inconsistent and ineffective. There is very low awareness as even law enforcement authorities are unfamiliar with child labour laws (GSS, 2003).

Underlying these statistics are important cultural norms. In Ghanaian society, children are expected to work and to contribute towards household income (GSS, 2003). Parents value education and children are therefore expected to be at school most of the time, but outside of school are expected to assist in domestic chores or help their parents on the farm. The cost of education makes up a large share of household expenditure. In Wave 4 (2022), the average household reports spending 1,487 Ghanaian cedis annually on educational expenses for each household member (GSPS, 2010).[3] This is equivalent to eight percent of total household consumption for each household member who attends school. Hence, children are expected to contribute to paying for it. More generally, child labour is seen as a passage to becoming a responsible, productive adult. Indeed, a child is considered "a deviant, lazy, or having poor upbringing" if they cannot perform basic chores (GSS, 2003).

These norms are not unique to Ghana. In many countries, most work undertaken by children has been explained as a form of education, training and play (Nieuwenhuys, 1996). In many African countries, there is a strong emphasis on the idea that children should help the family (Adu-Gyamfi, 2014), and that child labour has the cultural significance of evading idleness (Adonteng-Kissi, 2018). Some studies show that parents who are more deeply rooted in their culture tend to involve their children in riskier work (Adonteng-Kissi, 2018). Working and investing in human capital can coexist. However, when children work, it can conflict with their schooling because both generally occur during daylight hours. Some children might be allocated work over school due to parental preferences, a lack of nearby schools, or an inability to afford fees, leaving work as their only productive option (Bass, 2004). Allen IV (2022) shows empirically that the duration of overlap between the school calendar and the farming season influences the time children allocate to work and study. Anecdotal evidence from Mali, Chad, and Togo indicate that communities acknowledge this issue and respond by establishing community schools with flexible schedules that accommodate essential farming activities (Bass, 2004).

[3]These expenditures include school fees, Parent-Teacher Association fees, uniforms, books and school supplies, transport, food and board, extra classes and in-kind expenses.

In communities with limited opportunities for formal employment and skilled jobs, parents may value more practical forms of education. This could be particularly true when the traditional schooling system fails to meet their expectations or when their children struggle academically. In this context, parents may see children's involvement in work not only as a way to contribute to household income but also as a way to acquire important skills. Thus, this is a form of education complementary to formal schooling (Bass, 2004). In this sense, child labour can be seen akin to an apprenticeship

2.2 Data

We design and implement a survey module for 5,000 households in 2022 as part of Wave 4 of the EGC-ISSER-Northwestern Ghana Socioeconomic Panel Survey (GSPS, 2010). GSPS (2010) is a large-scale, nation-wide survey following households and their members, as well as communities through community-based questionnaires, from 2010 to 2022. The survey is administered every four years and currently has four waves. We have a sample size of 300 communities consisting of approximately 3,500 rural and 1,500 urban households from all 16 regions of Ghana. Exploiting the panel structure of the data, we are able to trace individuals from their childhood to the onset of early adulthood. We restrict the sample to all children aged 5 to 18 at the time of the survey to measure child labour on family farms and domestic work.

Our survey has three broad aims. First, we want to accurately measure child labour on family farms and in domestic work. We capture both parents- and self-reported child labour to improve truthful reporting as responses given by parents or by proxy are often understated due to social desirability bias (Lichand and Wolf, 2023; Dillon et al., 2012; Guarcello et al., 2010). In addition to broad measures of child labour participation, we also elicit information on when they work relative to school, hours worked, regular tasks, whether they are paid, and the age they first started work.

Second, we survey parents with a rich set of questions to test for intra-household work and investment allocations across siblings based on both present and future outcomes.

We capture parents' occupational aspirations for their children by asking them to rank, for each child, the likelihood of attaining the following four occupations in the future: Farm Owner, Business Owner or Market Trader, Teacher, and Private or Public Sector Employee. We select these four occupations to give us an informative, but not exhaustive, list of feasible occupations for poor rural families in Ghana.[4] This list is in ascending order of both desirability (i.e., remuneration and status) and ambition (i.e., ability and qualifications required to attain it). Specifically, completing at least Senior Secondary School is required to become a teacher or an employee in the public sector.

Third, we design a vignette to elicit parents' perceived returns to child labour in farming. Specifically, we ask the household head to rank future agricultural productivity of a hypothetical child under three scenarios. These scenarios differ only on the time allocation (schooling versus work on the farm) of the child. We describe in detail the vignette when we present the results in Section 4.3.

2.3 Measuring Child Labour

Definition We define child labour as any work performed on the farm or any domestic chore. The main sample are all children aged 5 to 18 at the time of the survey.[5] Specifically, for farm work we ask children: *Did you do any work, for pay or for free, or assist on a farm during the last seven days (even if it was for only one hour)?* For domestic chores, we ask children: *Did you do any domestic work or household chores during the last seven days (even if it was for only one hour)?* Our definition is broad and does not necessarily have a negative connotation as we are interested in a form of work that can co-exist with schooling and happens mostly within the household. In Appendix A we compare our questions and results with other large surveys in developing countries.

[4]We also selected occupations which are gender neutral. For example, we omitted community health workers as they tend to be female dominated.

[5]We choose age 18 because some of our survey questions relate to the past year and thus we aim to capture all individuals up to age 17 in the past year.

Summary Statistics Table 1 documents summary statistics for children aged 5 to 18 in Ghana in 2022.[6] 25% of children self-report working on a farm in the last seven days and, on average, for 13.2 hours a week. This is in contrast to parents reporting that 24% of children are a contributing worker on their family farm *in the last year*. Parents also report that their children work, on average, 12.7 hours a week. Appendix Figure A.3 compares the spread of hours worked reported by parents versus self-reported by the child. 76% of children self-report domestic work or household chores in the last seven days and, on average, for almost eleven hours a week.

In the survey, we also ask children aged 12 to 18 whether they were paid for their farm work and, if so, how much in Ghanaian cedis in the last seven days. Table B.1 shows that only three percent of children report being paid and 23 of these children report being paid an average of 58 Ghanaian cedis for a week's farm work. This is consistent with the context of child labour in rural Ghana where children largely work on family farms without pay, rather than as causal labourers on other people's land. Column (3) shows that children aged 12 to 18 report that they first started farm work at age 10, with some starting as young as 4. This stresses the importance of asking children to self-report child labour as age 10 is substantially lower than the age range in child labour surveys. The Ghana Living Standards Survey and the DHS Survey, for example, tend to target children aged 12 and above.[7] This could therefore lead to missing vital data from younger children who work.

Children also self-report the time of work for farming and domestic chores. Figure 1 shows that 17% of children report working on farms after school, 34% on the weekend, and 35% during school holidays. 9% report having permanently left school to work on a farm. 2% report working before school and, surprisingly, only 1% report working during class time.[8]

[6] All the results described in this section rely on data from Wave 4 of the GSPS (2010). This is the only wave for which the questions about work participation were asked directly to children. Moreover, this is the only wave for which we have information on which tasks and in which times children were working.

[7] We only ask more detailed child labour questions to children aged twelve and above because we followed the format of these other surveys.

[8] We expected this number to be much higher given the low school attendance rate. However, children may be under-reporting work during school time for social desirability bias because they understand that truancy could be frowned upon by international institutions.

The time of domestic work follows closely with that of farm work—12% of children report doing domestic work before school, 29% after school, 27% on the weekend, and 26% during school holidays. 4% report having permanently left school to carry out domestic chores and, again, only 1% report working during class time.

We also document regular tasks that children self-report doing during farming and domestic work (Figure 2). Figure 2a shows that the most frequent agricultural activities are planting seeds, weeding, and harvesting crops, occupying almost 80% of all tasks for children. In line with children having a comparative advantage working with crops such as cotton (Bau et al., 2021; Levy, 1985), we find that children perform tasks that are compatible with their smaller stature and weaker physical strength—such as being closer to the ground for planting seeds and weeding, and having "nimble fingers" to pluck crops during harvest. The remaining 20% of agricultural activities is split between post-harvest processing, e.g., cleaning and storing crops (8%), clearing and land preparation (7%), chemical application, e.g., spraying fertilisers and pesticides (4%), and ploughing (2%).

3 Descriptive Facts

In this section, we show the relevance of intra-household inequality in child labour supply and educational investments. We leverage rich microdata to present three descriptive facts. First, we show that there is a clear sorting between child labour and schooling *within* a family. Second, firstborns tend to work less and have more schooling than their younger siblings. Third, parents - while being highly aspirational on average - diversify the future occupations of their children.

3.1 Empirical Strategy

We exploit within-household variation to estimate child-specific determinants of child labour supply on family farms and domestic activity. Taking household fixed effects, we are able to interpret estimates as increments relative to the child's siblings. Moreover, household fixed effects enable us to eliminate selection bias arising from differences across households—e.g., wealth, ethnicity, religion, parental attitudes, and geography. A notable caveat is that this empirical specification cannot rule out reverse causality. Hence, results should be taken as descriptive facts.

We estimate the following fixed effects regression.

$$y_{ih} = \alpha + X'_{ih}\beta + W'_{ih}\delta + \gamma_h + \epsilon_{ih}, \tag{1}$$

where the outcome variable of interest, y_{ih}, is schooling or child labour for child i in household h. The independent variables of interest, represented by matrix X_{ih}, are divided into the following categories: demography (i.e., age, sex, birth order); ability (i.e., cognitive ability, Maths and English test scores); educational attainment (i.e., years of schooling); land inheritance (i.e., if the child will be bequeathed any household plots); and financial dependence in old age (i.e., if the child will be financially responsible for either parent).

W_{ih} is a matrix of controls for age, sex, and whether the child is a firstborn.[9] This enables us to parse out effects driven by family structure to isolate determinants of child labour. For instance, older children will mechanically have more years of schooling; by controlling flexibly for age, one can estimate the conditional effect of educational attainment on child labour supply relative to their siblings. γ_h signifies household fixed effects, α is the constant, and ϵ is the error term. We cluster standard errors at the household level[10].

3.2 Results

Fact 1: Sorting between School and Child Labour. We find that older and male children of lower birth order and with fewer years of schooling are more likely to be involved in farm work relative to their siblings. On the other hand, older and female children tend to do domestic work. See Appendix Table B.2. There is a clear sorting of children into school versus

[9]Naturally, these controls are not included when estimating the demographic determinants of child labour supply.

[10]The results described in this section rely on data from Wave 4 in 2022 of the GSPS (2010). This is the only wave for which the questions about work participation were asked directly to children. Moreover, this is the only wave for which we have information on parental beliefs about children's future occupations.

work *within* the same household—siblings with high ability (proxied by cognitive ability, and Maths and English test scores) have more years of schooling and are less likely to work on farms, controlling for age, sex, and birth order. Naturally, we cannot rule out reverse causality, where children's time allocation influences their cognitive ability. In Section 4.2, we use lagged child characteristics—which are unaffected by current time allocations—to show that parents seem to be aware of their children's traits and allocate those with higher cognitive ability to more school. Columns (1) and (2) of Table 2 show that an additional year of schooling decreases farm work by 0.7 to 1.3 percentage points, relative to their siblings, controlling for age, sex, and birth order. Similarly, column (4) shows that a one standard deviation rise in their Ravens' Matrices and Digit Span Test z-scores (these are common cognitive ability tests), increases educational attainment by 0.12 and 0.38 additional years of schooling. Moreover, a one standard deviation increase in their Maths and English test z-scores are associated to 0.26 and 0.09 additional years of schooling, respectively. These estimates, with the exception of English, are statistically significant at the 1% and 5% significance levels. Column (3) shows that children with an additional year of schooling are 1.2 percentage points more likely to engage in domestic work relative to their siblings. This may be a result of "inequality aversion" where parents try to allocate a fair burden of chores across their children—i.e., children who go to school more have to help out more at home, and children who work more instead of school can do less domestic labour. Importantly, domestic work coincides less with the school day whereas farming requires daylight and thus substitutes time from school. Hence, it is not the case that high-ability children are not without any responsibilities or chores—even those who attend school more than their siblings have to perform some household tasks. This is very much in line with the cultural norms detailed in Section 2.1 in which child labour is considered to be a necessary passage to a mature and productive life.

Fact 2: Parents Heavily Invest in Firstborns. Firstborns are overwhelmingly favoured within families in some developing countries (see Jayachandran and Pande, 2017). Table B.3 shows that, on average, relative to their younger siblings, firstborns are 2.9 percentage points less likely work (column 1), attain almost a quarter of a year more schooling (column 2), are 0.11 standard deviations larger for their age (column 3), and have 0.03 to 0.12 higher z-scores for cognitive, Maths and English tests (columns 4 to 7). Moreover, Figure 3 illustrates the z-scores residualised by household fixed effects and controlling for age and sex. It shows that the firstborn attains higher within-household deviations in z-scores for their anthropometric index[11] compared to the second- and third-born. There is a clear pattern of birth order being associated with more favourable outcomes.

Fact 3: Parents Diversify Occupations across Children. Parents are highly aspirational. We ask them to rank, for each child, the likelihood of attaining the following four occupations in the future: Farm Owner, Business Owner or Market Trader, Teacher, and Private or Public Sector Employee. We made sure to ask parents to rank occupations realistically (rather than ideally) to circumvent capturing desires that are not grounded in reality. Figure 4 is a stacked graph showing the full distribution of parents' occupational choices for their children. 61% of parents select the most ambitious occupation—a private or public sector employee—as the most likely occupation their child will attain in the future.[12] Conversely, the least aspirational occupation—a farm owner—is picked last for 79% of parents.

These are astounding beliefs. Parental beliefs on occupations in our sample are consistent[13] with Duflo, Dupas, and Kremer (2021) who survey secondary school students in Ghana to elicit their occupational beliefs. They find that, at baseline, 70% of students believed that they would be a government employee or teacher by age 25. Yet, when Duflo, Dupas, and Kremer (2021) revisit these students now aged 26, they find that only 6% have made it as a government employee or teacher. The overwhelming majority either returned to their villages to become farm workers or market traders, or are currently unemployed.

Table B.4 gives the determinants for parents ranking farm owner and either teacher and

[11] This is an unweighted index of the z-scores for height, weight, waist size, hip size, and arm circumference, by age and sex.

[12] Parents' first-choice occupations for their children are as follows: 4% for farm owner, 17% for business owner, 18% for teacher, and 61% for private or public sector employee.

[13] 79% of parents choose a private or public sector employee or teacher as their first choice.

private or public sector employee as the most likely occupation. Column (2) shows that, controlling for age, sex, and birth order, a one standard deviation increase in the Digit Span Test z-score decreases the likelihood of being designated to become a farm owner by 1.8 percentage points—which is 90% of the mean. Similarly, a one standard deviation increase in the English Test z-score decreases the likelihood of parents choosing farm owner as the first-choice occupation by 1.1 percentage points, or 55% of the mean.

On the other hand, a one standard deviation increase in the z-scores for Ravens' Matrices, Digit Span Test, Maths and English tests, raise the likelihood of parents earmarking teacher or a private or public sector employee as the most likely occupation by 1.9 to 5.3 percentage points (column 4, Table B.4).

Figure 5b compares the probability of parents choosing each of the four occupations as the most likely for their child if their child currently works on family farms. When a child is working on the farm, they are 7.4 percentage points more likely to have farm owner selected as their most likely future occupation. Similarly, they are 18.8 percentage points less likely to have private or public sector employment as their parents' first choice. On the other hand, children who are in school are 5.8 percentage points less likely to have farm owner selected as their most likely future occupation, and 16.5 percentage points more likely to have private or public sector employment as their parents' first choice (figure 5a).

Taken together, this shows that parents are in fact forming beliefs for their child's future occupations that are grounded in their child's ability (as measured by cognitive, numeracy and literacy skills) and that these beliefs are related to current investment allocations (as measured by school attendance and child labour on farms).

4 Determinants of Child Labour and Schooling

After having shown the relevance of intra-household inequality in child labour and schooling, this section focuses on uncovering the determinants of parental decisions on human capital investments.

