NBER WORKING PAPER SERIES

HUMAN CAPITAL INVESTMENT IN THE PRESENCE OF CHILD LABOR

Natalie Bau
Martin Rotemberg
Manisha Shah
Bryce Steinberg

Working Paper 27241
http://www.nber.org/papers/w27241

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2020, Revised June 2025

We are grateful to seminar and conference participants at Ashoka, BREAD, Dartmouth, Duke, IFPRI, Kent, NBER Children's and Development Meetings, Northwestern, NYU, UChicago, UCSD, NEUDC, and PacDev for insightful comments and suggestions. We would like to thank David Atkin, Eric Edmonds, Michael Gechter, and Abhijeet Singh for helpful advice. We thank Maria Elena Guerrero, David Henning, Gabriela Rubio, and Yiran Xie for their excellent research assistance. Bau, Shah, and Steinberg gratefully acknowledge funding from NSF grants #1658852 and #2049936. This paper uses secondary data and therefore does not have a pre-analysis plan. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Human Capital Investment in the Presence of Child Labor
Natalie Bau, Martin Rotemberg, Manisha Shah, and Bryce Steinberg
NBER Working Paper No. 27241
May 2020, Revised June 2025
JEL No. I2, J1, O12

ABSTRACT

Policies that improve human capital in early life are a promising tool for altering the lifelong trajectories of disadvantaged children. However, in many low-income countries, children and their parents face trade-offs between schooling and productive work. If there are positive returns to human capital in child labor, then children who receive more of some types of early life investments, such as those that improve health or nutrition, may attend less school. Exploiting early life rainfall shocks in India as a source of exogenous variation in early life investment, we show that child labor attenuates the positive effects of early life investment and increased early life investment increases school dropout in districts with high child labor. This lower investment in education has persistent long-term consequences in adulthood, resulting in lower household consumption. We show that our results are robust to instrumenting for child labor prevalence using local crop mix. Reductions in educational investment in response to positive early life shocks appear to reduce overall welfare in high-child labor districts.

Natalie Bau
University of California, Los Angeles
Department of Economics
and NBER
nbau@g.ucla.edu

Martin Rotemberg
New York University
mrotemberg@nyu.edu

Manisha Shah
University of California, Berkeley
Goldman School of Public Policy
Department of Economics
and NBER
ManishaShah@berkeley.edu

Bryce Steinberg
Brown University
and NBER
bryce_steinberg@brown.edu

1 Introduction

Policies that increase human capital investment during the critical period between the ages of zero to five, when the developing brain is most plastic, are a promising tool to increase long-term human capital attainment (Knudsen et al., 2006). The World Bank has made hundreds of investments in early childhood development around the world, spending billions of dollars (Sayre et al., 2015). One motivation for these programs is that the already meaningful direct benefits of early life investments may be amplified by "dynamic complementarities" in the human capital production function, as early skill investments increase the returns to later human capital investments, endogenously leading to increases in those investments (Cunha and Heckman, 2007).

In this paper, we show that the success of at least some types of these interventions in low-income countries – such as those to improve early childhood health and nutrition – likely depends critically on the economic environment. This is because investments in early life can also increase the returns to children or adolescents leaving school and generating income for their families, formally in the market or in home production (Bleakley, 2010a). This is a major concern, as the prevalence of child labor is still incredibly high in low-income countries. The International Labour Organization estimates there are approximately 265 million working children in the world—almost 17 percent of the worldwide child population (Ortiz-Ospina and Roser, 2020). Although much of the literature on early life investment has focused on settings where child labor is rare, understanding how parents and children respond to positive early life shocks is particularly important in low-income countries, where child labor is common (Edmonds, 2007; Bharadwaj et al., 2013). If children have opportunities to work productively and their early life human capital increases the returns to doing so, actions taken by parents and children in response to positive early life shocks can reduce or even reverse their positive direct effects on education. In the presence of market failures, these actions may lead to inefficiently low human capital investments. Understanding how early life investment affects opportunity costs in these particularly vulnerable settings is therefore crucial for the design of policies that seek to harness the benefits of early life investment to increase education.

To understand the interaction between child labor opportunities and early life investment, we build a model of human capital investment in a context where children work. The model shows that increased early human capital will not translate into increased schooling – even in the presence of dynamic complementarities – if there are positive returns to this type of human capital in the market for child labor and these positive returns dominate dynamic complementarities in the human capital production function. We test our model using changes in early life investment due to rainfall shocks in India; positive rainfall shocks increase yields and income (Jayachandran, 2006) and improve consumption. Hence, we view the investments we study as closest to interventions to improve health and nutrition during pregnancy and early in life. The test of our model builds on earlier work by Shah and Steinberg (2017), who show that in rural India, positive early life rainfall shocks act as an exogenous boost to initial human capital and on average lead to long-run increases in schooling. In this paper, we answer a different question: under what circumstances will a positive shock to early life human capital fail to translate into increases in overall schooling or even reduce schooling, harming long-term outcomes? We test the model empirically by estimating the effects of early life rainfall shocks on dropout (up to 16 years after the shock) and adult consumption (decades later), separately in places with and without a high prevalence of child labor.

We find that child labor mitigates the positive educational effects of early life shocks due to rainfall. In districts with a high prevalence of child labor, increased early life investment *reduces* schooling. We argue that the negative effect of increased early life investment on education in the presence of child labor is due to increases in the opportunity cost of schooling. Indeed, it is hard to rationalize this pattern with alternative explanations. While the human capital of children in high child labor districts could be less responsive to positive early life shocks if districts with high child labor also have worse schools or less demand for education, this does not explain why positive early life shocks *reduce* education in these districts. Nonetheless, to address the concern that high child labor districts differ from low child labor districts on a variety of dimensions, we help account for omitted variable bias by implementing an IV strategy exploiting a technological source of variation in the demand for child labor, local crops, which are primarily driven by regions' agroclimatic features. Children are known to have a comparative advantage in some crops, such as sugar and cotton (Levy, 1985). To choose crop-based instruments in a principled way, we follow Belloni et al. (2012) and use lasso to select instruments for child labor prevalence. The resulting instruments include sugar and cotton, along with several other crops. Across outcomes, the instrumental variables regressions deliver similar results to OLS.[1] Our IV results imply that in high child labor districts, switching a child from one early life drought to a year with good rainfall *lowers* his total schooling by 0.08 years.