Framework The starting point of our analysis is the trade-off emphasized by standard models of human capital investment. On one hand, higher currents returns from child labour make it more valuable to have the child working. This is because the child can bring home higher income. On the other hand, higher future returns to education make it more valuable to have the child in school. This is because in the future the child will be able to gain higher income; as parents are either altruistic or will rely on the child for their income when old, they internalize the utility from this higher income in the future.

In our analysis, we depart from the standard models in two ways. First, we augment the trade-off with a third component - the future returns to child labour. If parents perceive that child labour can make a child more productive as an adult, this increases the benefits from having the child working instead than in school. Second, returns to different time-allocation can vary depending on a child's future occupation.

Given the characteristics of our context, our analysis focuses on the choice between child-labour on family farms and schooling. Thus, the trade-off described above delivers the following implications. Higher agricultural productivity increases the income that the household can attain by having the child working on the farm - this should increase a child's labour supply. A child with higher returns from education should instead spend more time in school. On the contrary, if parents perceive that working on the farm will increase a child future agricultural productivity, children designated to become a farmer in the future should work more.

We employ this framework to guide our empirical analysis and interpret our results. We separately study the role of current returns to child labour, future returns to schooling, and future returns to child labour. First, in Subsection 4.1, we use harvest shocks to analyse the effect of a change in current returns to child labour on the prevalence of work on household farms and school attendance. Second, Subsection 4.2 shows how a child's cognitive ability in Wave 3 - a plausible proxy of future returns to schooling - can determine their child labour supply and schooling in Wave 4. Finally, Subsection 4.3 tests our novel hypothesis of parents perceiving future returns to child labour on family farms via a vignette survey design; it also investigates whether, and how, these beliefs shape children's time allocation.

4.1 Current Returns to Child Labour: Agricultural Shocks

One important determinant of child labour in rural households is the presence of agricultural shocks. Specifically, households who are exposed to positive, transitory shocks to their harvest, and thus experience increases to their agricultural productivity and income, face higher short-run returns to farming. There is therefore an advantage to temporarily substitute children away from schooling and into farming as a result of higher returns. In fact, children are less likely to be attending school, more likely to have dropped out altogether, and spend less time in class and on homework in the presence of positive harvest shocks (Lyu, 2022).

The Luxury Axiom in the models of Basu (1999) and Basu and Van (1998) states that parents prefer not to put their children to work but children do work when household consumption falls below a subsistence threshold. Yet, our results are a contradiction. Parents put their children to work on family farms when income is *high*. In our context, the substitution effect dominates the income effect. This is consistent with Shah and Steinberg (2017) and Bau et al. (2021) who also find that child labour supply increases when household income rises.

4.1.1 Empirical Strategy

The empirical strategy in this section follows the 2SLS approach used in Lyu (2022). Specifically, we want to estimate the following equation

$$\text{Child Labour}_{iht} = \alpha_v + \alpha_t + \beta \log \mathcal{H}_{ht} + X'_{iht}\delta + \epsilon_{iht} \tag{2}$$

where the dependent variable of interest is the prevalence of child labour for child i of household h in wave t; α_v is the village fixed effect to control for time-invariant community characteristics, whereas α_t control for wave-specific aggregate shocks; X_{iht} is a matrix of time-varying, child-specific characteristics including age fixed effects, gender, and birth order; and ϵ_{iht} is the error term.

The independent variable of interest, $\mathcal{H}_{ht}$, is the value of self-reported harvest output in Ghanaian cedis at the household level for each wave. However, there are endogeneity concerns. For example, household-level harvest values can be correlated with time-varying household characteristics and farming ability. To assuage these concerns, we follow Lyu (2022) and instrument harvest value, $\mathcal{H}_{ht}$, with deviations in rainfall, temperature, and soil moisture, and the same variables interacted with plot-level land characteristics. Hence, harvest value is driven exclusively by the exogenous variation induced by environmental conditions as well as the interaction between environmental conditions and land characteristics.[14] [15]

4.1.2 Results

Our first results use data from Waves 2 and 3 of GSPS (2010) to show that when faced with positive harvest shocks, children aged 5 to 18 are more likely to be put on farms for work and to substitute away from work at their family business or trade. This suggests that when returns to farming are higher (due to positive shocks), children increase their labour supply on family farms.

Table 3 presents the results. Column (1) shows that when household harvest values increase by one hundred percent, the likelihood of their child being a family farm worker increases by 4.5 percentage points, or 13% of the mean. This is significant at the 5% level. Conversely, a one hundred percent increase in harvest value leads to a 3.2 percentage point decline, or 3.7% of the mean, in the likelihood of the child having attended school in the last year (column 2). This is significant at the 5% level.

Our second set of results in Table 4 describes the effect of positive harvest shocks on the intensive margin of child labour using Wave 1 of GSPS (2010).[16] Column (1) gives the total number of hours worked for the whole farming season in 2010 while columns (2) to (5) give the individual agricultural stages. Note that these estimates are at the household level, i.e., the farm hours reported is the total for all children within a household.

We find that when harvest values increase by one hundred percent, children work 28.1

[14] For more details, please see Appendix C and Section 3 of Lyu (2022).

[15] All the results described in this section rely on data from Waves 1, 2, and 3 of the GSPS (2010). This is because plot-level data for Wave 4 were not available at the time of the analysis.

[16] We use Wave 1 because it is the only wave that reports total hours worked by children within the family on household farms.

hours more across the entire farming season (column 1). This coefficient size is considerable. It means that when harvest values double, as a result of positive shocks, children work almost 50% of the mean more. This is significant at the 1% level. Column (2) shows that during harvest shocks, children spend 4 additional hours in land preparation, or 40% of the mean. This is a smaller effect compared to other agricultural stages because land preparation generally includes clearing and ploughing the land which requires physical strength such that children are at a disadvantage. In land management, which includes child-friendly tasks such as weeding, children report working almost 10 more hours, or 66% of the mean (column 3). Column (4) shows that when harvest values increase by one hundred percent, children spend 7.9 more hours, or 60% of the mean, harvesting and this is significant at the 5% level. Finally, children spend 3.4 additional hours, or 56% of the mean, in post-harvest activities such as plucking and cleaning harvested crops. This is significant at 1%.

Persistent Effects on Employment Outcomes. We test whether exposure to positive harvest shocks in childhood affects the child's likelihood of remaining in agricultural employment in the future. Specifically, we take increases in log harvest value at the household level, instrumented by changes in environmental conditions and interacted with a rich set of land characteristics, in Wave 1 (2010), and regress this against employment outcomes in Wave 3 (2018).[17] Table 5 presents the results. We find that children and young adults aged 10 to 26 in 2018,[18] who experienced positive harvest shocks in 2010, are more likely to remain working in agriculture, are more likely to substitute away from non-agricultural jobs such as paid employment and small businesses or market trades, and are also less likely to remain as full-time students.

First, when Wave 1 harvest values increase by one hundred percent, a child is 0.1 and 19.1 percentage points more likely to become a farm owner and a contributing farm worker, respectively, eight years later (columns 4 and 5). These effects are sizeable—38% and 43% of

[17] Because of the small sample that we obtain when merging Wave 1 (or Wave 2) with Wave 4, we are unable to look at longer term outcomes.

[18] We chose age 10 as Table 1 in Section 2 shows that children on average report starting farm work for the first time at this age. 26 is the maximum age in 2018 because children can be at most 18 years old in 2010 inour sample.

the mean, respectively. The coefficient for Farm Worker is statistically significant at 1% while the coefficient for Farmer Owner is not significant—this could largely be due to the fact that in Wave 3 the individuals are still 8 to 26 years old and many have not yet inherited or purchased their own farm.

Second, children exposed to harvest shocks during childhood are more likely to substitute away from non-agricultural work eight years later. Specifically, a one hundred percent increase in harvest value leads to a decline of 2.0, 2.0, and 0.3 percentage points in the likelihood of becoming a paid employee, business owner, and business worker, respectively (columns 1 to 3). The coefficients for Paid Employed and Business Owner are significant at the 10% and 5% levels. Again, these are large effects—86% to 142% of the mean, respectively.

Third, experiencing positive harvest shocks makes it much less likely that the child stays in full-time education. In fact, the likelihood of maintaining the status of a full-time student decreases by 14.1 percentage points, or 22% of the mean, and is statistically significant at the 1% level (column 6).

Taken together, this shows that exogenous harvest shocks during childhood can have persistent effects on the child's future outcome. This channel can be driven by two mechanisms. First, positive harvest shocks lead to increased child labour on family farms (see Section 4.1) due to greater short-run returns to farming in the presence of a productivity shock. Children accumulate agricultural skills through learning-by-doing.[19] Thus, children who carried out more farm work have a comparative advantage in farming in the future. They are therefore more likely to choose agricultural jobs due to higher returns relative to non-agricultural work. This is the mechanism we focus on in this paper.

The second mechanism could be that a household who experiences an increase in harvest value would mechanically accumulate more assets and land which facilitates more farming in the future. The children therefore simply grow up in a more agricultural household where there is more farm work to be done. A third mechanism is that households may also perceive the rise

[19]Since we model human capital as multi-dimensional, one can consider an equivalent to this mechanism as a child facing higher returns to education if they attended more classes or read more books during childhood.

in their harvest income in Wave 1 as a personal success and therefore encourage their children to pursue farming which they believe is remunerative.[20]

4.2 Future Returns to Schooling: Child Characteristics

In this section, we present evidence on the role of future returns to schooling as proxied by higher cognitive skills. The reasoning behind this is that children with higher cognitive skills benefits more from going to school - for instance because they learn more and can obtain better grades - and this translates into higher income in the future.

4.2.1 Empirical Strategy

Our empirical strategy tests how a child's lagged characteristics in Wave 3 (2018) of the GSPS (2010) can determine the likelihood of farm work and school attendance in Wave 4 (2022). We use past characteristics to predict present child labour supply to assuage concerns of reverse causality. For example, one may be concerned with a child scoring poorly on tests to be a consequence of working on farms in place of attending school, instead of parents allocating children who are performing poorly at school to farm work. A caveat with our approach is that lagged characteristics may be influenced by past parental decisions on human capital investments and we do not have any exogenous shock to these characteristics.

With this caveat in mind, we regress parents- and child-reported farm and domestic work in 2022 on four test scores administered in 2018. These are the Digit Span test and Raven's Matrices to measure cognitive ability, and Maths and English tests to measure numeracy and literacy. Test scores are measured as age- and sex-specific z-scores.

In particular, we estimate the following regression model:

$$\text{Child Labour}_{i,2022} = \alpha + \beta \text{Test z-Score}_{i,2018} + X'_{i,2022}\delta + \epsilon_{i,2022} \tag{3}$$

[20] Note that increases in harvest value are driven by exogenous shocks, and not tangible forces such as ability or assets. This specific mechanism is therefore only supported by parental beliefs.

where the dependent variable of interest is the prevalence of child labour in domestic and farm work for child i of household h in 2022; the independent variable is the lagged z-score for tests measuring cognitive ability, Maths and English in 2018; $X_{i,2022}$ is a matrix of child-specific characteristics including age fixed effects, gender, birth order, relationship to the household head, religion, and ethnicity; α is the constant. Standard errors $\epsilon_{i,2022}$ are clustered at the household level.

4.2.2 Results

We find that lagged cognitive ability negatively predicts future child labour supply (Table 6). A one standard deviation increase in the Digit Span test score leads to a 4.2 to 5.8 percentage point decrease in the likelihood of working on family farms four years later (columns 1 and 2). On the other hand, a one standard deviation increase in the Digit Span test score leads to a 2.7 percentage point increase in the likelihood of attending school four years later (column 3) and to 0.49 more years of schooling (column 4). Except the coefficient on school attendance, the other coefficients are statistically significant at the 10% or higher level. Similarly, a one standard deviation increase in the Raven's Matrices test score leads to a 1.2 to 4.8 percentage point decrease in the likelihood of working on family farms four years later (columns 1 and 2). On the other hand, a one standard deviation increase in the Digit Span test score leads to a 1.7 percentage point increase in the likelihood of attending school four years later (column 3) and to 0.77 more years of schooling (column 4).The coefficients on farm work in the last year and educational attainment are statistically significant at the 1% level. We find a similar relationship with the Maths and English test scores (see Appendix Table B.5). Higher lagged scores are associated with less future farm work and higher educational attainment.

Overall, these results show that children's cognitive characteristics in Wave 3 - a plausible proxy for future returns to schooling - can determine their child labour supply and schooling in Wave 4.

Finally, we also test whether lagged physical size predicts future child labour. This relationship is explained in Bau et al. (2021) in which they find height (alongside test scores) as an important determinant for child wages and child labour. We use an anthropometric index (an average of age- and sex-specific z-scores for height, weight, hip size, waist size, and arm circumference) to measure a child's physical size. Contrary to the literature, we find no relationship between a child's physical size and their likelihood of participating in farm and domestic work four years later. See Appendix Table B.6 for the results. This suggests that in the context of rural Ghana, parents view child labour on family farms as an avenue for developing crucial farming skills and experience. Indeed, parents select children to work on farms who they believe will become farmers in the future and who also have the lowest returns to schooling. Parents are therefore concerned with accumulating the appropriate human capital for their children—and not simply selecting their strongest or largest child to work to maximise farm profits.

4.3 Future Returns to Child Labour: Farm Work as Human Capital

The previous subsections have provided evidence on the role of current returns to child labour and future returns to schooling for shaping human capital investments. This subsection delves into the third component of the trade-off between child labour and schooling: future returns to farm work.

We tests the following hypotheses[21]: (1) Parents believe there are positive returns to child labour for farming; and (2) Parental investment allocations for their children are consistent with these beliefs. We test hypothesis (1) using a novel vignette survey design. We test hypothesis (2) by combining the beliefs elicited in the vignette with data on child labour and schooling.

4.3.1 Do parents believe there are positive returns to child labour for farming?

Vignette Design We design a vignette to elicit parents' beliefs on the human capital returns to child labour. We provide the household head with three hypothetical scenarios regarding a hypothetical fifteen year old child (we randomise the gender across households: Ama or Kofi).[22]

[21] Results in this section rely on Wave 4, because this is the wave in which we elicited parental beliefs on returns to child labour and on children future occupational paths.

[22] We choose the names, Ama and Kofi, because they are common and not specific to a wealth distinction.

We add that this child comes from a family with similar wealth to their own household and lives in a similar village. We want the respondent to report beliefs as truthfully and accurately to their own circumstances as possible. We use a hypothetical child, rather than their own, to circumvent social desirability bias. Appendix Figure A.5 details the full vignette.

We propose three hypothetical scenarios in order to measure the household head's *perceived* returns to child labour for farming as well as perceived returns to schooling for farming. A priori, it is not obvious whether parents value farming experience or schooling more in becoming a successful farm owner. Working on farms provides hands-on learning of specific tools and techniques. But, attending school can help a child attain financial literacy and social skills which may be important in managing a farm. Both activities serve to accumulate human capital. This vignette is designed to precisely distinguish parental beliefs on these two different dimensions of human capital. The scenarios for a hypothetical child are:

A. Completes Junior Secondary School, then attends two more years of schooling, and runs errands in their spare time.

B. Completes Junior Secondary School, and runs errands in their spare time.

C. Completes Junior Secondary School, and helps parents on farms in their spare time.

For each scenario, we ask the household head to rank, from 1 to 10, what they would expect Ama or Kofi's harvest value would be when they are older and are farm owners. The higher the value, the more productive the household head believes the child will be as a farmer in the future. This therefore captures parents' perceived returns to child labour and schooling in future agricultural productivity.

Vignette Results Table 7 shows that, on average, parents believe that a child who works on family farms (C) will, in the future, obtain a harvest whose value 1.5 times larger than if they had attained two more years of schooling with no farming experience (A). In fact, they believe the child with farming experience (C) will be over *twice* as productive as having no farming experience (B), holding educational attainment constant.[23]

Appendix Figure B.4 gives the mean expected harvest value for each of the three hypothetical scenarios by gender of the hypothetical child: Ama or Kofi. Expected harvest values are slightly higher (0.13 to 0.19) for a hypothetical boy (Kofi) than for a girl (Ama), and this difference is statistically significant for all three scenarios.[24] This is consistent with our first Descriptive Fact in Section 3 that boys are more often put to work on farms than girls. This productivity gap between Kofi and Ama suggests that parents view boys as having a comparative advantage in farming.

4.3.2 Do parents allocate children's time consistently with their beliefs?

The interplay of parents' perceived returns to farm work and actual child labour on farms only has bite if their child will actually become a farmer in later life. Put differently, it does not matter if parents resolutely believe that having hands-on farm experience is fundamental for future agricultural productivity if their child will never become a farmer and therefore never put these agricultural skills to use.

Empirical Strategy Intuitively, if parents behave consistently with their beliefs, children who are designated to become a farm owner should work more (less) on the farm and attend school less (more) if the parent believes in high future returns for farming in farm work (schooling). To test Hypothesis (2), we estimate the two following regressions.

$$\begin{aligned} \mathrm{y}_{ihv} = {} & \beta_A \text{Scenario A}_{hv} + \beta_B \text{Scenario B}_{hv} + \beta_C \text{Scenario C}_{hv} \\ & \pi \text{Farmer}_{ihv} + \theta_A \left(\text{Farmer}_{ihv} \times \text{Scenario A}_{hv} \right) \\ & + Z'_{ihv}\delta + \gamma_v + \epsilon_{ih}, \end{aligned} \tag{4}$$

[23]We are unable to test which of these beliefs are accurate, i.e., whether children who received more hands-on experience indeed become more productive farmers as adults. This is largely because the children we observe in Waves 1 and 2 are still fairly young 8 to 12 years later in Wave 4 and therefore do not have their own household for us to measure agricultural productivity. As such, we only obtain a very small sample when merging the children from Wave 1 (or Wave 2) to Wave 4.

[24]See Appendix Table B.7 for the full results.

and

$$
\begin{aligned}
\mathrm{y}_{ihv} = \beta_A \text{Scenario A}_{hv} + \beta_B \text{Scenario B}_{hv} + \beta_C \text{Scenario C}_{hv} \\
\pi \text{Farmer}_{ihv} + \theta_C \left(\text{Farmer}_{ihv} \times \text{Scenario C}_{hv}\right) \\
+ Z'_{ihv}\delta + \gamma_v + \epsilon_{ih}, \quad (5)
\end{aligned}
$$

where the outcome of interest, y_{ihv} for child i in household h of village v, is the following: $\mathbb{1}\{\text{Work on Farms}\}_{ihv}$, an indicator of whether the child has self-reported work on farms in the last seven days or the parents reported it; $\mathbb{1}\{\text{Attended School}\}_{ihv}$, an indicator of whether the child has attended school in the last year; the number of hours of class attended in the last week; the number of hours of missed classes in the last week; and years of schooling.

Hypothetical Scenarios A, B, and C are measured at the household level. They give the harvest value from 1 to 10 of what the household head would expect the hypothetical child to achieve for each of the three scenarios. Each scenario can therefore be interpreted as the effect on the outcome variable (e.g., the probability of farm work) for each incremental increase in the hypothetical harvest value for that scenario, holding constant the parents' perceived returns on the other two scenarios.

$\text{Farmer}_{ihv} \in \{0, 1\}$ is a binary variable equal to 1 if parents rank farm owner as the most likely occupation for their child.[25] This is interacted with either Scenario A or Scenario C to capture returns to schooling and child labour, respectively, in future farm work.

We include a matrix, Z_{hv}, which controls for gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects, γ_v. Standard errors are clustered at the household level. ϵ_{ih} is the error term.

[25]This is an indicator of ranking farm owner first, as opposed to business owner, teacher, and private or public sector employee. For more details on the questions on parents' occupational rankings see Sections 2.2 and 3.

Results Table 8 reports results from the estimation of equation 4. This focuses on how beliefs on returns to schooling for farming affect the time allocation for children who are designated to become a farmer. Column (3) shows that the effect of a one unit increase in expected harvest value for Scenario A[26] on school attendance in the last year is 3.3 percentage points higher for a child designated to become a farm owner compared to one who is not. Put simply, parents who strongly believe in the returns to *schooling* for farming would invest more in schooling for children who will become farmers than those who will not. This is significant at the 1% level. Similarly, columns (4) to (6) indicate that children who will become farmers have more years of schooling, more class hours, and fewer missed school hours. Since there is a trade-off between school and farm work (i.e., they are substitutes in a child's time), we observe parents, who strongly believe that schooling raises agricultural productivity, reduce farm work for children designated to be farm owners in contrast to those who are not (columns 1 and 2).

Table 9 follows the same structure but interacts Farmer with Hypothetical Scenario C.[27] The coefficient for the interaction term, Scenario C $\times$ Farmer, shows that children who have parents with high perceived returns to *child labour* in farming and have farm owner as their highest ranked occupation, face more time on farm work and substitute away from school. Specifically, a child who is designated to become a farm owner is 1.8 to 2.1 percentage points more likely to have worked on the farm than a child not designated to become a farmer (columns 1 and 2). Moreover, they spend fewer hours in class and also miss more hours of school (columns 5 and 6). They also have, on average, a quarter of a year shorter in educational attainment. These interacted coefficients are large relative to the magnitude of the uninteracted Scenario C coefficients. However, they are not statistically significant due to large standard errors. This suggests that there may be underlying heterogeneity in the effects. One possibility is that the effect depends on the child's age. Indeed, we observe striking differences when we split the sample between children age 5 to 11 (table B.8) and children age 12 to 18 (table

[26] Scenario A refers to a child who completes Junior Secondary School, then attends two more years of schooling, and runs errands in their spare time.

[27] Scenario C refers to a child who completes Junior Secondary School, and helps parents on farms in their spare time.

10). While there is no clear pattern for younger children, the interaction coefficients for older children is aligned with our hypothesis. A child who is designated to become a farm owner is 5.5 to 6.6 percentage points more likely to have worked on the farm than a child not designated to become a farmer (columns 1 and 2); albeit, we note that the coefficient on child-reported work is not statistical significant at conventional levels. Moreover, they spend fewer hours in class and also miss more hours of school (columns 5 and 6). While less pronounced, such heterogeneity between younger and older children is present also when looking at the interaction between Scenario A and occupational beliefs (Tables B.9 and B.10).

A caveat with our empirical strategy is that we do not have any exogenous variation in parental beliefs about returns to farm work and occupational paths. While we control for children and household observable characteristics, as well as for village characteristics, we cannot rule out that there other omitted variables which may bias our results. To strengthen the credibility of our results, we perform a placebo exercise. Specifically, we interact Scenario C with an indicator equal to 1 for children designated to get a white collar job (teacher or employee). The results are reported in Tables B.11 and B.12. We observe that the interaction coefficients do not follow the same pattern: beliefs on future returns for farming in farm work are only relevant for children designated to become a farm owner. Put simply, if parents believe that their child will be a white-collar worker in the future, they would not put into pratice their beliefs on farming - regardless of how strong these beliefs are.

Results from testing Hypotheses (1) and (2), through the design of a vignette, show that parental beliefs on the returns to farm work can be strong drivers for child labour. This is in contrast to previous literature as we show that parents' allocation for work on family farms can be motivated by their desire to accumulate their child's human capital in farming. The results are stronger for children aged 12 to 18. This is the age in which parents make crucial decisions on children's time allocation. Indeed, compulsory schooling ends after Junior Secondary School, and intense studying is required to pass the exams (i.e., the BECE) which is necessary to enrol in Senior Seconday School.[28]

5 Conclusion

This paper studies the relationship between child labour and schooling in rural Ghana. We first present some descriptive facts to motivate our main results. We find that there is a clear sorting of children into schooling versus farm work within a household. We highlight the role of child characteristics and birth order. Moreover, we document that parents are highly aspirational for their children, but they simultaneously diversify future occupations of their children by evaluating their individual ability.

In the second part of the paper, we present evidence on the determinants of child labour, studying separately the role of current returns to child labour, future returns to schooling, and future returns to child labour. Using an instrumental variable approach, we show that higher harvest value increases child labour supply and reduce school attendance. Moreover, children with higher future returns to schooling—as proxied by higher lagged cognitive ability—are more likely to attend school and less likely to work on the farm. Finally, we use a vignette design to confirm that parents do view child labour on family farms as one dimension of human capital formation. Specifically, they believe that having hands-on farming experience in childhood has a positive and direct effect on a child's agricultural productivity in later life. Parental beliefs are consistent with their behaviour as parents who have higher perceived returns to child labour are more likely to put their own children, who are designated to become a farmer, to work on farms.

These results provide a novel explanation of high child labour participation rates on family farms and businesses, particularly in sub-Saharan Africa. It offers a more benign view on parents' choice of intra-household allocations as evidence supports the hypothesis that parents are trying to accumulate human capital for their children and they are consciously considering their child's future occupation in their present decision-making.

[28]Moreover, we note that our vignette design focuses exactly on the choice of time allocation after JSS.

An important caveat is that this paper focuses on one dimension of human capital formation, namely, agricultural skills. It does not speak to the social returns of education and that schooling can have many benefits beyond academic achievements such as acquiring important social skills and positive peer effects. Moreover, high child labour participation rates can lead to inter-generational persistence in agriculture, making it harder to adapt to climate change and prevents migration from rural to urban regions. This can delay structural transformation and perpetuate poverty traps.

References

Adonteng-Kissi, Obed (2018). "Parental perceptions of child labour and human rights: A comparative study of rural and urban Ghana." In: *Child abuse & neglect* 84, pp. 34–44.

Adu-Gyamfi, Jones (2014). "Childhood construction and its implications for children's participation in Ghana." In: *Working Paper*.

Allen IV, James (2022). "Double-Booked: Effects of Overlap between School and Farming Calendars on Education and Child Labor." In: *Working Paper*.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena (2020). "Bride Price and Female Education." In: *Journal of Political Economy* 128.8, pp. 597–641.

Atkin, David (2016). "Endogenous skill acquisition and export manufacturing in Mexico." In: *American Economic Review* 106.8, pp. 2046–2085.

Attanasio, Orazio, Teodora Boneva, and Christopher Rauh (2022). "Parental beliefs about returns to different types of investments in school children." In: *Journal of Human Resources* 57.6, pp. 1789–1825.

Baland, Jean-Marie and James A. Robinson (2000). "Is Child Labor Inefficient?" In: *Journal of Political Economy* 108.4, pp. 663–679.

Banerjee, Abhijit (2004). "Educational Policy and the Economics of the Family." In: *Journal of Devel- opment Economics* 74.1, pp. 3–32.

Barker, Nathan (2022). "The Effects of Female Land Inheritance on Economic Productivity in Ghana." In: *University of Chicago, Chicago*.

Bass, Loretta Elizabeth (2004). *Child labor in sub-Saharan Africa*. Lynne Rienner Publishers.

Basu, Kaushik (1999). "Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards." In: *Journal of Economic Literature* 37.3, pp. 1083–1119.

Basu, Kaushik and Pham Hoang Van (1998). "The Economics of Child Labor." In: *American Economic Review* 88.3, pp. 412–427.

Bau, Natalie (2021). "Can Policy Change Culture? Government Pension Plans and Traditional Kinship Practices." In: *American Economic Review* 111.6, pp. 1880–1917.

Bau, Natalie, Martin Rotemberg, Manisha Shah, and Bryce Steinberg (2021). "Human Capital Investment in the Presence of Child Labor." In: *NBER Working Papers*. Working Paper 27241.

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti (2009). "Why Should We Care About Child Labor? The Education, Labor Market, and Health Consequences of Child Labor." In: *Journal of Human Resources* 44.4, pp. 871–889.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012). "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain." In: *Econometrica* 80.6, pp. 2369–2429.

Belloni, A. and V. Chernozhukov (2013). "Least squares after model selection in high-dimensional sparse models." In: *Bernoulli* 19, pp. 521–547.

Berman, Nicolas, Lorenzo Rotunno, and Roberta Ziparo (2020). "Sweet Child of Mine: Income, Health and Inequality." In: *CEPR Discussion Paper No. DP14444*.

Borusyak, Kirill and Peter Hull (2020). "Non-Random Exposure to Exogenous Shocks: Theory and Applications." In: *NBER Working Paper* w27845.

Canagarajah, Sudharshan and Helena Skyt Nielsen (2001). "Child Labor in Africa: A Comparative Study." In: *The Annals of the American Academy of Political and Social Science* 575, pp. 71–91.

Carneiro, Pedro, Imran Rasul, and Francesca Salvati (2023). *Families as Drivers of Inequality: Experimental Evidence from an Early Childhood Intervention*. Tech. rep. Working Paper.

Chiappori, Pierre-André and Costas Meghir (2015). "Intrahousehold inequality." In: *Handbook of income distribution*. Vol. 2. Elsevier, pp. 1369–1418.

Cunha, Flávio, Irma Elo, and Jennifer Culhane (2013). *Eliciting maternal expectations about the technology of cognitive skill formation*. Tech. rep. National Bureau of Economic Research.

Cunha, Flávio and James J. Heckman (2007). "The Technology of Skill Formation." In: *The American Economic Review* 97.2, pp. 31–47.

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman (2013). "School inputs, household substitution, and test scores." In: *American Economic Journal: Applied Economics* 5.2, pp. 29–57.

DHS (2018). *DHS Multiple Indicator Cluster Survey 6*. Unicef.

Dillon, Andrew, Elena Bardasi, Kathleen Beegle, and Pieter Serneels (2012). "Explaining Variation in Child Labor Statistics." In: *Journal of Development Economics* 98.1. Symposium on Measurement and Survey Design, pp. 136–147. ISSN: 0304-3878.

Doepke, Matthias and Fabrizio Zilibotti (2005). "The Macroeconomics of Child Labor Regulation." In: *American Economic Review* 95.5, pp. 1492–1524.

Duflo, Esther (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." In: *American Economic Review* 91.4, pp. 795–813.

Duflo, Esther, Pascaline Dupas, and Michael Kremer (2021). "The Impact of Free Secondary Education: Experimental Evidence from Ghana." In: *NBER Working Paper*.

Duflo, Esther and Rohini Pande (2007). "Dams." In: *The Quarterly Journal of Economics* 122.2, pp. 601–646.

Edmonds, Eric and Norbert Schady (2012). "Poverty Alleviation and Child Labor." In: *American Economic Journal: Economic Policy* 4.4, pp. 100–124.

Edmonds, Eric, Theodore Schultz, and John Strauss (2008). *Handbook of development economics*.

Edmonds, Eric and Caroline Theoharides (2020). "The short term impact of a productive asset transfer in families with child labor: Experimental evidence from the Philippines." In: *Journal of Development Economics* 146.102486.

Ejrnæs, M. and C. C. Pörtner (2004). "Birth order and the Intrahousehold Allocation of Time and Education." In: *Review of Economics and Statistics* 86.4, pp. 1008–1019.

Giannola, Michele (2024). "Parental investments and intra-household inequality in child human capital: evidence from a survey experiment." In: *The Economic Journal* 134.658, pp. 671–727.

GSPS, EGC-ISSER-Northwestern (2010). *Ghana Socioeconomic Panel Survey*. URL: https://egc.yale.edu/data/egc-isser-northwestern-ghana-panel-survey.

GSS (2003). *GLSS Child Labour Report*. Ghana Statistical Service.

— (2014). *GLSS6 Child Labour Report*. Ghana Statistical Service.

— (2017). *GLSS7 Labour Force Survey*. Ghana Statistical Service.

Guarcello, Lorenzo, Irina Kovrova, Scott Lyon, Marco Manacorda, and Furio C. Rosati (2010). "Towards Consistency in Child Labour Measurement: Assessing the Comparability of Estimates Generated by Different Survey Instruments." In: *Understanding Children's Work Programme Working Paper*.

ILO (2015). *World Report on Child Labour*. International Labour Organization.

— (2016). *Global Estimates of Child Labour: Results and trends, 2012-2016*. International Labour Organization.

— (2020). *Child Labour: Global Estimates 2020, Trends and The Road Forward*. International Labour Organization.

IPUMS (2010). *IPUMS Population and Housing Census*. Ghana Statistical Service.