Decreased educational attainment does not necessarily imply that reduced schooling is welfare-reducing in a net present value sense, even when associated with long-run decreases

[1] To further rule out the possibility of omitted variable bias, in supplementary analysis, we control for a battery of other local characteristics, including average incomes, literacy rates, and measures of school quality, as well as household-specific socioeconomic controls and household fixed effects. The inclusion of this rich set of controls has little effect on our results.

in consumption (Baland and Robinson, 2000). Perhaps increased dropout reflects efficient households choosing additional income today in lieu of greater income in the future. Indeed, it is not even obvious if there are costs to dropping out, as it is possible that the returns to schooling over working are very low or even zero (Beegle et al., 2009), especially given criticisms of the Indian educational system (Pritchett, 2013). We address the key question of efficiency in two ways. First, to compare the costs and benefits of the educational choices, we study consumption. We find that there are contemporaneous benefits to dropping out and working: in places with high child labor, households whose children have experienced more positive early life shocks have higher current consumption. However, there are also long-run costs: in places with high child labor, households whose male household heads have experienced more positive early life shocks have lower current consumption. Our IV estimates show that, in high-child labor districts, switching a household head from one early life drought to a year with good rainfall lowers his family's consumption as an adult by 0.9%. Based on these estimates, we estimate the discount factor for a unitary household that would rationalize increasing current child labor at the expense of future consumption. The estimated discount factor, 0.88, is lower than social discount factor estimates and is also inconsistent with Indian interest rates over the same period. Since the decrease in long-run consumption is too large to be rationalized with discounting the future, this suggests that even though these reductions in education were demand-driven, they may represent an inefficient loss of human capital investment.

Our second test directly considers the mechanism underlying inefficient human capital investment. Parents may inefficiently underinvest in the education of their children in response to positive shocks because there are incomplete contracting problems between parents and children. These problems occur when imperfectly altruistic parents do not make efficient educational investments because they cannot capture the returns to these investments in the future (Banerjee, 2004; Ashraf et al., 2020; Bau, 2021). To test whether incomplete contracting problems lead to inefficient investment in response to positive early life shocks, we leverage the fact that in rural India, oldest sons traditionally stay home and care for parents in their old age (Jayachandran, 2015; Jayachandran and Pande, 2017). This allows parents to capture more later-in-life benefits from educational investments in oldest sons, incentivizing imperfectly altruistic parents to invest in these sons' education. Consistent with the importance of incomplete contracting, we find that for oldest sons (but not their siblings), parents reinforce early life investments in human capital regardless of child labor prevalence. Altogether, our results suggest that declines in human capital investment due to opportunity costs related to youth and child labor are inefficient.

Our findings contribute to several literatures. First, they are in line with the growing literature that emphasizes the importance of opportunity costs as a driver of human capital investment, both in developed and developing countries (Atkin, 2016; Charles et al., 2018; Cascio and Narayan, 2022; Shah and Steinberg, 2017, 2021). Our contribution to this literature is two-fold. First, while there is a broad consensus that greater labor market opportunities for low-skill workers can reduce schooling, we show that, because there are returns to human capital even in low-skill work, these labor market opportunities can induce dropout among precisely those students for whom schooling is more productive. Second, the policy-relevance of the literature is limited by a dearth of evidence on whether these dropouts are efficient. This paper provides evidence that at least some of the dropouts related to opportunity costs are inefficient.

We also contribute to a growing literature on the medium and long-run effects of early life shocks, largely though not entirely in higher-income countries.[2] This literature has suggested that "good" early life shocks lead to "good" later life outcomes. In line with this, we do find that early life rainfall shocks have positive effects when we average across all districts. However, focusing only on average effects masks important heterogeneity. Understanding the effects of early life investments in low-income countries requires taking opportunity costs into account. Even in places with a moderate level of child labor, the long-run impacts of early life investments on human capital will be smaller than would be predicted from interventions in low child labor settings, where parents and children do not trade-off work and education. When child labor is high, early life investments can negatively affect education and even adult consumption and welfare. While Bleakley (2007) discusses this theoretical possibility, to our knowledge this is the first empirical evidence that positive early life shocks decrease schooling for a large population. The heterogeneous effects of early life shocks that we identify also have important implications for policy. Interventions like conditional cash transfers can help policymakers harness the benefits of early childhood investment in settings where opportunity costs are also responsive to early life human capital, while the targeting of early life investment policies to maximize either educational outcomes or welfare depends critically on the prevalence of child labor.

The setting and variation in this paper are most similar to Shah and Steinberg (2017), who contribute to both the opportunity costs and long-run effects literatures with two distinct analyses. Shah and Steinberg (2017) estimate the average long-run effects of early life rainfall on educational attainment in India *and* separately estimate the contemporaneous effects of

[2] While this literature is too large to do justice to, notable contributions include Maluccio et al. (2009), Attanasio et al. (2020), Bleakley (2010b), and García et al. (2020). See Heckman (2007), Almond and Currie (2011), Currie and Vogl (2013), Currie and Rossin-Slater (2015), and Almond et al. (2018) for reviews of the literature and Maccini and Yang (2009) and Shah and Steinberg (2017) for two closely-related papers exploiting rainfall shocks in low-income countries.

rainfall on educational investment for older children. They conclude that early life rainfall shocks increase educational attainment on average due to increased early life human capital investment, while contemporaneous shocks decrease enrollment due to increased opportunity costs. This is because contemporaneous shocks increase agricultural productivity, raising wages and directly increasing the opportunity cost of school. Shah and Steinberg (2017) do not link early life rainfall shocks to later-life opportunity costs. In contrast, this paper examines whether early life shocks, which increased household income during a critical period for child investment, affect the opportunity cost of schooling in adolescence. We show that positive early life shocks can *reduce* later educational investment in some places due to increased opportunity costs – a wholly novel result. In addition, unlike Shah and Steinberg (2017), this paper also provides evidence on the efficiency of the educational response to increased opportunity costs and investigates very long-run average and heterogeneous effects of early-life shocks on consumption, an outcome more closely associated with well-being.

We additionally contribute to the literature on human capital investment and dynamic complementarities (Cunha and Heckman, 2008; Gilraine, 2018; Aizer and Cunha, 2012), taking into account an important feature of developing countries that has previously been omitted from this literature – that children work (Schultz, 1960; Basu and Pham, 1998; Basu, 1999; Edmonds and Pavcnik, 2005). This literature often takes a revealed preference approach to argue that an increase in educational investment in response to an earlier positive shock is evidence of dynamic complementarities in human capital production.[3] In our setting, the presence of dynamic complementarities does not guarantee that early boosts to human capital will be reinforced by additional investments later on. If children can use their human capital to earn more in the labor market or produce more on the family farm, they may even *reduce* schooling investment in response to a positive shock to early human capital.