Jayachandran, Seema and Rohini Pande (2017). "Why Are Indian Children So Short? The Role of Birth Order and Son Preference." In: *American Economic Review* 107.9, pp. 2600–2629.

Jensen, Robert (2010). “The (Perceived) Returns to Education and the Demand for Schooling.” In: *The Quarterly Journal of Economics* 125.2, pp. 515–548.

Krueger, Alan and Mikael Lindahl (2001). “Education for Growth: Why and For Whom?” In: *Journal of Economic Literature* 39.4, pp. 1101–1136.

Levy, Victor (1985). “Cropping Pattern, Mechanization, Child Labor, and Fertility Behavior in a Farming Economy: Rural Egypt.” In: *Economic Development and Cultural Change* 33.4, pp. 777–791.

Lichand, Guilherme and Sharon Wolf (2023). “Measuring Child Labor: the Who’s, the Where’s, the When’s, and the Why’s.” In: *Working Paper*.

Lise, Jeremy and Fabien Postel-Vinay (2020). “Multidimensional Skills, Sorting, and Human Capital Accumulation.” In: *American Economic Review* 110.8, pp. 2328–2376.

Lyu, Kaman (2022). “Agricultural Incomes and Human Capital Formation.” In: *Working Paper*.

Mbiti, Isaac, Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, and Rakesh Rajani (2019). “Inputs, incentives, and complementarities in education: Experimental evidence from Tanzania.” In: *The Quarterly Journal of Economics* 134.3, pp. 1627–1673.

Mokyr, Joel (2019). “The economics of apprenticeship.” In: *Apprenticeship in Early Modern Europe*, pp. 20–43.

Moscona, Jacob and Awa Ambra Seck (2021). *Social structure and redistribution: Evidence from age set organization*.

Nieuwenhuys, Olga (1996). “The paradox of child labor and anthropology.” In: *Annual review of anthropology* 25.1, pp. 237–251.

Ortiz-Ospina, Esteban and Max Roser (2016). “Child Labor.” In: *Our World in Data*. URL: `https://ourworldindata.org/child-labor`.

Rosenzweig, Mark R. and Kenneth I. Wolpin (1985). “Specific Experience, Household Structure, and Intergenerational Transfers: Farm Family Land and Labor Arrangements in Developing Countries.” In: *The Quarterly Journal of Economics* 100, pp. 961–987.

Schultz, Theodore (1960). “Capital formation by education.” In: *Journal of political economy* 68.6, pp. 571–583.

Shah, Manisha and Bryce Steinberg (2017). “Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital.” In: *Journal of Political Economy* 125.2, pp. 527–561.

Sviatschi, Maria Micaela (2022). “Making a narco: Childhood exposure to illegal labor markets and criminal life paths.” In: *Econometrica* 90.4, pp. 1835–1878.

Truffa, Francesca and Ashley Wong (2018). “Effect of Beliefs and Gender Roles on Girls’ Math Education.” In: *Working Paper*. Registered in the AEA RCT Registry, unique identifying number: AEARCTR-0003054 [Data collection completed].

UNESCO (2016). *Leaving no one behind: How far on the way to universal primary and secondary education?* UNESCO Institute for Statistics and Global Education Monitoring.

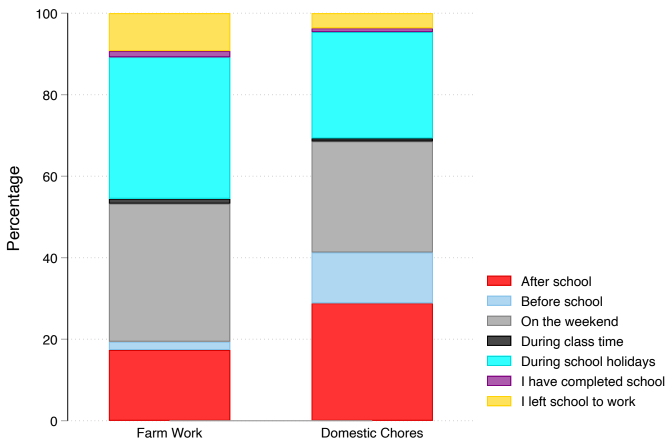
Figures and Tables

Table 1: Child Labour Supply

	Parents-reported		Child-reported			
	Farm		Farm		Domestic	
	Worked last year (1)	Hours worked (2)	Worked last 7 days (3)	Hours worked (4)	Worked last 7 days (5)	Hours worked (6)
A. All						
Mean	0.24	12.7	0.25	13.2	0.76	10.9
Standard Deviation	0.21	38.9	0.43	12.1	0.43	9.1
N	6636	1569	4762	799	4762	1946
B. Age 5 to 11						
Mean	0.15	11.4	0.16		0.66	
Standard Deviation	0.36	47.4	0.37		0.47	
N	3325	494	2506		2506	
C. Age 12 to 18						
Mean	0.32	13.4	0.35	13.2	0.87	10.9
Standard Deviation	0.47	34.3	0.48	12.1	0.33	9.1
N	3311	1075	2256	799	2256	1946

Notes: Variables, shown in the column headings, are categorised into whether it is reported by parents or self-reported by the child. Parents are asked about economic activity in the last year for all household members aged 5 and above—we simply restrict responses for ages 5 to 18. Children are asked to self-report work in the last 7 days only. Only children aged 12 and above are asked hours worked in the past week. This data is from Wave 4 of the GSPS (2010).

Figure 1: Time of Work

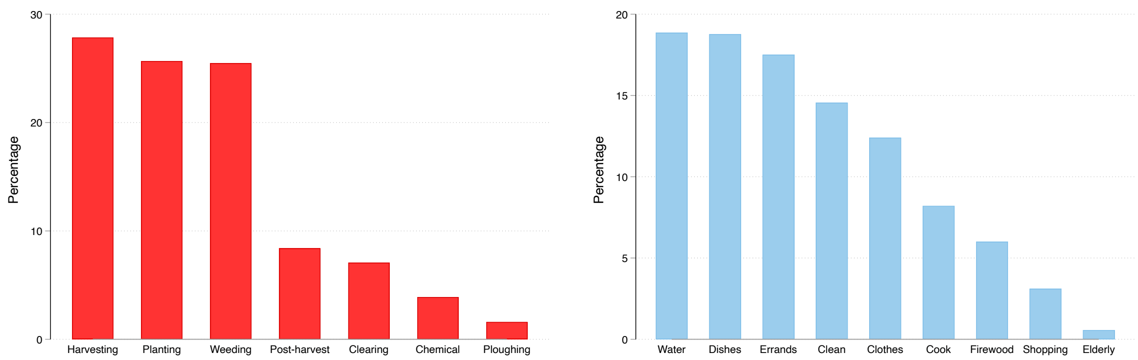


Notes: Time during which children engage in farm work (left bar) or perform domestic chores (right bar). Each child can select multiple options, and we count all responses as separate observations. We report percentages of respondents selecting an option. The sample includes all children aged 5 to 18 at the time of the Wave 4 survey in GSPS (2010).

Figure 2: Tasks

(a) Farming

(b) Domestic Work



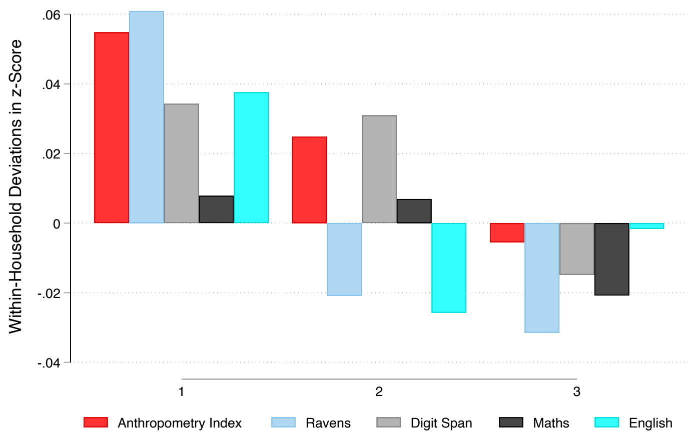
Notes: Activities performed by children engaging in farm work (left panel) or in domestic chores (right panel). The farming activities are: clearing and land preparation, ploughing, weeding, planting, chemical application, harvesting, and post-harvest processing. The domestic activities are: fetching water, washing dishes, running errands, cleaning dwelling, washing and mending clothes, cooking food, collecting firewood, shopping, and taking care of the elderly. Each child can select multiple options, and we count all responses as separate observations. We report percentages of respondents selecting an option. The sample includes all children aged 5 to 18 at the time of the Wave 4 survey in GSPS (2010).

Table 2: Sorting into School versus Child Labour

	Dependent Variable: Work $\in \{0,1\}$			
	Parents-reported	Child-reported		
	Farm Work	Farm Work	Domestic Work	Years of Schooling
	(1)	(2)	(3)	(4)
Age	0.036***	0.034***	0.032***	0.781***
	(0.003)	(0.005)	(0.005)	(0.014)
Male	0.041***	0.047***	-0.079***	-0.144**
	(0.009)	(0.013)	(0.015)	(0.070)
Firstborn	-0.014	-0.001	-0.017	0.089
	(0.013)	(0.017)	(0.019)	(0.085)
Years of schooling	-0.013***	-0.007	0.012**	
	(0.004)	(0.005)	(0.006)	
Ravens z-score				0.123**
				(0.050)
Digit span z-score				0.375***
				(0.076)
Maths z-score				0.263***
				(0.059)
English z-score				0.091
				(0.056)
Adjusted R^2	0.590	0.564	0.338	0.819
Household FE	Y	Y	Y	Y
Mean Dep. Var.	0.217	0.244	0.770	6.669
Observations	4897	3291	3291	1996

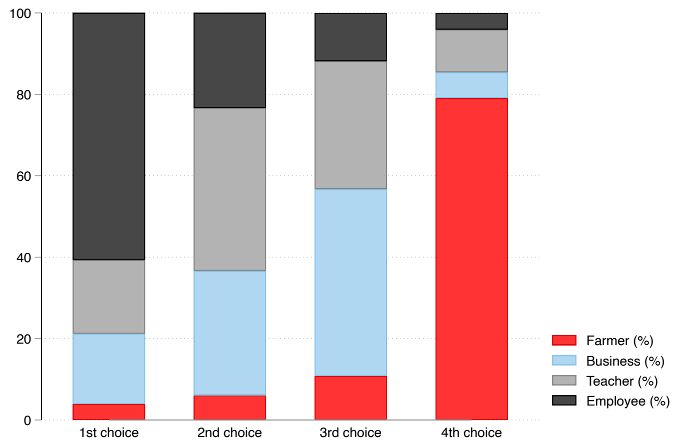
Notes: The dependent variables, shown in the column headings, are child labour participation which are binary variables coded 0 or 1, and are categorised into farm work and domestic work, and whether it is reported by parents or self-reported by the child, and years of schooling. The independent variables, Age, Male and Firstborn, are binary variables coded 0 or 1, whilst Ravens, Digit span, Maths, and English test scores are standardised z-scores by age and sex. We take household fixed effects. Standard errors in parentheses are clustered at the household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 3: Residualised z-Scores and Birth Order



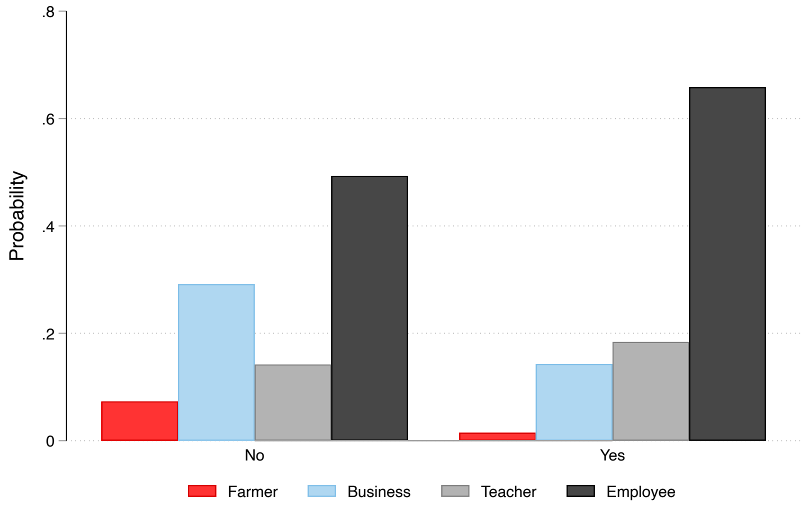
Notes: The x-axis denotes the birth order, i.e., 1 means the firstborn, 2 the second born and so forth. Estimates are residualised by household fixed effects and controlling for age and sex. They should therefore be interpreted as within-household deviations in the z-score. We restrict the sample to households with at least 3 children born at any time and report the results for the first three children.

Figure 4: Distribution of Parents' Occupational Choices for Children

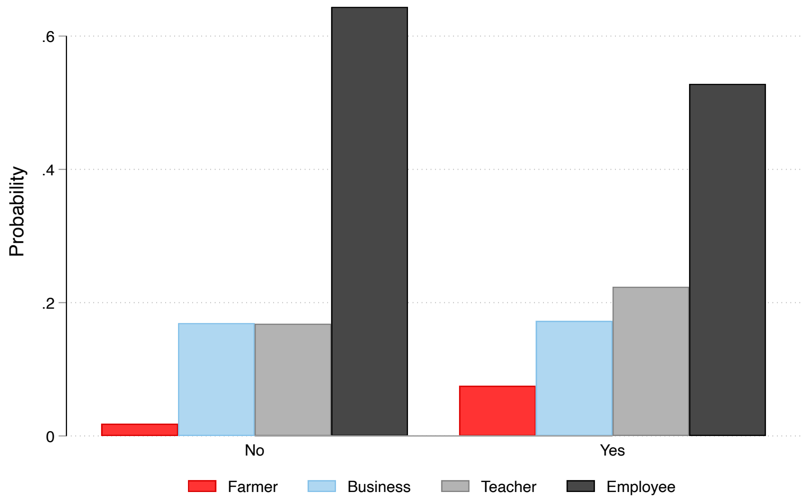


Notes: We report the percentage of parents' responses for each of the four occupations of each ranking. The four occupations are: Farm Owner, Business Owner or Market Trader, Teacher, and Private or Public Sector Employee.

Figure 5: Current time allocation and 1st Choice Occupations



(a) Attended school in the last 12 months?



(b) Worked on farm in the last 12 months?