Finally, we contribute to the relatively new literature showing that childhood location has significant and long-lasting effects on adult outcomes (Chetty and Hendren, 2018; Chyn and Katz, 2021). While this literature has focused on relatively wealthy countries (Deutscher, 2020; Laliberté, 2021; Nakamura et al., 2022), our results from rural India show similar patterns. Being born in a district with high child labor has important long-term impacts on human capital accumulation and later life consumption, and, as in Aloni and Avivi (2024), we emphasize the importance of heterogeneity in place effects for different types of children.

[3]See for example, Aizer and Cunha (2012); Gilraine (2018); Johnson and Jackson (2019); Rossin-Slater and Wüst (2020); Duque et al. (2023); Adhvaryu et al. (Forthcoming); Agostinelli and Wiswall (Forthcoming); Goff et al. (2023) among others. A related literature, primarily in developing countries, estimates the extent to which parents invest unequally in their children in order to reinforce or mitigate early differences in human capital (Behrman et al., 1994; Adhvaryu and Nyshadham, 2016; Bharadwaj et al., 2018; Dizon-Ross, 2019), whether due to dynamic complementarities or other convexities in human capital returns, and finds mixed results.

The paper proceeds as follows. To guide the empirical analysis, Section 2 introduces a theoretical framework for human capital investment and child labor in the presence of dynamic complementarities and derives testable predictions. Section 3 provides further background on child labor in India and describes the data used in the analysis. Section 4 describes both the OLS and the lasso IV empirical strategies, and Section 5 reports the primary results on education using a variety of specifications. Section 6 reports the long-run effects of early life investment on adult outcomes in the presence of child labor. Section 7 reports the results from a series of robustness tests. Section 8 discusses the results and provides evidence that parents are not choosing welfare-maximizing education levels, and Section 9 concludes.

2 Theoretical Framework

To develop testable predictions about the effects of early life human capital investment on education and child labor, we develop a simple partial equilibrium model. Intuitively, this model brings together the theoretical literature on child labor (e.g. Basu and Pham, 1998), the trade-off between child labor and human capital formation (Basu, 1999; Baland and Robinson, 2000; Dessy, 2000; Hazan and Berdugo, 2002; Ravallion and Wodon, 2000; Beegle et al., 2009), and the literature on dynamic complementarities (Cunha and Heckman, 2007). Doing so allows us to clarify the circumstances under which positive early life human capital investments can reduce schooling, even in the presence of dynamic complementarities in the human capital production function.

The model captures the following intuition. With only dynamic complementarities, increased early life human capital investment positively affects the returns to later schooling investment, incentivizing parents to invest more in later education. This is the standard effect of dynamic complementarities posited by Cunha and Heckman (2007). However, the new feature of our model is that, in places where child labor is prevalent, early life investments also affect children's productivity at work, raising the opportunity cost of schooling. Thus, a novel prediction of our model is that this countervailing force attenuates and can even reverse the positive effect of early life investment on schooling. If early life investments increase opportunity costs more than they increase the expected utility the parent derives from the increased return to education, schooling and potentially long-run consumption will fall. Furthermore, if the parent is imperfectly altruistic, or she underestimates the size of dynamic complementarities relative to the effect of early life human capital investments on wages, reductions in education due to early life investments can be inefficient and total welfare-reducing.

2.1 Set Up

The decision-maker in the model is a parent, and each parent has one child. The decision-maker is indexed by her child's exogenous educational ability, α, which is distributed according to the continuously differentiable function F. She is also indexed by her type of district, $d \in \{low, high\}$, which denotes whether a district has high or low child labor. To simplify exposition, at the risk of abusing notation, subscripts for these indices are suppressed when not relevant. There are three periods in the child's life: early life, school age, and adulthood. Exogeneous ability α becomes observable in period 2, when a child is old enough to attend school. In period 1, the parent decides how much to invest in a child's early life human capital, h. In period 2, the parent makes a discrete decision whether or not to educate the child, $e \in \{0, 1\}$, or have the child work. If the child works, the parent receives $w^c_{2,d}(h)$, which depends on h and d; $w^c_{2,d}(h)$ can either be a child's actual wage on the market or the value of what she produces at home.

The parent's consumption in the first two periods – when the parent is making human capital investment decisions – is explicitly included in the model. In addition, the parent also places some weight on the child's third period adult utility. This can be thought of as capturing both altruism and a reduced-form representation of the parent's third period consumption, which is determined by the fraction of the child's third period utility that the parent captures as old age support. A parent's preferences in period 1 are represented by

$$U^p_1(h) = u(c^p_1(y_1, h)) + \rho E\left(\max_e u(c^p_2(y_2, e, h)) + \delta U^c(c^c_3(e, h))\right),$$

where c^p_1 and c^p_2 are the parent's consumption in periods 1 and 2, c^c_3 is the child's adult consumption in period 3, u is the parental utility function, U^c is the child's adult utility, which depends on educational and early life investments, and $\delta = \rho\gamma$ is the product of the parent's discount factor ρ and γ, where γ captures both the parent's altruism toward the child and her ability to resolve incomplete contracting problems by extracting utility from the child in the third period. The expectation is taken over realizations of α. Both u and U^c are assumed to have diminishing marginal returns in consumption.

The parent's period 2 utility is given by

$$U^p_2(e, h) = u(c^p_2(y_2, e, h)) + \delta U^c(c^c_3(e, h)).$$

For simplicity, the model abstracts away from borrowing and saving.[4] In line with this

[4] In the empirical analysis, we show that the results are robust to specifications that explicitly control for saving by controlling for household fixed effects. These specifications compare siblings within the same household (with the same budget constraint) who received different shocks. We also test directly for whether

assumption, formal banking was rare during our study period. In 2011, the World Bank Findex data report that only 35% of Indians had bank accounts, and financial access was likely much lower in the rural population we study.

Parental consumption in period 1 is equal to some exogenous income y_1 net the cost of human capital investment h. Parental consumption in period 2 is total income y_2 net the cost of schooling if $e = 1$ or plus the wages from child labor if $e = 0$. Thus,

$$c_1^p = y_1 - c_h h$$
$$c_2^p = y_2 + (1 - e)w_{2,d}^c(h) - c_e e$$
$$c_3^c = w_3^c(e, h) + \alpha e$$

where c_h is a cost of the human capital investment and c_e is the cost of education. The expression $w_3^c(e, h) + \alpha e$ is what the parent believes to be the child's total adult wage, where the function $w_3^c(e, h)$ allows for a flexible relationship in adult wages between e and h and does not directly depend on d, and the returns to education also depend on exogenous schooling ability α.[5] Parents may have incorrect beliefs about $w_3^c(e, h)$, such that $w_3^c(e, h) \neq \tilde{w}_3^c(e, h)$, where $\tilde{w}_3^c(e, h)$ is the true relationship. Following Cunha and Heckman (2008), parents perceive that there are dynamic complementarities in the adult wage function if $\frac{\partial w_3^c(1,h)}{\partial h} > \frac{\partial w_3^c(0,h)}{\partial h}$. This captures the idea that early life investments in human capital make educational investments more productive.

Before solving the model, we make two assumptions. First, for expositional simplicity, we assume that $w_{2,low}^c(h) = 0$, so that if child labor in a district is negligible, child output is always equal to zero. This assumption simplifies algebra but is not necessary for our main propositions (Propositions 1, 2, 3a, and 4a-b). Second, in places where child labor is high, we assume $\frac{\partial w_{2,high}^c}{\partial h} > 0$. This assumption captures the idea that child wages are increasing in child human capital – our key mechanism of interest – and that this effect is stronger in high child labor districts. We provide empirical support for this relationship in subsection 3.4.

early life shocks change household investment patterns in Section 7.2.

[5]The assumption that α is not directly affected by h does not imply that early life shocks cannot directly affect the returns to education since the function $w_3^c(e, h)$ flexibly allows h to affect the returns to e. Rather the existence of an exogenous α merely captures the fact that there is heterogeneity in the returns to education across students, allowing for differences across students of the same incomes' response to the same shock.

2.2 Propositions

We now solve for the parent's equilibrium investment decisions and relate them to changes in first period income y_1.

Proposition 1. *Denote h^* as the parent's equilibrium choice of h. If $w^c_{2,d}(h)$ and $w^c_3(e,h)$ have constant or diminishing marginal returns in h, then $\frac{\partial h^*}{\partial y_1} > 0$ for all d.*

Proof. See Appendix A.

The first proposition simply delivers the classic result that a positive income shock in early life will increase early life human capital investment. The intuition for this prediction is straightforward. When y_1 increases, the marginal utility of first period consumption falls, increasing the parent's incentive to invest in her child's human capital. This proposition is consistent with the previous findings of Maccini and Yang (2009) and Shah and Steinberg (2017), who show that an early life shock increases test scores and weight.

Building on Proposition 1, the next set of propositions describe the key empirical results in this paper – that early life shocks increase education rates in places with low child labor and have smaller positive or even negative effects on education rates in places with high child labor. Proposition 2 delivers a standard prediction in the dynamics complementarities literature.

Proposition 2. *Denote $\lambda_d(y_1)$ to be the share of children educated in a district of type d given y_1. $\frac{\partial \lambda_{low}(y_1)}{\partial y_1} > 0$ only if $\frac{\partial w^c_3(1,h)}{\partial h} > \frac{\partial w^c_3(0,h)}{\partial h}$.*

Proof. See Appendix A.

This proposition captures the fact that, in low child labor places, increased h only positively affects the parent's educational decisions through its effect on the returns to later life educational investments. Therefore, if an early life shock increases educational investments in low child labor markets, this is evidence in favor of dynamic complementarities.

The remaining propositions introduce the novel predictions of this paper. Proposition 3a shows that the standard dynamic complementarity results can be reversed by opportunity costs. In high child labor markets, positive early life investments can have *negative* effects, despite their potential positive effect on the returns to education due to dynamic complementarities. Proposition 3b (presented in Appendix A) describes the conditions under which opportunity cost effects are not strong enough to reverse the positive effect of early life investment on education but nonetheless dampen that positive effect.

Proposition 3a. *If $\frac{\partial w^c_{2,high}(h^*(y_1))}{\partial h}$ is sufficiently great, then $\frac{\partial \lambda_{high}(y_1)}{\partial y_1} < 0$ even if $\frac{\partial \lambda_{low}(y_1)}{\partial y_1} > 0$.*

Proof. See Appendix A.

Proposition 3a shows that when the effect on parental utility of the increase in child wages due to an increase in y_1 is sufficiently large in high child labor places, it outweighs the effect of the increase in the returns to education (weighted by the parents' altruism and discount factor). Then, positive income shocks that increase early life investments can lead to reduced education. The model also clarifies that (1) observing these effects does not depend on the magnitude of the increase in early life human capital but rather the size of the derivative of the child wage with respect to h, as the change in the returns to education vs. child labor is what matters for the marginal parent's decision, and (2) even for small shocks, there will be some households on the margin of education vs. child labor as long as the prevalence of child labor is neither 0 or 100%.

Finally, our last two propositions consider some plausible circumstances under which these reductions in education will be inefficient. These sources of inefficiency in educational investment appear in other work (for example, Banerjee (2004) and Bau (2021) on intergenerational incomplete contracting and Jensen (2010) on systematic under-estimation of the returns to schooling). Our contribution is showing that, in conjunction with the existence of child labor, these forces can cause increased early life investments to have perverse effects and reduce total welfare. Additionally, modeling these sources provides us with tests for whether reductions in schooling in response to increases in early life human capital in high child labor districts are inefficient. We view an educational investment decision as inefficient if it does not maximize total welfare $W_2(e;\alpha,h)$, which is the sum of the parent's and child's utilities (equivalent to setting $\gamma = 1$ in $U_2^p(e,h)$).

Proposition 4a. *If* $\gamma < 1$ *or* $\frac{\partial w_3^c(h,1)}{\partial h} < \frac{\partial \tilde{w}_3^c(h,1)}{\partial h}$ *(where* $w_3^c(h,0) = \tilde{w}_3^c(h,0)$*), then an increase in* y_1 *may inefficiently reduce education.*

Proof. See Appendix A.

Proposition 4a captures two intuitive circumstances under which the reductions in education due to the increase in y_1 (under Proposition 3a) may be inefficient. The first case $\gamma < 1$ captures the idea that an imperfectly altruistic parent who cannot perfectly contract with her child to capture the returns to her investments during childhood will underweight the increase in a child's utility in the future relative to the increase in consumption today. Thus, an increase in y_1 will reduce the parent's payoff to educating the child, even though the increase in y_1 increases the payoff from education for total household utility. The second case $\frac{\partial w_3^c(h,1)}{\partial h} < \frac{\partial \tilde{w}_3^c(h,1)}{\partial h}$ captures the idea that dynamic complementarities are hard to observe and even a perfectly altruistic parent may underestimate them. Thus, the parent will underestimate the increase in the returns to education for a child's adult wages due to an increase in y_1 relative to the increase in the child wages, again leading the reduction in education to be inefficient.

The final proposition focuses on the case where $\frac{\partial w_3^c(h,1)}{\partial h} = \frac{\partial \tilde{w}_3^c(h,1)}{\partial h}$ and motivates a test for whether reductions in investment in response to increases in y_1 are inefficient.