Notes: The four occupations are: Farm Owner, Business Owner or Market Trader, Teacher, and Private or Public Sector Employee. The bars report the probability that parents rank the occupation as their first choice. The top-left panel is if the child attended school in the last 12 months for the child and the top-right panel is if the child attended school in the last 12 months for the child. The bottom-left panel is if the parent did not report farm work in the last 12 months for the child and the bottom-right panel is if the parent did report farm work in the last 12 months for the child

Table 3: Log Harvest on Child Farm Work

	Farm Worker	Attended school
	(1)	(2)
Log Harvest	0.045**	-0.032**
	(0.020)	(0.014)
Adjusted R^2	-0.068	-0.074
2SLS	Y	Y
Community and Wave FE	Y	Y
Controls	Y	Y
KP Wald F-Stat.	206.264	225.369
Mean Dep. Var.	0.342	0.875
Observations	4561	4937

Notes: The dependent variables, shown in the column headings, are binary variables coded 0 or 1. The independent variable, Log Harvest, is the logged harvest value in each wave at the household level. This is a 2SLS regression where Log Harvest is instrumented by changes in environmental conditions (i.e., deviations in rainfall, temperature and soil moisture) interacted with a rich set of land characteristics to predict plot-level harvest values. Lasso is applied to resolve the many-instruments problem. We report the Kleibergen-Paap rk Wald F-statistic for the first stage. The controls are age, gender, and firstborn, which are binary variables coded 0 or 1. We take village and wave fixed effects. Standard errors in parentheses are clustered at the household level. We use data from Waves 2 and 3 in GSPS (2010). The number of observations vary across columns because of missing values for the dependent variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Log Harvest on Farm Hours Worked by Children

	Whole Season	Land Preparation	Land Management	Harvest	Post-Harvest
	(1)	(2)	(3)	(4)	(5)
Log Harvest	28.114***	4.084*	9.776**	7.911**	3.464***
	(10.548)	(2.389)	(4.198)	(3.083)	(1.114)
Adjusted R^2	-0.189	-0.186	-0.201	-0.223	-0.195
2SLS	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y
Community FE	Y	Y	Y	Y	Y
KP Wald F-Stat.	129.878	129.878			
Mean Dep. Var.	58.774	10.287	14.781	13.223	6.102
Observations	1648	1648	1648	1648	1648

Notes: The dependent variables, shown in the column headings, are hours spent by all children on a farming stage at the household level. Column (1) is for the whole farming season while columns (2) to (5) are farming stages. The independent variable, Log Harvest, is the logged harvest value in each wave at the household level. This is a 2SLS regression where Log Harvest is instrumented by changes in environmental conditions (i.e., deviations in rainfall, temperature and soil moisture) interacted with a rich set of land characteristics to predict plot-level harvest values. Lasso is applied to resolve the many-instruments problem. We report the Kleibergen-Paap rk Wald F-statistic for the first stage. The controls are the number of children per household, religion, ethnicity, the main crop farmed, and plot size. We take community fixed effects. Standard errors in parentheses are clustered at the household level. We use data from Wave 1 in GSPS (2010). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Persistent Effects on Employment Outcomes

	Paid Employed (1)	Business Owner (2)	Business Contributor (3)	Farm Owner (4)	Farm Contributor (5)	Full-time Student (6)
Log Harvest W1	-0.020*	-0.020**	0.002	0.010	0.191***	-0.142***
	(0.012)	(0.008)	(0.016)	(0.023)	(0.063)	(0.049)
Adjusted R^2	-0.055	-0.068	-0.033	-0.040	-0.240	-0.226
2SLS	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
KP Wald F-Stat.	34.790	34.790	34.790	34.790	34.790	34.790
Mean Dep. Var.	0.023	0.014	0.047	0.026	0.440	0.648
Observations	1956	1956	1956	1956	1956	1956

Notes: The dependent variables, shown in the column headings, are binary indicators for different employments in Wave 3 (2018). The independent variable, W1 Log Harvest, is the logged harvest value in Wave 1 (2010) at the household level. This is a 2SLS regression where W1 Log Harvest is instrumented by changes in environmental conditions (i.e., deviations in rainfall, temperature and soil moisture) interacted with a rich set of land characteristics to predict plot-level harvest values in Wave 1. Lasso is applied to resolve the many-instruments problem. We report the Kleibergen-Paap rk Wald F-statistic for the first stage. The controls are age, gender, firstborn, relationship to the household head, religion, and ethnicity. Standard errors in parentheses are clustered at the household level. We use data from Waves 1 and 3 in GSPS (2010). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Cognitive Ability and Child Labour

	(1) Farm Work	(2) Farm Work (last 7d)	(3) Attended school last year	(4) Years schooling
Digit Span	-0.042*	-0.058**	0.027	0.495***
z-Score	(0.022)	(0.026)	(0.019)	(0.095)
Raven's Matrices	-0.048***	-0.012	0.017	0.177***
z-Score	(0.015)	(0.018)	(0.010)	(0.055)
adj. R^2	0.064	0.050	0.043	0.616
Controls	Y	Y	Y	Y
Mean Dep. Var.	0.545	0.483	0.828	7.030
N	1322	892	1209	1207

Notes: The dependent variables, shown in the column headings, are: farm work , a binary variables coded 1 if the child worked on the farm in in the last year- reported by either the child or the parent (column 1), farm work , a binary variables coded 1 if the child worked on the farm in in the last seven days- reported by the child (column 2), attended school , a binary variables coded 1 if the child worked on the farm in in the last year (column 3), the number of years of schooling completed by the child (column 4). These variables are measured in Wave 4 (2022). The independent variables are age- and sex-specific z-scores for a Digit Span and Raven's Matrices test from Wave 3 (2018). The controls are age, gender, firstborn, relationship to the household head, religion, and ethnicity. Standard errors in parentheses are clustered at the household level. We use data from Waves 3 and 4 in GSPS (2010). The different number of observations across columns is due to missing values in the dependent variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Summary Statistics

	Scenario A (1)	Scenario B (2)	Scenario C (3)
Mean	5.035	3.707	7.990
Standard Deviation	2.645	2.253	2.155
Lower Quartile	3	2	7
Median	5	3	8
Upper Quartile	7	5	10
N	4816	4816	4816

Notes: The scenarios are: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario has harvest value ranked from 1 to 10. The respondent is the household head of each household in Wave 4 of GSPS (2010).

Table 8: Parents' Perceived Returns to Education for Farming, occupational beliefs and time allocation

	(1) Worked farm	(2) Worked farm (7d)	(3) Attended school	(4) Years of schooling	(5) Class hrs	(6) Missed hrs
Scenario A	0.005	0.003	0.010***	-0.008	0.453***	-0.116
	(0.004)	(0.004)	(0.003)	(0.015)	(0.152)	(0.097)
Farmer	0.100	0.198	-0.335***	-2.100***	-8.513*	3.899
	(0.096)	(0.125)	(0.084)	(0.758)	(4.464)	(2.389)
Farmer ×	-0.001	-0.010	0.028*	0.281**	1.285*	-0.629*
Scenario A	(0.017)	(0.022)	(0.015)	(0.127)	(0.666)	(0.375)
Scenario B	0.004	0.007	-0.006*	0.007	0.150	0.003
	(0.004)	(0.005)	(0.003)	(0.017)	(0.179)	(0.102)
Scenario C	0.002	0.000	-0.001	-0.001	0.290	-0.027
	(0.004)	(0.005)	(0.003)	(0.018)	(0.203)	(0.121)
Adjusted R^2	0.379	0.351	0.254	0.802	0.422	0.154
Controls	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.294	0.247	0.880	5.744	27.536	2.343
Observations	5026	3595	5026	5026	3113	3102

Notes: The dependent variables are: Farm Work (reported by either child or parents), Farm Work (reported by child), coded 0 or 1, Attended School Last Year coded 0 or 1, Years of Schooling, Class Hours and Missed Hours per week. The independent variables are the hypothetical scenarios: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario is ranked from 1 to 10. Farmer is an indicator equal to 1 if parents indicated Farm Owner as most likely occupation for the child. The controls are: gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects. Standard errors in parentheses are clustered at the household level. Sample is all children age 5 to 18. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Parents' Perceived Returns to Child Labour for Farming, occupational beliefs and time allocation

	(1) Worked farm	(2) Worked farm (7d)	(3) Attended school	(4) Years of schooling	(5) Class hrs	(6) Missed hrs
Scenario A	0.005	0.003	0.011***	-0.000	0.477***	-0.127
	(0.004)	(0.004)	(0.003)	(0.015)	(0.152)	(0.096)
Scenario B	0.004	0.007	-0.006*	0.005	0.142	0.007
	(0.004)	(0.005)	(0.003)	(0.017)	(0.179)	(0.102)
Scenario C	0.001	-0.000	-0.001	0.004	0.306	-0.034
	(0.004)	(0.005)	(0.003)	(0.018)	(0.204)	(0.121)
Farmer	-0.082	-0.001	-0.347*	0.576	7.166	-2.559
	(0.213)	(0.253)	(0.208)	(1.050)	(8.962)	(4.738)
Farmer $\times$	0.021	0.018	0.017	-0.161	-1.030	0.357
Scenario C	(0.025)	(0.032)	(0.025)	(0.142)	(1.089)	(0.564)
Adjusted R^2	0.379	0.351	0.253	0.801	0.422	0.154
Controls	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.294	0.247	0.880	5.744	27.536	2.343
Observations	5026	3595	5026	5026	3113	3102

Notes: The dependent variables are: Farm Work (reported by either child or parents), Farm Work (reported by child), coded 0 or 1, Attended School Last Year coded 0 or 1, Years of Schooling, Class Hours and Missed Hours per week. The independent variables are the hypothetical scenarios: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario is ranked from 1 to 10. Farmer is an indicator equal to 1 if parents indicated Farm Owner as most likely occupation for the child. The controls are: gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects. Standard errors in parentheses are clustered at the household level. Sample is all children age 5 to 18. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Parents' Perceived Returns to Child Labour for Farming, occupational beliefs and time allocation: age 12 to 18

	(1) Worked farm	(2) Worked farm (7d)	(3) Attended school	(4) Years of schooling	(5) Class hrs	(6) Missed hrs
Scenario A	0.004	0.006	0.013***	0.008	0.439**	-0.084
	(0.005)	(0.006)	(0.004)	(0.023)	(0.220)	(0.109)
Scenario B	0.007	0.009	-0.011**	-0.013	0.276	-0.173
	(0.006)	(0.008)	(0.005)	(0.027)	(0.248)	(0.128)
Scenario C	0.004	0.002	-0.001	0.033	0.026	0.100
	(0.006)	(0.007)	(0.005)	(0.029)	(0.302)	(0.205)
Farmer	-0.322	-0.369	-0.140	1.289	18.891	-10.029**
	(0.233)	(0.367)	(0.304)	(1.552)	(12.229)	(4.462)
Farmer=1 ×	0.055*	0.066	-0.014	-0.326	-2.959**	1.254**
Scenario C	(0.030)	(0.046)	(0.037)	(0.211)	(1.424)	(0.523)
Adjusted R^2	0.417	0.369	0.227	0.526	0.365	0.047
Controls	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.397	0.355	0.832	8.262	28.121	1.918
Observations	2533	1667	2533	2533	1367	1362

Notes: The dependent variables are: Farm Work (reported by either child or parents), Farm Work (reported by child), coded 0 or 1, Attended School Last Year coded 0 or 1, Years of Schooling, Class Hours and Missed Hours per week. The independent variables are the hypothetical scenarios: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario is ranked from 1 to 10. Farmer is an indicator equal to 1 if parents indicated Farm Owner as most likely occupation for the child. The controls are: gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects. Standard errors in parentheses are clustered at the household level. Sample is all children age 12 to 18. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix

A Appendix: Details on data and measurement

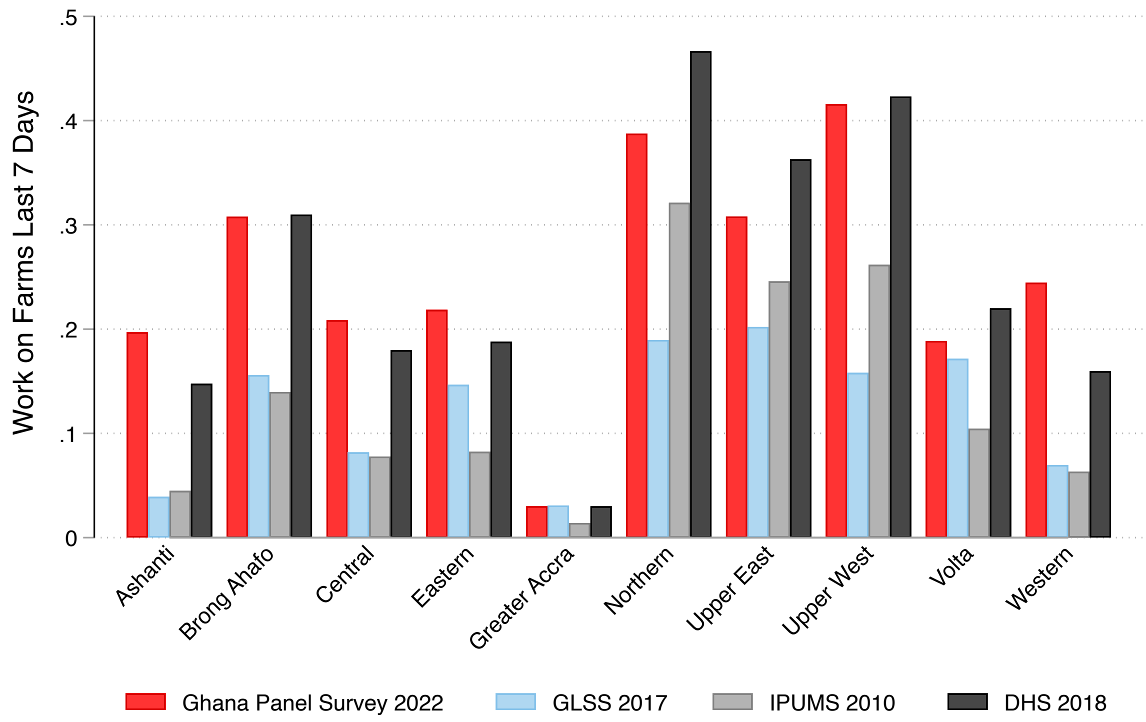
Comparison with Other Surveys There are substantial inconsistencies across child labour statistics in surveys administered by developing countries (Dillon et al., 2012; Guarcello et al., 2010). To get a measurement as precise as possible, in addition to relying on labour supply questions asked to the household head, we also measure child labour by asking the question directly to the child. We also phrase the question to be terse and inclusive of marginal and household-oriented tasks (Figure A.4).

We compare our child labour participation rates with three other surveys administered in Ghana: GLSS 7 Labour Force Survey (GSS, 2017), IPUMS Population and Housing Census (IPUMS, 2010), and DHS Multiple Indicator Cluster Survey 6 (DHS, 2018). Specifically, we compare the survey response to whether a child aged 5 to 18 has worked on a farm in the last seven days. The precise way the question is asked, respondent type, and survey length are different across surveys and can have an impact on truthful reporting. The exact wording of the survey questions are documented in Figure A.4. Seasonality may also differ among the surveys which can have an impact on actual participation rates given that agricultural work is highly seasonal.

Crucially, our survey is the only one in Ghana that specifically targets the child to self-report work on farms in contrast to other surveys which direct questions to adults or by proxy, who may misreport due to social desirability bias or misinformation. Indeed, Lichand and Wolf (2023) find that in cocoa-producing regions of Cote d'Ivoire, parents under-report child labour by a factor of at least 60%. On the other hand, children accurately self-report their labour as verified by third-party data and satellite imagery of cocoa production.

Figure A.1 shows the prevalence of child labour on farms in the last seven days across the ten regions of Ghana for all four surveys (including our own Ghana Socioeconomic Panel Survey (GSPS, 2010) from Wave 4). GSS (2017) and IPUMS (2010) document much lower

Figure A.1: Survey Responses across Regions in Ghana



Notes: The survey question "Any Work on Farms in the Last 7 Days" is a dummy variable coded 0 or 1. We do not account of seasonality across the four surveys. The sample is children aged 5 to 18.

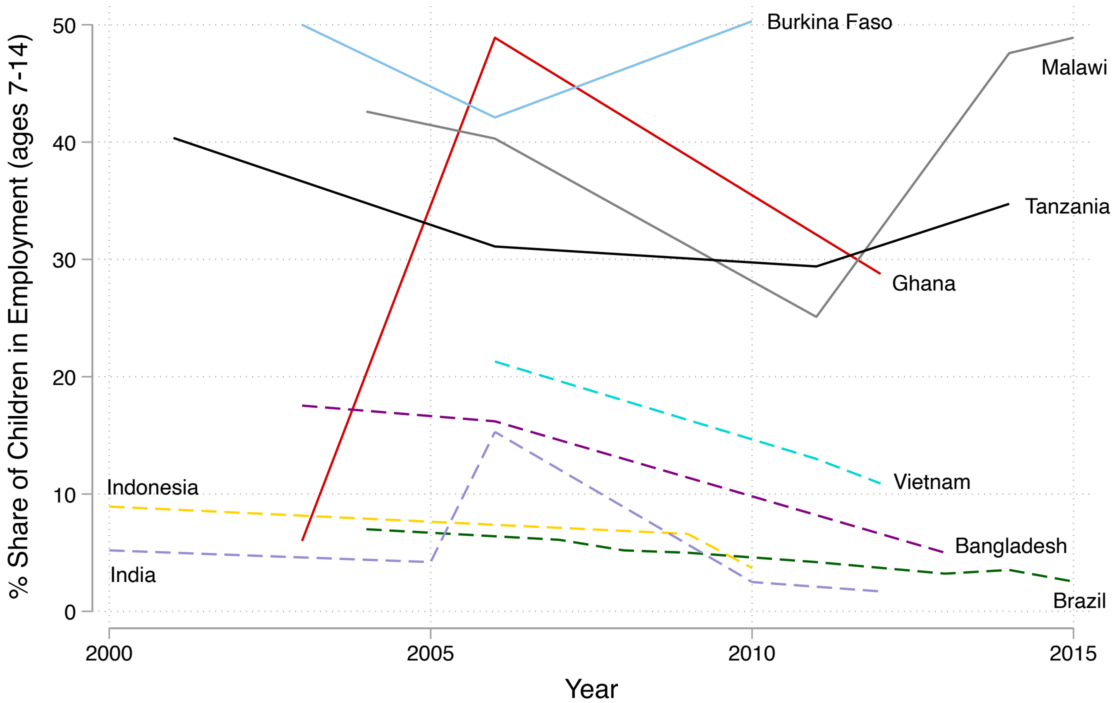
levels of farm work compared to our survey (11.1% and 13.6%)—twice as many children are reporting work in the last week in our survey (25.4%). This may in large part be due to the respondent being a parent and by proxy in GSS (2017) and IPUMS (2010), respectively.