Proposition 4b. *Define the cut-off value for α above which a child is educated as α^*. If $\frac{\partial w_3^c(h,1)}{\partial h} = \frac{\partial \tilde{w}_3^c(h,1)}{\partial h}$ and $\frac{\partial W_2(1;\alpha^*,h)}{\partial h} > \frac{\partial W_2(0;\alpha^*,h)}{\partial h}$, then there exists a value $\bar{\gamma}$ such that for $\gamma > \bar{\gamma}$, $\frac{\partial \alpha^*}{\partial h} < 0$.*

Proof. See Appendix A.

This proposition focuses on the case where an increase in h increases the net value of educating the marginal child more than it increases the value of the child working. In this case, it would be efficient for the marginal parent to respond to an increase in h by investing in education. If the parent is sufficiently altruistic or sufficiently able to contract with the child to capture the return to her investment ($\gamma > \bar{\gamma}$), the value of α needed for a child to be educated will fall and the marginal child will be educated. In other words, if γ is sufficiently high, the parent values the child's adult utility enough that she responds to the increase in the returns to education by efficiently increasing educational investment. This is easy to see in the extreme case where $\gamma = 1$ and a parent is either perfectly altruistic or perfectly able to contract with her child. In that case, education levels are never inefficient, and there will never be an inefficient reduction in educational investment in response to an increase in h. This proposition indicates that if there is incomplete contracting between parents and children, we should see heterogeneity in the response to early income shocks across parent-child pairs with different underlying incomplete contracting problems.

3 Background and Data

In this section, we describe the datasets used in this paper and report basic facts about child labor in India. We then provide evidence that early life rainfall shocks increase early life human capital and that this effect is similar in both high and low child labor districts. Finally, we show that, consistent with the mechanisms in the theoretical framework, greater human capital (in the form of both height and lagged test scores) is associated with higher child wages.

3.1 Data

This paper utilizes a variety of datasets from India, which we summarize in Table 1. The summary statistics for our main outcomes of interest are reported in Table 2. For all datasets, we restrict the sample to include only rural households since rainfall shocks affect incomes through crop yields. We describe the data in more detail in Appendix B.

Our primary outcome, dropout, comes from the Annual Status of Education Report (ASER), an annual household survey of approximately 500,000 children across every rural district in India since 2005. We supplement these data with the ASER school-level survey, which provides data on classroom-level attendance in a given village, measured by ASER surveyors. In addition to ASER, we also use data from the National Sample Survey (NSS) from 2004-2012 on the number of meals taken at school, which provides an additional measure of school attendance. Finally, we use NSS data on household consumption both as a long-run outcome for adults and to measure the impact of children's work contemporaneously on their family's consumption.

In order to measure the district-wide prevalence of child labor, we again turn to the NSS. For each district and year, we measure the percent of children age 5-17 who report working as their primary activity, where work includes both salaried and wage work, as well as unpaid work in home enterprises and domestic work. We then calculate the share of years the district is above median, omitting the current round in order to avoid any mechanical correlations. This is our primary measure of the prevalence of child labor.

We also use the NSS to create a pool of potential instruments for the prevalence of child labor. We exploit the fact that the NSS round 68 asks respondents for their "principal industry" at a dissagregated level. Our pool of potential instruments is then the share of adults working in agriculture who report working in each disaggregated agricultural industry. We outline the construction of the potential instruments in detail in Appendix B (see "National Social Sample, Schedule 10: Local Crop Mix").

The data on rainfall shocks are from the University of Delaware Gridded Rainfall Data (version 5) for 1957-2008. Following the literature (Jayachandran, 2006; Shah and Steinberg, 2017), we define a "rainfall shock" as equal to one if rain is in the top 20th percentile for the district, -1 if it is in the bottom 20th percentile, and 0 otherwise.[6] We form an aggregate early life rainfall shock measure denoted ELR by summing over the shocks when the child is in utero (age$= -1$), in her first year of life (year of birth), and in her second year of life. Thus, the aggregate shock variable ranges from -3 to $+3$. We match these data to individuals in all the other datasets using their birth year and district.

Lastly, we use several supplemental data sets for additional outcomes and robustness, which are outlined in detail in Appendix B.

[6]In India, though flooding occasionally occurs in rural areas, more rain is essentially always better for crop yields. See Jayachandran (2006), Kaur (2019) and Santangelo (2019) for more discussion of the direct relationship between rainfall and crop yields.

3.2 Background on Child Labor in India

Officially, child labor for children aged 14 and under has been banned in India since 1986. However, the ban covers only certain industries and has not been well-enforced.[7] The main employers of child labor, agriculture and family-run businesses, are exempted from the ban. Beyond the various exemptions, the ban itself may have increased child labor through negative income effects (Bharadwaj et al., 2013).

Overall, child and adolescent labor are common in India, as is the case in many low-income countries. In our data, 10% of children aged 5–17 report working as their primary or secondary activity. This statistic can be considered a lower bound for the extent of child labor since it typically captures children who spend the majority of their time working and may not include those who still spend a substantial fraction of their time on work or domestic labor. While 30% of individuals 15–17 report working as their primary activity, child labor is not entirely driven by older children who have aged out of school. Among adolescents, 18% of children aged 13–15 work. Likewise, UNICEF (2011) estimates that 28 million children in India aged 5-14 are engaged in work.[8]

Figure 1 maps the variation in the percent of children 5-17 (across all NSS rounds) who report working as their primary activity across Indian districts. The figure shows that there is substantial geographic variation in child labor and that areas with a high prevalence are scattered throughout the country. The most common industries for these children are agriculture and domestic duties, and children both work in the labor market for pay and part-time at home or on family farms. Among children aged 5–17 who report working as a primary activity, the most frequent types of activities are domestic duties (28%), working as an unpaid helper in a household enterprise (24%), and working as a causal laborer outside of the household (21%).

3.3 Early Life Rainfall and Human Capital

Before turning to our main empirical strategies, we provide evidence on two key preliminaries. In this subsection, we document the link between early life rainfall and early life human capital investment, which we will exploit for identification. In the next subsection, we examine whether children with greater human capital appear to have a greater opportunity

[7]Industries where child labor is banned include occupations involving the transport of passengers, catering establishments at railway stations, ports, foundries, handling of toxic or inflammable substances, handloom or power loom industry, and mines. Processes banned include hand-rolling cigarettes, making or manufacturing matches, explosives, shelves, and soap, construction, automobile repairs, and the production of garments (Bharadwaj et al., 2013).

[8]For domestic work to count under this definition, a child must be engaged in domestic work for over 28 hours per week.

cost of schooling, consistent with the key mechanism in the theoretical framework.