The estimate, 24.8%, from DHS (2018) is similar, or slightly smaller, compared to our survey. DHS (2018) is a very long and detailed questionnaire targeted at the mother or primary caregiver of the child to accurately elicit the prevalence of child labour. The fact that the closest results to ours come from a survey which specifically aims to measure child labour prevalence gives us confidence on the quality of our measurement.[29] In fact, seeking out and

[29]A note of caution is that one assumes that estimates from DHS (2018) is accurate; therefore our estimates being close implies accuracy of our survey's measurement. Certainly, without an independent dataset for verification, it is not guaranteed that this measurement is the truth. However, the under-reporting of child labour is well established so one can expect the bias to be attenuating estimates (Lichand and Wolf, 2023; Dillon et al., 2012; Guarcello et al., 2010). Thus, a higher child labour participation rate can be seen as an improvement in truthful reporting.

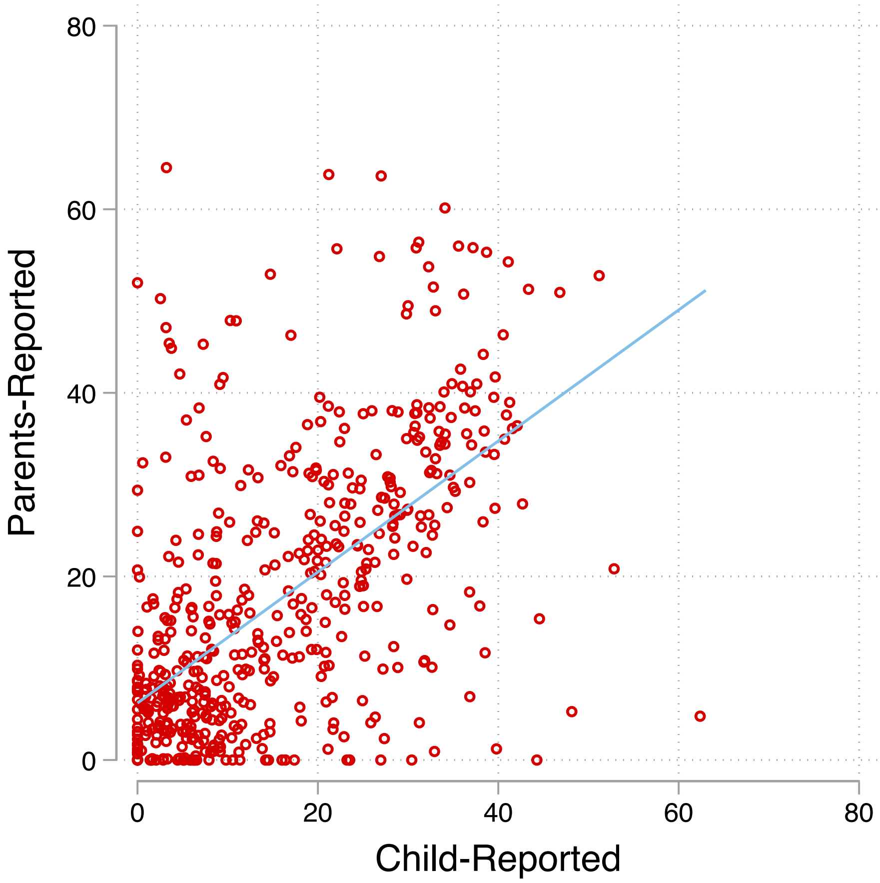
then interviewing the primary caregiver separately, as DHS (2018) does, is expensive and time-intensive. Moreover, the specific question asked is extensive as it includes many examples of farming-related activity to preclude marginal or non-paying work from being excluded (Appendix Figure A.4). We find posing a short question, inclusive of household work, directly to the child attains as high a labour participation rate as that of a long, detailed, and potentially costly, survey module.

Figure A.2: Comparing Child Labour across Countries



Notes: Statistics obtained from the World Bank.

Figure A.3: Comparing Reported Hours



Notes: The red circles measure the reported hours of farm work per week. The line indicates the line of best fit between the parents-reported and child-reported hours.

Figure A.4: Survey Questions and Respondents

- Ghana Panel Survey 2022: Child only
  - *Did you do any work, for pay or for free, or assist on a farm during the last 7 days (even if it was for only one hour)?*
- GLSS 2017: Parents if child absent
  - *Did [NAME] do any work for pay, profit, family gain of did [NAME] produce anything for barter or home use during the last 7 days even if it was for only one hour?*
- IPUMS 2010: Proxy
  - *During the 7 days before census night, did [NAME] engage in any activity for pay (cash or kind) or profit or family gain for at least one hour? (This includes helping in the family business/farm, trading, street vending, etc)*
- DHS 2018: Mother/Primary Caretaker
  - *Since last (day of the week), did [NAME] do any of the following activities, even for only one hour? Did [NAME] do any work or help on [his/her] own or the household's plot, farm, food garden or looked after animals? For example, growing farm produce, harvesting, or feeding, grazing or milking animals?*

Notes: Each survey question on child labour is extracted directly from the respective questionnaire.

Figure A.5: Vignette

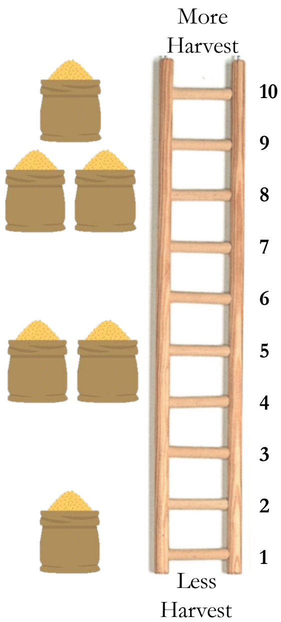
Ama/Kofi is 15 years old and lives in a village similar to your village. Ama/Kofi comes from a family with wealth similar to yours. When she/he finishes school, she/he will become a farmer. Please listen to the following hypothetical scenarios A, B, or C for Ama/Kofi.

A. Completes Junior Secondary School, then attends 2 more years of schooling, and runs errands in their spare time.

B. Completes Junior Secondary School, and runs errands in their spare time.

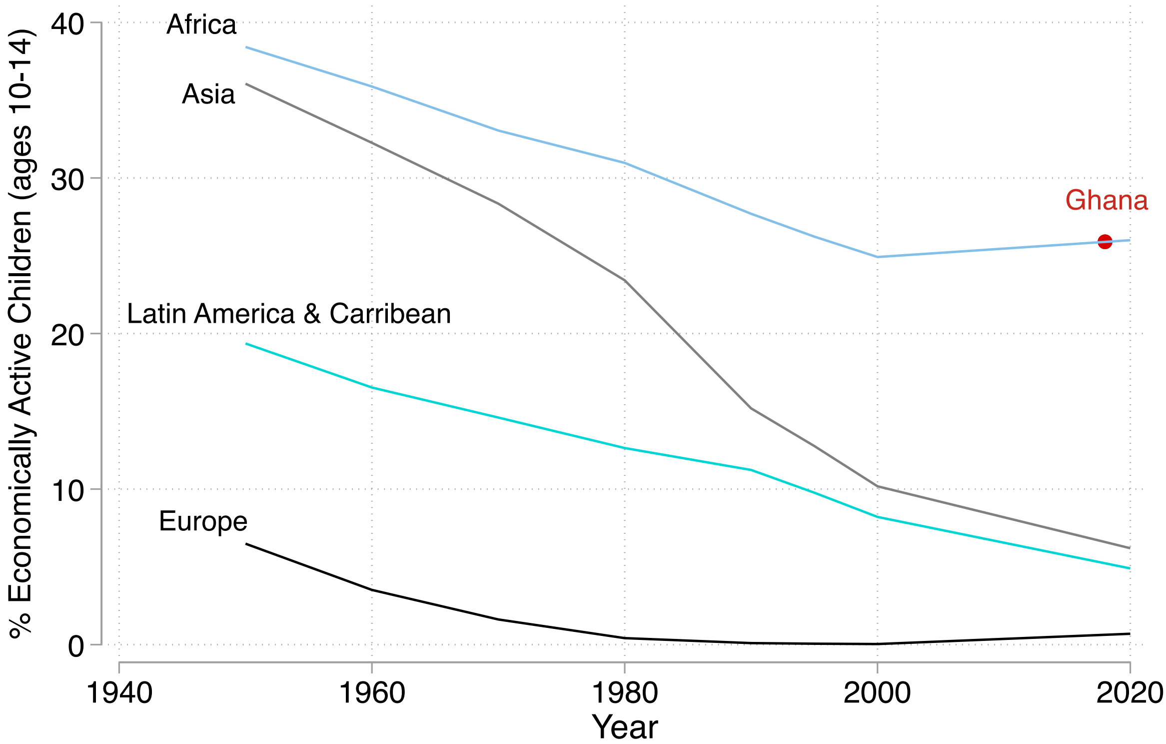
C. Completes Junior Secondary School, and helps parents on farms in their spare time.

When Ama/Kofi becomes a farm owner, what do you think the value of their harvest would be from an average farming season? Place scenarios A, B and C on the ladder from 1 to 10. Moving up the ladder means having a higher value of harvest.



Notes: The image of a ladder with increasing sacks of harvest is shown to the household head alongside reading out the hypothetical setting. The respondent is encouraged to physically point on the ladder where they believe the expected harvest value would be for each of the three hypothetical scenarios. We randomise the gender of the hypothetical child (Ama or Kofi) across households to test for heterogeneity. We choose Ama and Kofi because they are common and neutral names.

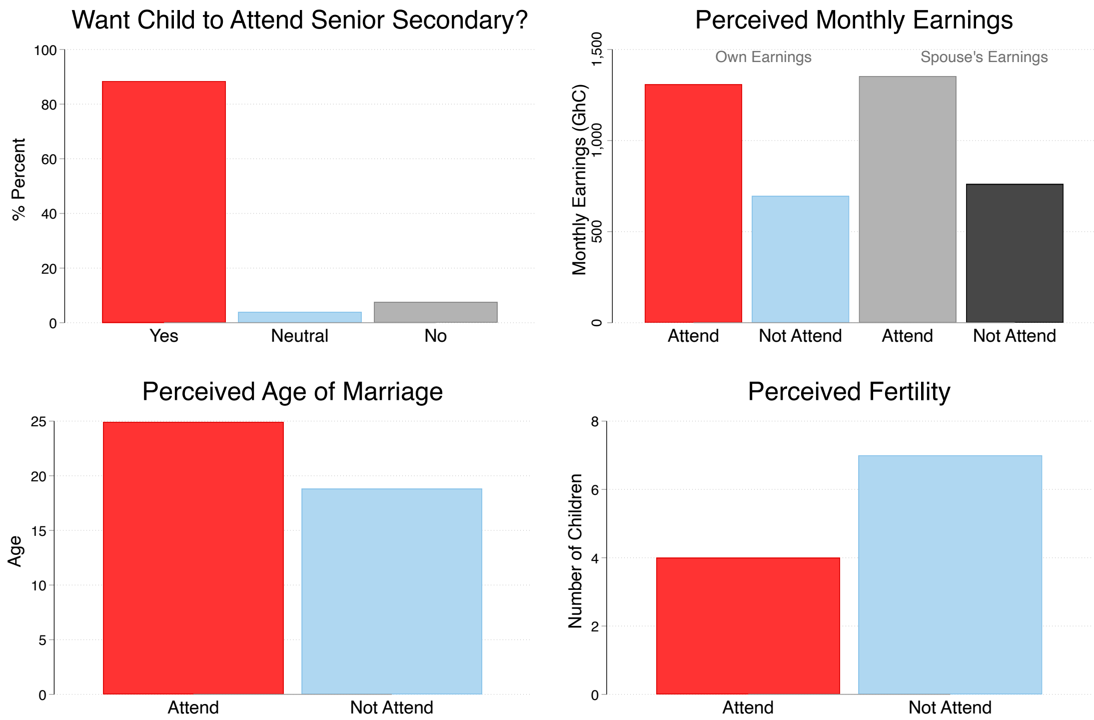
Figure B.1: Historical Trends in Child Labour by Continent



Notes: Statistics obtained from the International Labour Organization (ILO).

B Appendix: Additional results

Figure B.2: Parents' Perceived Returns to Attending Senior Secondary School



Notes: Responses obtained from Wave 3 of GSPS (2010) from a survey module designed by Truffa and Wong (2018).

Table B.1: Additional Child Labour Questions

	Child-reported		
	Paid for work $\in \{0,1\}$	Pay last 7 days (GhC)	Age started (years)
	(1)	(2)	(3)
Mean	0.03	58	10
Standard Deviation	0.17	70	2.32
N	799	23	1127

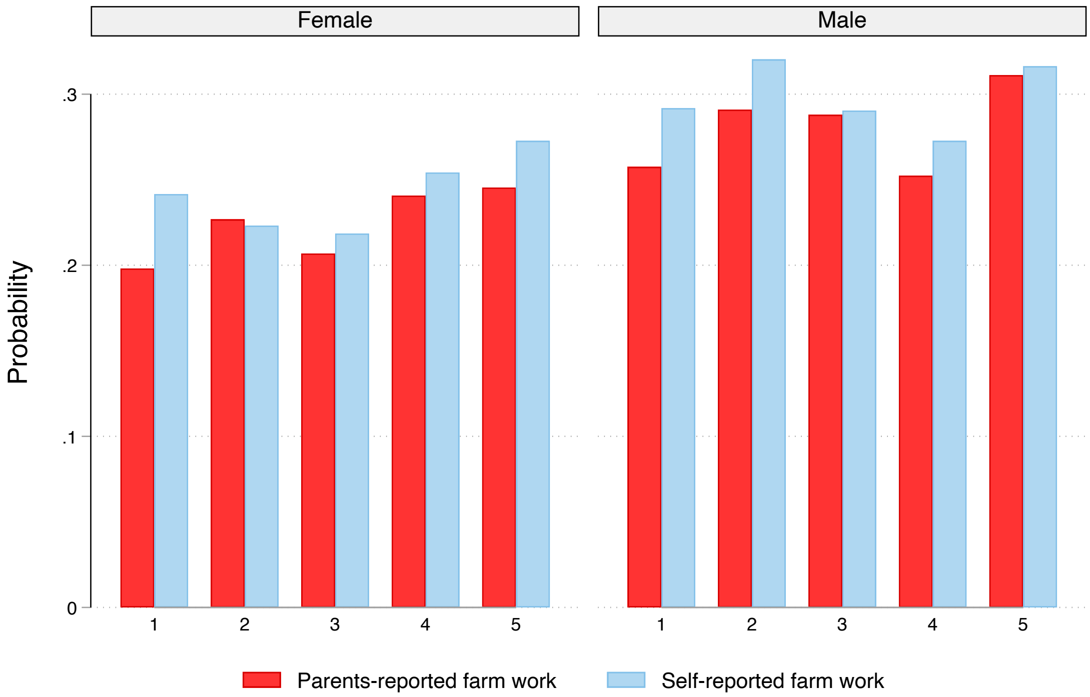
Notes: Variables, shown in the column headings, are farm work self-reported by the child. Only children aged 12 and above are asked whether they were paid for farm work, their pay in Ghanaian cedis in the last 7 days (if paid), and age they first started work. This data is from Wave 4 of the GSPS (2010).

Table B.2: Child Labour and Demographic Characteristics

	Parents	Self	
	Farm Work	Farm Work	Domestic Work
	(1)	(2)	(3)
Age	0.030***	0.032***	0.039***
	(0.002)	(0.002)	(0.002)
Male	0.045***	0.052***	-0.103***
	(0.010)	(0.013)	(0.015)
Firstborn	-0.029**	-0.009	-0.018
	(0.012)	(0.017)	(0.018)
Adjusted R^2	0.580	0.545	0.348
Household FE	Y	Y	Y
Mean Dep. Var.	0.251	0.261	0.759
Observations	5572	3683	3683

Notes: The dependent variables, shown in the column headings, are binary variables coded 0 or 1, and are categorised into farm work and domestic work, and whether it is reported by parents or self-reported by the child. The independent variables, Age, Male and Firstborn, are binary variables coded 0 or 1. We take household fixed effects. Standard errors in parentheses are clustered at the household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.3: Child Labour and Birth Order



Notes: Variables are categorised into farm work and domestic work, and whether it is reported by parents or self-reported by the child. These are summary statistics taken across the sample without fixed effects or controls.

Table B.3: Advantages of the Firstborn

	Works on Farm (1)	Years of Schooling (2)	Anthropometry z-Score (3)	Ravens z-Score (4)	Digit Span z-Score (5)	Maths z-Score (6)	English z-Score (7)
Male	0.042***	-0.150***	-0.199***	0.077**	0.026	-0.020	-0.017
	(0.009)	(0.051)	(0.024)	(0.032)	(0.026)	(0.047)	(0.042)
Firstborn	-0.023*	0.198***	0.114***	0.118***	0.031	0.059	0.102*
	(0.013)	(0.068)	(0.035)	(0.045)	(0.034)	(0.061)	(0.057)
Adjusted R^2	0.584	0.841	0.410	0.369	0.465	0.383	0.499
Household FE	Y	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.251	5.566	-0.032	-0.022	-0.027	-0.031	-0.043
Observations	5572	4897	3926	3683	3683	2023	2023

Notes: The dependent variables, shown in the column headings, are child labour participation on family farms coded 0 or 1, years of schooling, and z-scores for Anthropometry Index, Ravens, Digit span, Maths, and English test scores standardised by age and sex. We take household fixed effects. Standard errors in parentheses are clustered at the household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4: Determinants of Parents' Occupational Ranking

	Farm Owner (1)	Business Owner (2)	Teacher (3)	Employee (4)
Ravens z-score	-0.005	-0.014	-0.001	0.019
	(0.006)	(0.015)	(0.015)	(0.018)
Digit span z-score	-0.018**	0.001	-0.014	0.031
	(0.008)	(0.019)	(0.019)	(0.024)
Maths z-score	0.005	-0.034**	-0.006	0.036*
	(0.007)	(0.016)	(0.016)	(0.019)
English z-score	-0.011*	-0.046***	0.004	0.053***
	(0.006)	(0.017)	(0.016)	(0.020)
Male	0.011	-0.025	0.029	-0.015
	(0.008)	(0.022)	(0.022)	(0.026)
Firstborn	0.001	-0.028	-0.047*	0.074**
	(0.011)	(0.027)	(0.028)	(0.036)
Adjusted R^2	0.169	0.203	0.207	0.247
Household FE	Y	Y	Y	Y
Mean Dep. Var.	0.020	0.183	0.186	0.610
Observations	1884	1884	1884	1884

Notes: The dependent variables, shown in the column headings, are the probability of parents selecting this occupation as the most likely occupation for their child in the future. The independent variables, Age, Male and Firstborn, are binary variables coded 0 or 1, whilst Ravens, Digit span, Maths, and English test scores are standardised z-scores by age and sex. We take household fixed effects. Standard errors in parentheses are clustered at the household level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.5: Test Scores and Child Labour

	(1) Farm Work	(2) Farm Work (last 7d)	(3) Attended school last year	(4) Years schooling
English	-0.064***	-0.089***	0.025	0.429***
z-Score	(0.024)	(0.029)	(0.019)	(0.092)
Maths	-0.007	0.042	0.025	0.398***
z-Score	(0.023)	(0.027)	(0.019)	(0.092)
adj. R^2	0.046	0.059	0.012	0.395
Controls	Y	Y	Y	Y
Mean Dep. Var.	0.551	0.505	0.776	8.506
N	757	467	741	739

Notes: The dependent variables, shown in the column headings, are: farm work , a binary variables coded 1 if the child worked on the farm in in the last year- reported by either the child or the parent (column 1), farm work , a binary variables coded 1 if the child worked on the farm in in the last seven days- reported by the child (column 2), attended school , a binary variables coded 1 if the child worked on the farm in in the last year (column 3), the number of years of schooling completed by the child (column 4). These variables are measured in Wave 4 (2022). The independent variables are age- and sex-specific z-scores for a Maths and English test from Wave 3 (2018). The controls are age, gender, firstborn, relationship to the household head, religion, and ethnicity. Standard errors in parentheses are clustered at the household level. We use data from Waves 3 and 4 in GSPS (2010). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: Physical Size and Child Labour

	(1) Farm Work Farm Work	(2) Farm Work (last 7d)	(3) Attended school last year	(4) Year schooling
Anthropometric	-0.005	-0.023	0.011	0.473^{***}
Index	(0.016)	(0.021)	(0.014)	(0.064)
adj. R^2	0.125	0.122	0.046	0.738
Controls	Y	Y	Y	Y
Mean Dep. Var.	0.471	0.408	0.851	5.913
N	1799	1223	1574	1572

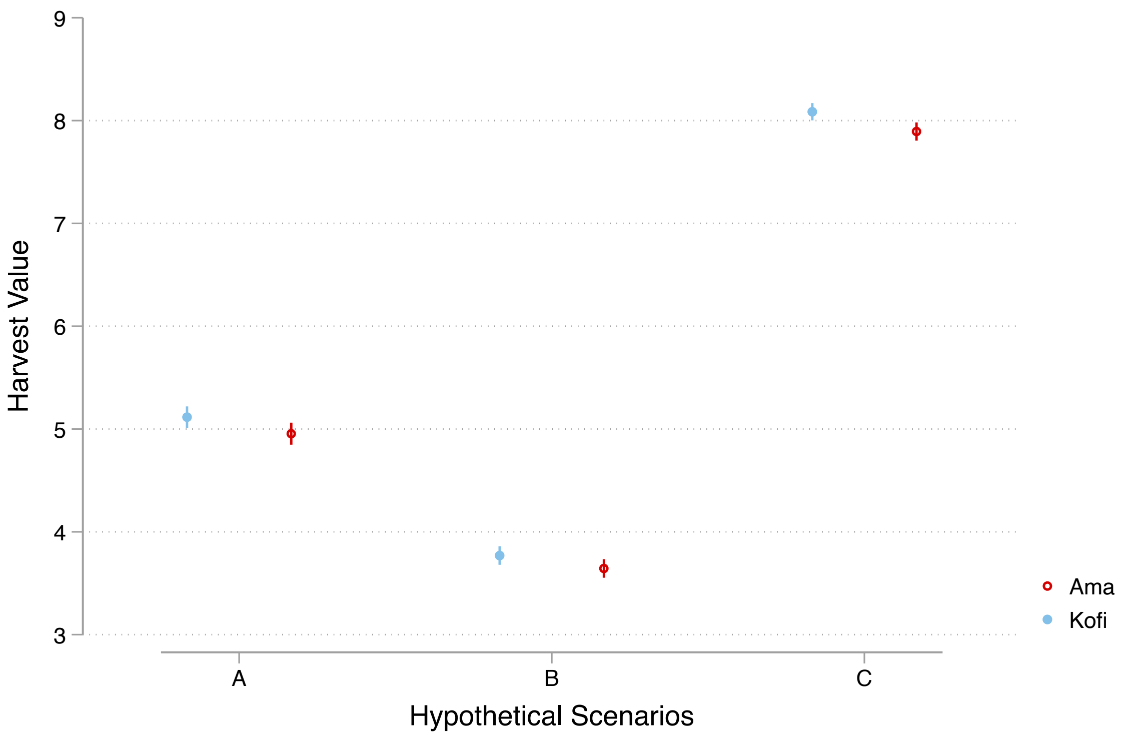
Notes: The dependent variables, shown in the column headings, are: farm work , a binary variables coded 1 if the child worked on the farm in in the last year- reported by either the child or the parent (column 1), farm work , a binary variables coded 1 if the child worked on the farm in in the last seven days- reported by the child (column 2), attended school , a binary variables coded 1 if the child worked on the farm in in the last year (column 3), the number of years of schooling completed by the child (column 4). These variables are measured in Wave 4 (2022). The independent variable, Anthropometric Index, is an index of age- and sex-specific z-scores for height, weight, hip size, waist size, and arm circumference from Wave 3 (2018). The controls are age fixed effects, gender, firstborn, relationship to the household head, religion, and ethnicity. Standard errors in parentheses are clustered at the household level. We use data from Waves 3 and 4 in GSPS (2010). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Parents' Perceived Returns to Child Labour and Education for Farming

	Dependent Variable: Harvest Value $\in [1, 10]$					
	Scenario A		Scenario B		Scenario C	
	(1)	(2)	(3)	(4)	(5)	(6)
Ama	4.955^{***}		3.644^{***}		7.894^{***}	
	(0.055)	(0.046)	(0.045)			
Kofi		5.115^{***}		3.769^{***}		8.086^{***}
		(0.053)		(0.046)		(0.043)
H_0: Ama = Kofi (p-value)	0.035		0.053		0.002	
Adjusted R^2	0.378	0.406	0.352	0.379	0.453	0.479
N	4816	4816	4816	4816	4816	4816

Notes: This table corresponds to Figure B.4. The dependent variables, shown in the column headings, are Hypothetical Scenarios A, B, and C. The independent variable of interest is an indicator for the sex of the hypothetical child: Ama is female and Kofi is male. The scenarios are: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario has harvest value ranked from 1 to 10. Robust standard errors are used. We report p-values for the t-test of Ama = Kofi. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.4: Parents' Perceived Returns to Child Labour in Farming



Notes: Coefficient estimates from the regression, Hypothetical Scenario$_h$ $= \mathbb{1}\{\text{Kofi}\} + \epsilon$, are plotted to show the average harvest value by the sex of the hypothetical child: Ama is female and Kofi is male. The scenarios are: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario has harvest value ranked from 1 to 10. Robust standard errors are used. Vertical bars indicate the 95% confidence intervals. See Appendix Table B.7 for the full results.

Table B.8: Parents' Perceived Returns to Child Labour for Farming, occupational beliefs and time allocation: age 5 to 11

	(1) Worked farm	(2) Worked farm (7d)	(3) Attended school	(4) Years of schooling	(5) Class hrs	(6) Missed hrs
Scenario A	0.007	-0.001	0.007**	0.004	0.545***	-0.168
	(0.005)	(0.005)	(0.003)	(0.016)	(0.186)	(0.165)
Scenario B	0.003	0.008	0.000	0.015	0.098	0.150
	(0.006)	(0.006)	(0.004)	(0.019)	(0.238)	(0.181)
Scenario C	-0.005	-0.005	-0.000	-0.039**	0.542**	-0.170
	(0.006)	(0.006)	(0.004)	(0.018)	(0.267)	(0.181)
Farmer	0.264	0.141	-0.672*	0.154	-10.836	13.571
	(0.460)	(0.395)	(0.357)	(1.156)	(16.788)	(8.795)
Farmer ×	-0.021	-0.001	0.066	-0.009	1.296	-1.325
Scenario C	(0.052)	(0.048)	(0.041)	(0.148)	(2.058)	(1.070)
Adjusted R^2	0.288	0.288	0.305	0.632	0.449	0.129
Controls	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.192	0.157	0.932	2.918	26.983	2.770
Observations	2229	1696	2229	2229	1522	1511

Notes: The dependent variables are: Farm Work (reported by either child or parents), Farm Work (reported by child), coded 0 or 1, Attended School Last Year coded 0 or 1, Years of Schooling, Class Hours and Missed Hours per week. The independent variables are the hypothetical scenarios: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario is ranked from 1 to 10. Farmer is an indicator equal to 1 if parents indicated Farm Owner as most likely occupation for the child. The controls are: gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects. Standard errors in parentheses are clustered at the household level. Sample is all children age 5 to 11. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.9: Parents' Perceived Returns to Education for Farming, occupational beliefs and time allocation: age 5 to 11

	(1) Worked farm	(2) Worked farm (7d)	(3) Attended school	(4) Years of schooling	(5) Class hrs	(6) Missed hrs
Scenario A	0.006	-0.000	0.008**	0.002	0.533***	-0.164
	(0.005)	(0.005)	(0.003)	(0.016)	(0.186)	(0.171)
Farmer	0.047	0.196	-0.014	-0.253	-5.031	5.329
	(0.161)	(0.188)	(0.140)	(0.712)	(5.654)	(4.028)
Farmer×	0.007	-0.012	-0.019	0.065	1.104	-0.664
Scenario A	(0.028)	(0.034)	(0.028)	(0.147)	(0.985)	(0.669)
Scenario B	0.003	0.008	-0.000	0.015	0.106	0.145
	(0.006)	(0.006)	(0.004)	(0.019)	(0.238)	(0.182)
Scenario C	-0.005	-0.005	0.001	-0.040**	0.540**	-0.173
	(0.006)	(0.006)	(0.004)	(0.018)	(0.265)	(0.180)
Adjusted R^2	0.288	0.288	0.302	0.632	0.449	0.129
Controls	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.192	0.157	0.932	2.918	26.983	2.770
Observations	2229	1696	2229	2229	1522	1511

Notes: The dependent variables are: Farm Work (reported by either child or parents), Farm Work (reported by child), coded 0 or 1, Attended School Last Year coded 0 or 1, Years of Schooling, Class Hours and Missed Hours per week. The independent variables are the hypothetical scenarios: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario is ranked from 1 to 10. Farmer is an indicator equal to 1 if parents indicated Farm Owner as most likely occupation for the child. The controls are: gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects. Standard errors in parentheses are clustered at the household level. Sample is all children age 5 to 11. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.10: Parents' Perceived Returns to Education for Farming, occupational beliefs and time allocation: age 12 to 18

	(1) Worked farm	(2) Worked farm (7d)	(3) Attended school	(4) Years of schooling	(5) Class hrs	(6) Missed hrs
Scenario A	0.005	0.005	0.012***	-0.002	0.393*	-0.074
	(0.005)	(0.006)	(0.004)	(0.023)	(0.221)	(0.111)
Farmer	0.115	0.160	-0.442***	-2.791***	-17.587**	1.112
	(0.134)	(0.176)	(0.111)	(1.070)	(8.612)	(2.677)
Farmer=1 ×	0.003	0.005	0.042**	0.314*	2.162*	-0.218
Scenario A	(0.024)	(0.032)	(0.021)	(0.174)	(1.242)	(0.577)
Scenario B	0.007	0.009	-0.011**	-0.009	0.305	-0.177
	(0.006)	(0.008)	(0.005)	(0.027)	(0.249)	(0.128)
Scenario C	0.005	0.003	-0.001	0.024	-0.018	0.117
	(0.006)	(0.007)	(0.005)	(0.029)	(0.298)	(0.206)
Adjusted R^2	0.416	0.368	0.229	0.527	0.365	0.045
Controls	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.397	0.355	0.832	8.262	28.121	1.918
Observations	2533	1667	2533	2533	1367	1362