To test the relationship between the early stock of human capital and schooling and child labor, we use early life rainfall shocks as a proxy for shocks to early life human capital. It has been well-documented that positive rainfall shocks increase crop yields, increasing wages (Epstein, 1962). To give a sense of the magnitude of our shocks, Jayachandran (2006) uses the same measure and finds that a positive rainfall shock is associated with an increase in crop yields of 7%. We expect that rainfall shocks lead to higher early life (and prenatal) consumption, and that in this context, increased early life investment likely takes the form of increased calorie consumption by mothers and infants.

This is consistent with existing work showing that positive weather shocks early in life positively impact human capital in developing countries (Maluccio et al., 2009; Maccini and Yang, 2009; Shah and Steinberg, 2017). There is substantial evidence of a link between weather shocks and undernutrition specifically in children and adults (Carpena, 2019; Le and Nguyen, 2022), and between the in utero environment and adult height (Well, 2007; Behrman et al., 1994; Alderman et al., 2006). In rural India, Kumar et al. (2016) find that early life shocks have lasting effects on children's health and argue that an important mechanism is changes in food consumption. Agarwal (1992) argues that adult men bear relatively less of the burden of reduced consumption in response to negative shocks. It is possible that other parental investments in this critical period, such as medical care, time spent engaging with children, or vaccines could also respond to income shocks. However, given the plethora of evidence for the importance of nutrition, we posit that this is likely the main channel that drives the link between weather shocks and early life human capital in this population.

In Figure 2, we establish that early life rainfall affects a measure of health-related human capital. The figure shows the relationship between the aggregate early life rainfall shock (ELR) and height for children and adolescents aged 5 to 17 in the IHDS 2012 separately for places with more and less child labor. To allow us to visually evaluate whether the relationship between ELR and height varies by child labor, we divide districts up into those that have above median child labor either more or less than 50% of the time. This figure plots the relationship using residual variation after conditioning on age and district fixed effects. There is a clear positive relationship between early life rain and height in childhood, which is consistent with increased nutrition investments for children who experienced greater early life rain. This effect does not vary with child labor prevalence, indicating that differences between the effects of early life rainfall on medium and long-term outcomes across districts are unlikely to be driven by differences in the effects of early life rainfall on human capital investment.

3.4 Human Capital and Child Wages

Having established that ELR increases early life human capital, we next explore whether greater human capital is associated with greater child productivity. While productivity is hard to observe in most cases, we observe a proxy for productivity – wages – for the selected sample of children who work for wages. We focus on test scores and height as observable proxies for human capital – which are likely to be correlated with human capital (including health) more broadly. Appendix Table A1 reports results from hedonic regressions of child wages (conditional on working for pay) in the IHDS 2012 on height and lagged test scores. The substantial reduction in the sample size in column 3 is due to the fact that we use test scores from the first round of the survey as a control. This restricts us to children aged 15-17 in 2012 because test scores were only collected for children 8-11 in the first (2005) wave of the survey.

For both measures of human capital, we find a strong positive association between early human capital and child wages: a 1 sd (18 cm) increase in height is associated with a 11% increase in wages, while a child who answers one more math question correctly receives a 5% higher wage. While we caution that, due to selection into working, these regressions should not be interpreted as causal measures of the effect of human capital on the opportunity cost of schooling, these descriptive results are consistent with greater human capital increasing the opportunity cost of schooling. While wages provide us with a useful observable measure of productivity, the same mechanisms are likely to be important for children who work without wages (e.g., on family farms). Though we cannot observe marginal products in these cases, it is still likely that labor productivity grows with human capital.

We cannot directly test if early life rainfall causally increases children's productivity or whether this effect is greater in higher child labor districts. This is because, as we will show, early life rainfall (1) affects the probability of a child working, and (2) does so with a different sign in high and low child labor districts. As a result, estimates from regressing wages on early life rainfall and its interaction with whether a child is in a high or low child labor district may be driven by marginal children (with lower wages) moving into and out of working.

4 Empirical Strategy

The theoretical framework predicts that the effects of early human capital investments on later schooling investment will depend on the opportunity cost of children's time. In subsection 3.3, we established that rainfall shocks experienced in utero and in the first two years of life provide exogenous variation in the stock of early human capital. In this section, we outline the OLS and IV empirical strategies for the remainder of the paper.

OLS Strategy. In the primary OLS specifications, we estimate the following regression

$$y_{idmtag} = \beta_1 ELR_{dta} + \beta_2 ELR_{dta} \times CL_{dt} + \tau_{dmt} + \tau_a + \tau_g + \epsilon_{idmtag} \tag{1}$$

where y_{idmtag} is an outcome measure (i.e. dropped out, attends school, consumption) for individual i in district d in month m and year t at age a of gender g, ELR_{dta} is individual i's early life aggregate rainfall shock, CL_{dt} is a measure of child labor in district d in year t, τ_{dmt} is a district-month-year fixed effect, τ_a is an age fixed effect, and τ_g is a gender fixed effect. As mentioned above, CL_{dt} is a variable for the share of rounds (leaving out round t) that an above median share of the district's children work. β_2 can be interpreted as the differential effect of early life shocks in a district that always has an above median share of children working. We refer to these districts where $CL_{dt} = 1$ as "high child labor districts" and districts where $CL_{dt} = 0$ as "low child labor districts." We also report β_1, the average effect of the aggregate early rainfall shock in low child labor districts, and $\beta_1 + \beta_2$, the total effect of a positive early rainfall shock in high child labor districts.

For educational outcomes, the estimates of β_1, β_2, and $\beta_1 + \beta_2$ provide tests of the different propositions from the theoretical framework. In the case of dropout, where a positive coefficient indicates that a child receives less education, Proposition 2 states that $\beta_1 < 0$ is evidence in favor of dynamic complementarities. Proposition 3b indicates that β_2 should be positive, consistent with the opportunity cost effect at least attenuating the positive effects of early life rainfall on human capital in high child labor places. Finally, Proposition 3a predicts that $\beta_1 + \beta_2 > 0$, indicating that a positive shock *reduces* human capital in high child labor districts if the returns to human capital in child labor are sufficiently great.

District-time and age fixed effects ensure that the estimates are identified from within-district within-cohort variation. Thus, fixed differences across districts (such as experiencing drought more or less often) will not drive the results. Rather, the identifying assumption for β_1 is that, conditional on country-level changes in rainfall patterns over time, deviations from district-level average rainfall are not associated with other time-varying district-level characteristics that may affect children's outcomes. This is the standard identifying assumption from Shah and Steinberg (2017) and Maccini and Yang (2009).

Interpreting β_2 as measuring the causal effect of the interaction between child labor and early life rainfall shocks requires the additional assumption that there is no important district-level characteristic associated with child labor that also leads rainfall to have different effects in high and low child labor districts. While this assumption is strong, we note that it is hard to develop an alternative explanation for why increasing ELR, which we have shown increases early life measures of human capital in both high and low child labor districts, would have *net negative* effects on educational outcomes other than through the opportunity cost channel. Hence, even if the point estimates are biased, we view an estimate of net negative effects of ELR on education in high child labor districts as strongly favoring our theoretical framework. Nonetheless, to ensure our results are robust, we pursue two additional strategies. First, we introduce an instrumental variables identification strategy below. Second, in Section 7, we describe a series of robustness tests, including the inclusion of a large array of controls in both the OLS and IV strategies, supporting the argument that β_2 is driven by the interaction between child labor and ELR rather than ELR's interaction with other district characteristics.

Finally, one potential concern for interpreting the effects of ELR as being driven by earlier human capital investments is that there may be a direct long-run effect of early life rainfall shocks on individuals' outcomes (e.g., if families use the windfall to buy investment assets). District-time fixed effects help control for this since the fixed effects compare families who have faced the same series of rainfall shocks (but whose children received different shocks during the critical period around birth). We also show our results are robust to the inclusion of household-time fixed effects, which fully capture all income and savings differences across households. For our analysis of consumption, district-time fixed effects have the additional benefit of controlling for seasonality (Merfeld and Morduch, 2023).

IV Strategy. We also present results from an IV strategy. In addition to helping address the concern that the OLS results are driven by omitted variable bias, the IV addresses the possibility of attenuation bias due to measurement error. As is well-known, classical measurement error in the explanatory variable will bias the coefficient toward zero in an OLS regression (but not an IV). In our setting, since child labor prevalence is measured with noise, we expect the IV to deliver larger estimates that do not suffer from attenuation bias.

The IV exploits technological variation in children's comparative advantage in working. Children have a relative advantage in some crops due to the nature of the tasks associated with planting, weeding, and harvesting. For example, cotton is known as a child labor crop because it is low to the ground and very lightweight (Levy, 1985). Crop mix across regions in India is mainly driven by agroclimatic conditions, such as average temperatures and rainfall, as well as soil requirements (Krishna, 2014). Thus, agroclimatic conditions create variation in the prevalence of child labor.

Figure 3 graphs the coefficients from a regression of CL_{dt} on measures of crop importance (the potential instruments) for the crops that make up a non-negligible share of adult agricultural employment (>1%). Crop importance is measured as the share of adult agricultural employment in a district in an agricultural industry at the 4-digit NIC code level. We show more common crops since there are a large number of crops with close to zero agriculture employment shares. The size of the markers denotes the share of adult agricultural labor in a given crop. Reassuringly, cotton is the strongest predictor of child labor among these crops. In contrast, crops that require brawn and height, such as tree crops (coconuts, rubber, bamboo) are negatively associated with child labor.

Having confirmed that adult crop mix predicts child labor in ways that are consistent with children's comparative advantage, to identify instruments in a principled way and maximize statistical power, we follow the IV-lasso methodology proposed by Belloni et al. (2012) and use lasso to choose the set of instruments that best predict CL_{dt}. To maintain consistency, we select crops once using our main outcome, dropout in ASER (the first stage of the IV regression in column 3 of Table 3) and then use the same set of instruments throughout the analysis. The second stage regression is the same as equation (1) above. We describe the instrumental variables strategy in more detail in Appendix A2

In practice, the lasso selects cotton, sugar cane, rice, cattle and buffaloes, sheep and goats, wheat, jowar/bajra/millets, other cereals, and other oil seeds as instruments. Notably, cotton and sugar – for which children are thought to have a comparative advantage – are selected. Appendix Figure A1 maps the geography of predicted child labor given the crop shares and the crops selected as instruments. For comparability with the raw data in Figure 1, we predict the mean child labor prevalence rather than the share of rounds child labor is above median. The figure confirms that the instruments generate substantial geographic variation.

5 Key Outcomes During Childhood: Education, Work, and Consumption

In this section, we test the main propositions of the model. Based on Proposition 2, we expect that if there are dynamic complementarities, early life shocks will increase educational investment in districts with low child labor. In districts with high child labor, this effect will be attenuated (Proposition 3b) and may even be reversed (Proposition 3a). We test these predictions for our key outcome, dropout, as well as a measure that may be more sensitive to intensive margin changes in educational investment, classroom-level attendance, in the ASER data. We then verify that the same patterns appear in self-reported enrollment and meals in school (a proxy for attendance) in the NSS data. Furthermore, we show that the results are not sensitive to the specific choice of measure of child labor prevalence. Finally, we provide additional evidence in favor of the mechanisms in the model by examining the heterogeneous effects of rainfall shocks on working for a wage and household consumption.

Dropout (ASER). Table 3 reports the effects of early life shocks on dropout in the ASER data, as well as their interaction with the measure of child labor prevalence (see equation (1)). Column 1 reports the average effects of early life rainfall, while column 2 reports the differential effects using OLS, and column 3 reports the differential effects using the IV specification. The "total effect" row at the bottom of the table reports the aggregate effect of rainfall shocks in districts whose CL_{dt} measure is equal to 1. The results confirm the predictions of the theoretical framework. Consistent with Proposition 2, and with the presence of dynamic complementarities, an increase in ELR_{dta} reduces dropout in low child labor districts (β_1). In contrast, the interaction of ELR_{dta} with high child labor prevalence (β_2) is positive. The human capital-boosting effects of early life rainfall shocks in low child labor districts are attenuated as child labor becomes more prevalent. Indeed, this effect is strong enough that on net, children who experience positive early life human capital shocks in high child labor districts are more likely to have dropped out than their counterparts who did not experience these shocks ($\beta_1 + \beta_2 > 0$). The IV and OLS results are qualitatively similar, though the IV estimates are larger in magnitude for both high and low child labor districts.[9] The fact that early life shocks increase dropouts on net is strong evidence for the importance of opportunity costs. Many of the other potential differences between low and high child labor districts, such as school quality or norms about educational investment, might attenuate the positive effects of early life shocks but are unlikely to *reverse* them.

The estimates in column 3 of Table 3 imply that getting one positive early life rainfall shock relative to a negative one (a change in ELR of +2), in a low child labor district, reduces dropout by 1 percentage point (30%). Adding the enrollment effects over a child's life implies that total years of schooling increases by 0.14 years. In contrast, a positive rainfall shock instead of a negative one increases dropout by 0.64 percentage points (18%) in high child labor districts. Adding up the enrollment effects implies a reduction in years of education of 0.08. Shocks in this range are extremely common in the ASER sample: 28% of children have a ELR_{dta} measure of -1 in our sample, while 17% have a shock of +1.[10] Indeed, even

[9]Motivated by the fact that girls tend to receive less educational investment than boys in India (Lancaster et al., 2008; Himaz, 2009; Azam and Kingdon, 2013), Appendix Table A2 estimates the effects of the rainfall shocks on dropout separately by gender. Both boys and girls are significantly affected by the interaction of early life rainfall shocks in high child labor districts, though the effects are more pronounced for girls. One potential explanation for this heterogeneity is that girls are – on average – less likely to care for parents in their old age and are therefore more likely to experience incomplete contracting problems with parents. We further explore whether incomplete contracting problems lead to inefficient investment in response to increased human capital in Section 8.