Notes: The dependent variables are: Farm Work (reported by either child or parents), Farm Work (reported by child), coded 0 or 1, Attended School Last Year coded 0 or 1, Years of Schooling, Class Hours and Missed Hours per week. The independent variables are the hypothetical scenarios: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario is ranked from 1 to 10. Farmer is an indicator equal to 1 if parents indicated Farm Owner as most likely occupation for the child. The controls are: gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects. Standard errors in parentheses are clustered at the household level. Sample is all children age 12 to 18. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.11: Placebo: Parents' Perceived Returns to Child Labour for Farming, white collar jobs and time allocation

	(1) Worked farm	(2) Worked farm (7d)	(3) Attended school	(4) Years of schooling	(5) Class hrs	(6) Missed hrs
Scenario A	0.006	0.004	0.010***	0.031	0.494***	-0.127
	(0.004)	(0.004)	(0.003)	(0.026)	(0.153)	(0.096)
Scenario B	0.002	0.006	-0.004	-0.033	0.113	-0.001
	(0.004)	(0.005)	(0.003)	(0.033)	(0.178)	(0.103)
Scenario C	0.004	-0.007	-0.003	0.079	-0.077	0.175
	(0.008)	(0.009)	(0.008)	(0.064)	(0.332)	(0.233)
White Collar	-0.010	-0.106	0.108*	0.593	-3.409	2.221
	(0.065)	(0.071)	(0.065)	(0.588)	(2.627)	(1.990)
White Collar ×	-0.003	0.011	0.003	-0.105	0.448	-0.234
Scenario C	(0.008)	(0.008)	(0.008)	(0.071)	(0.314)	(0.236)
Adjusted R^2	0.314	0.289	0.227	0.131	0.421	0.153
Controls	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.294	0.247	0.880	5.744	27.536	2.343
Observations	5026	3595	5026	5026	3113	3102

Notes: The dependent variables are: Farm Work (reported by either child or parents), Farm Work (reported by child), coded 0 or 1, Attended School Last Year coded 0 or 1, Years of Schooling, Class Hours and Missed Hours per week. The independent variables are the hypothetical scenarios: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario is ranked from 1 to 10. Farmer is an indicator equal to 1 if parents indicated Farm Owner as most likely occupation for the child. The controls are: gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects. Standard errors in parentheses are clustered at the household level. Sample is all children age 5 to 18. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.12: Placebo: Parents' Perceived Returns to Child Labour for Farming, white collar jobs and time allocation: age 12 to 18

	(1) Worked farm	(2) Worked farm (7d)	(3) Attended school	(4) Years of schooling	(5) Class hrs	(6) Missed hrs
Scenario A	0.005	0.005	0.013***	0.027	0.439**	-0.086
	(0.005)	(0.006)	(0.004)	(0.029)	(0.221)	(0.109)
Scenario B	0.006	0.009	-0.009*	-0.039	0.245	-0.167
	(0.006)	(0.008)	(0.005)	(0.035)	(0.254)	(0.129)
Scenario C	0.012	-0.001	-0.007	0.117	-0.254	0.128
	(0.010)	(0.013)	(0.011)	(0.072)	(0.523)	(0.362)
White Collar	0.046	-0.031	0.121	1.239**	-1.843	0.530
	(0.085)	(0.106)	(0.096)	(0.606)	(4.810)	(3.138)
White Collar ×	-0.009	0.005	0.008	-0.093	0.314	-0.020
Scenario C	(0.010)	(0.013)	(0.012)	(0.075)	(0.569)	(0.366)
Adjusted R^2	0.415	0.359	0.218	0.150	0.362	0.047
Controls	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y
Mean Dep. Var.	0.397	0.355	0.832	8.262	28.121	1.918
Observations	2533	1667	2533	2533	1367	1362

Notes: The dependent variables are: Farm Work (reported by either child or parents), Farm Work (reported by child), coded 0 or 1, Attended School Last Year coded 0 or 1, Years of Schooling, Class Hours and Missed Hours per week. The independent variables are the hypothetical scenarios: (A) Completes Junior Secondary School (JSS), 2 more years of schooling, and runs errands; (B) Completes JSS, and runs errands; and (C) Completes JSS, and helps parents on farms. Each scenario is ranked from 1 to 10. Farmer is an indicator equal to 1 if parents indicated Farm Owner as most likely occupation for the child. The controls are: gender, whether the child is the firstborn, age fixed effects, family size, whether the household head is female, whether there is a father or mother present in the household, and dummy variables for the father's occupation and education level. We take village fixed effects. Standard errors in parentheses are clustered at the household level. Sample is all children age 12 to 18. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Appendix: Instrumenting harvest value

C.1 Empirical Strategy

We use self-reported harvest output at the household level. This level of variation can capture more granular income volatility which is usually missed in more aggregated regional-level income data. The drawback is that transitory fluctuations in agricultural incomes cannot necessarily be taken to be random. The endogeneity issue is explained with the following equations.

The baseline regression for child i in household h at wave $t = 1, 2, 3$ is

$$\text{y}_{iht} = \alpha + \beta \log \mathcal{H}_{ht} + \gamma_h + X'_{iht}\delta + \epsilon_{iht} \tag{6}$$

where y_{iht} is a vector of child outcomes: anthropometry, cognitive ability, test scores, schooling, educational expenditures, and mental health. $\mathcal{H}_{ht}$ is the value of harvest output in Ghanaian cedis. γ_h is the household fixed effect to control for time-invariant household characteristics. X_{iht} is a matrix of time-varying, child-specific characteristics. α is the constant and ϵ_{iht} is the error term.

The endogeneity issue arises with harvest value, $\mathcal{H}_{ht}$, which comprises of an exogenous component, ζ_{ht} (e.g., level of rainfall), and an endogenous component, h_{ht}, which reflects time-varying household characteristics and farming ability.[30] These two components act as a scalar to the farming production function, $f(\cdot)$, which depends on capital, k, and land characteristics, l. Thus, the value of harvest can be written as

$$\mathcal{H}_{ht} = (\zeta_{ht} + h_{ht}) f(k, l). \tag{7}$$

h_{ht} is possibly correlated with the error term, ϵ_{iht}, in Equation (6). For example, if a child is sick during the farming season, and thus works less, their parents may dedicate more time to care for the child and neglect their crops and thereby reduce their harvest income.

[30]For brevity, these components enter linearly and additively separable. These assumptions can be lifted without changes to the result.

Consequently, this causes one to overstate β, the effect of harvest value on child labour. On the other hand, we have $\text{Cov}(\zeta_{ht}, \epsilon_{iht}) = 0$. The exogenous component, ζ_{ht}, therefore acts as a productivity shock.

To address these endogeneity concerns, we adopt a two-stage least squares (2SLS) approach by instrumenting harvest value, $\mathcal{H}_{ht}$, with deviations in rainfall, temperature, and soil moisture, and the same variables interacted with plot-level land characteristics. Hence, harvest value is driven exclusively by the exogenous variation induced by environmental conditions, ζ_{ht}, as well as the interaction between environmental conditions and land characteristics, $\zeta_{ht} * l_{ph}$.

Exploiting the panel structure of the data, we take household fixed effects to account for bias arising from time-invariant household characteristics. The identifying assumption is that changes in environmental conditions (i.e., deviations in rainfall, temperature, and soil moisture) that occur during plot-specific crop seasons can only affect child outcomes through the variation in harvest value, conditional on household fixed effects and child-specific, time-varying controls.

A potential violation of the exclusion restriction could be that parents manipulate the season start and end times in response to child outcomes. For example, if their child is about to undertake their BECE (junior secondary school diploma equivalent), they may delay planting to redirect capital for their child's academic resources (e.g., textbooks), and thus endogenising the season's length. This would understate the effect of harvest value on schooling outcomes. We overcome this potential violation by fixing the crop seasons across households.

C.2 Harvest Value Prediction

The first stage, at the plot level, is

$$\hat{\mathcal{H}}_{pht} = \hat{\Pi}(\zeta_{pht} + \zeta_{pht} * l_{ph}) \tag{8}$$

where all variables are demeaned to account for household fixed effects and $\hat{\Pi}$ is a matrix of estimators.

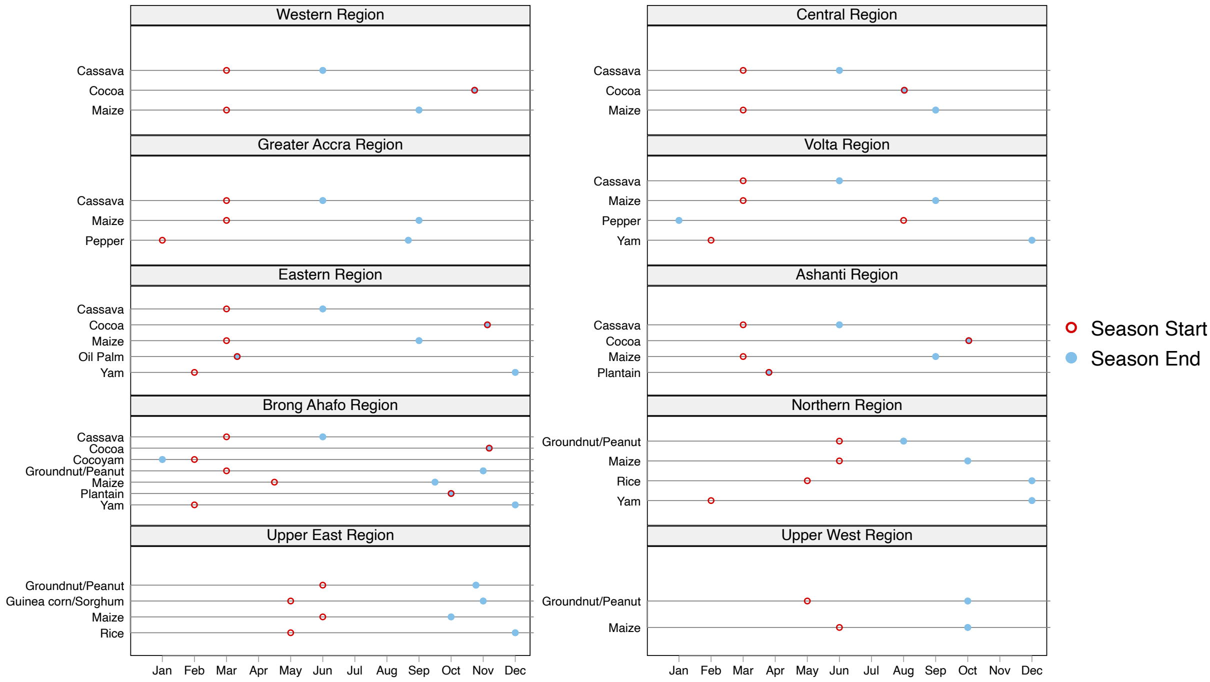
Environmental conditions, ζ_{pht}, are specific to each plot, p, due to households having different crops (i.e., different farming seasons) over multiple plots. ζ_{pht} is a vector measuring shocks in rainfall, temperature and soil moisture. More specifically, it includes their deviations from the long-run mean, alongside its quadratic to capture nonlinearities, and binary variables for extreme values below the 20th and above the 80th percentile. The seasons are further partitioned into thirds to capture the planting, growing and harvest seasons. This allows for a more flexible functional form as different phases of the farming season could respond differently to environmental conditions.

Secondly, l_{ph} is a vector of plot-level land characteristics: soil colour (e.g., red, brown, black), soil type (e.g., loamy, sandy, dry), soil depth, soil wetness, soil drainage, plot size, and plot distance from home. It is interacted with environmental conditions as particular characteristics can make weather shocks more or less salient for harvest incomes. For instance, if the plot naturally drains well then it is less prone to flooding during heavy rainfall. Naturally, these endogenous land characteristics do not individually enter the first stage without the interaction with the exogenous environmental conditions (see Borusyak and Hull (2020) for more details).

All these variables and their interactions amount to over a hundred instruments. There is therefore concern for overfitting. In order to select instruments in a data-driven way, we follow Belloni, Chen, et al. (2012) and apply Lasso to obtain optimal instruments to form first-stage predictions. This setting of "high p, low n" data makes Lasso a suitable and efficient solution to alleviate the many-instruments problem (Belloni, Chen, et al., 2012). Specifically, we use Lasso to obtain a data-dependent set of instruments. Then, we refit the first stage via ordinary least squares (OLS) to alleviate Lasso's shrinkage bias (Belloni, Chen, et al., 2012; Belloni and Chernozhukov, 2013).

We use split sampling to choose the optimal method in selecting the penalty parameter, λ^*. The plugin estimator performed the best in minimising mean-squared error (MSE) in the testing sample. Thus, Lasso selects 18 final instruments which are used to estimate the first

Figure C.1: FAO Crop Seasons across Regions



Notes: This figure shows the season lengths for the most common crops grown across the ten regions of Ghana. Note that cassava takes 15 months to grow from March to June the following year. Cocoa, maize and plantains are fixed at one year.

stage.[31]

C.3 Plot- to Household-level Aggregation

The richness of the farm enterprise data enables analysis to be conducted at the *plot* level. This is important as this level of granularity allows for the range of crops grown and plot-specific characteristics within households to be accounted for. In rural Ghana, the average household owns 1.69 cultivated plots whilst the upper quartile owns at least two. To reduce risk, many households diversify their crop choice over plots. Hence, the different crop seasons on each plot need to be considered when instrumenting with the correct period of environmental conditions. Figure C.1 shows the variety of crops grown, and their respective seasons, across the ten regions of Ghana.

[31]The first stage has an F-statistic of 36.4.

To preclude endogenous start and end times of crop seasons, we take the crop-specific calendar for each wave as given by the Food and Agriculture Organisation of the United Nations (FAO).[32] For less nationally representative crops, such as cocoyam and okro, we take the farmer's self-reported season.

Regressing harvest value on plot-level instruments improves precision. However, the regression of interest is at the household level—how does transitory household income, driven by exogenous environmental conditions, affect child labour? Hence, we require an additional step to aggregate harvest values from the plot to the household level. Since household harvest is the direct sum of plot-level harvests, we can simply sum up predicted plot-level harvest values to arrive at the aggregate predicted harvest value at the household level.[33]

For clarity, the empirical strategy is outlined by the following four steps. This method closely follows the strategy employed in Duflo and Pande (2007).[34]

1. Regress plot-level harvest values on crop-season-specific environmental conditions including their interactions with plot-level land characteristics. Instruments selected by Lasso.

2. Sum over predicted plot-level harvest values to form predicted household-level harvest.

3. Use predicted household-level harvest as an instrument for actual household-level harvest. This is the first stage.[35]

4. Apply 2SLS on child outcomes. This is the second stage.

Section C.2 outlines the prediction exercise in Step 1. This section details Steps 2 to 3. First, we compute the predicted harvest value for the household level. We aggregate across

[32]There is typically a calendar for the North and South since the Northern regions have only one major season whilst the South also has a minor season.

[33]Plot characteristics that are categorical and cannot be summed up (e.g., soil colour) are aggregated by weighting the characteristic by plot size.

[34]The authors use the sum of the predicted number of dams in every district to construct the aggregated predicted number of upstream dams for each district. They then use these predicted values as instruments for the actual number of dams.

[35]Standard errors are correct via OLS.

all plots $p \in P$ which is equivalent to summing over the plot-specific predicted harvest values given by Step 1 in Equation (8). Thus, Step 2 is

$$\hat{\mathcal{H}}_{ht} = \sum_{p \in P} \hat{\mathcal{H}}_{pht} \tag{9}$$

$$= \sum_{p \in P} \hat{\Pi}\big(\hat{\zeta}_{pht} + \hat{\zeta}_{pht} * \hat{l}_{ph}\big). \tag{10}$$

Then, in Step 3 the first stage is regressing actual household-level harvest value on predicted household-level harvest as given by Step 2.[36] Values are again demeaned to account for household fixed effects. Now, all variables are at the household level. Finally, we arrive at Step 4 which is Equation (6) with harvest value, $\mathcal{H}_{ht}$, instrumented by its predicted counterpart, $\hat{\mathcal{H}}_{ht}$. Note that, in practice, Steps 3 and 4 are executed concurrently using 2SLS to obtain correct standard errors.

[36]The first stage produces an F-statistic of 346.81.