[10]Mechanically, from 1957 through 2014, 20% of district/years have positive shocks and 20% have negative shocks. Our sample children happened to have been born during a relatively dry time in India. For children

larger variation in ELR_{dta} affects a still substantial part of the sample; for nearly 15%, the absolute value of ELR_{dta} is greater than or equal to 2. Hence, moving from the 5^{th} to 95^{th} percentile of ELR_{dta} in a high child labor district would increase the annual likelihood of having dropped out by 1 percentage point (30%).

To put these effect sizes into perspective, Duflo (2001) finds that receiving one more school per 1,000 children in a district in Indonesia increased male education by 0.12 years. Thus, the reduction in education in high child labor districts caused by receiving a positive early life rainfall shock instead of a negative one is on the order of two-thirds the effect of receiving another school per 1,000 children in Indonesia. Altogether, these effects are economically meaningful but, unsurprisingly, not as substantial as those of a large-scale school construction program. Indeed, we would not expect a single year's rainfall in early childhood to have dramatic effects on a child's outcomes. However, by studying these shocks, we hope to not only identify an economically important shifter of human capital investment but also improve our understanding of households' human capital investment decision-making.

We also explore how the effects of early life rainfall on drop out change with age. We allow the effects of ELR_{dta} to vary by age in equation (1) and include controls for age-by-sex, age-by-district, and age-by-year fixed effects to account for potentially differential effects of age across different types of districts or over time. Figure 4 reports the estimated total effect of a 1 unit increase in ELR_{dta} in a district that always has above median child labor for both the OLS and IV specifications. In both specifications, at younger ages, an additional positive early life shock has small effects on dropout, but these effects increase dramatically for older children over the age of 14 (the group that is more likely to be on the margin of child labor). While we do not directly model dropout decisions by age in our theoretical framework, we interpret this figure as evidence in favor of our proposed mechanism; dropout is concentrated among older children for whom the returns to working are likely to be relatively high.

Attendance (ASER). In rural India, actual attendance rates may be low even if children are officially enrolled in school. Thus, we also estimate effects on the school-grade-level measure of attendance, which captures both extensive and intensive margin changes in educational investment. Note that since attendance is measured at the school-grade-survey level, we base the timing of the shock for a grade on the statutory age for that grade. The results are again consistent with the predictions of the model. Columns 4-6 of Table 3 show that a positive early life shock increases attendance in low child labor districts (0.35 children per classroom in the IV specification, around 2% of the average enrollment). These positive effects attenuate as child labor becomes more prevalent, and in high child labor districts,

born after our sample, rainfall reverted back to its long-run likelihood of positive shocks.

positive shocks significantly decrease attendance (0.46 children per classroom, or 2%).

Alternative Schooling Measures (NSS). We also reproduce our benchmark results from the ASER data using alternative schooling measures in the NSS. In columns 1–3 of Table 4, the outcome measure is an indicator variable for responding "attends school" to the question about a child's primary activity.[11] The results are similar to those in ASER. Early life rainfall shocks increase self-reported enrollment in low child labor places, this positive effect declines with increases in child labor, and in high child labor places, the effect reverses and becomes negative (albeit the net effect is not significant). Lower precision in the NSS is not altogether surprising since it contains data for many fewer individuals than ASER.

Columns 4-6 report the results for eating any meals at school in the previous month, an alternative proxy for school attendance. Columns 7-9 report results for eating over twenty meals at school in a month, a proxy for more intensive attendance. The same pattern emerges for meals as self-reported school attendance, though the total effect in high child labor places is now significant in our IV specification. One natural question is how much of the effect of early life shocks matters at the intensive versus extensive margin. The total effects in our IV specification are similar for any meal and more than 20 meals (around a .01 decrease in high child labor districts), suggesting that the response to increased human capital is likely driven by extensive margin dropout.

Alternate Specifications for Child Labor Prevalence While the baseline regressions measure CL_{dt} as the leave-own out share of surveys where child labor is above-median, our results are robust to alternative choices. In Figure 5, we report our primary estimate from Table 3, the total effect of an early life shock on dropout in a high child labor district, using different cutoffs besides the median for calculating the share of rounds a district is above the cutoff. In this case, a "high child labor district" is a district that is always above the cut-off given by the x-axis. We plot the total effect using 20th, 40th, 60th, and 80th percentile cut-offs. As the cut-offs grow, the total effect on dropout in high child labor districts monotonically increases. In the IV specification, there is an insignificant effect of 0.0005 with a (very low) cutoff at the 20th percentile that grows to an effect of 0.009 for districts above the 80th percentile of the child labor distribution. This implies that districts with the most child labor experience the strongest increase in dropout in response to positive early life shocks. While not a formal prediction of the model, this is in line the

[11] Because virtually all children are either coded as attending school or working (as we define it) for their primary activity, the results with working as the primary activity would be nearly identical except with flipped signs.

logic of Proposition 3b that the positive effects of early life shocks on dropout are increasing with the returns to child labor.

One potential concern is that our results are an artifact of the specific zero-one measure of the district-level prevalence of child labor. In Table 5, we report the results using the continuous variation in the share of children in each district who report working as a primary activity in the NSS. The qualitative pattern of the results is the same: children who receive positive shocks in low child labor districts are less likely to dropout, while the effects of these shocks attenuate and are eventually reversed in higher child labor districts. We can also evaluate if the estimates are quantitatively similar to those with our preferred measure. A district that always has a below median share of children engaged in child labor has an average share of children engaged in child labor of 3.9%, while a district that is always above median has an average share of 16.5%. Comparing the OLS estimates for dropout across the two tables shows that the implied magnitudes from the two measures are almost identical. Based on Table 5, a 1 unit increase in the early life rainfall shock in a low child labor district reduces the likelihood of dropout by 0.28 percentage points; the estimated effect in Table 3 is 0.29 percentage points. Similarly, the implied effect from Table 5 for the always above median districts is +0.12, identical to the value in Table 3. As the two measures of child labor prevalence deliver nearly identical results, for ease of exposition, we only report results for the share of rounds a district is above-median for the remainder of the paper.

As described previously, for our main specification, we use child labor shares in other survey rounds and exclude the current round to proxy for current child labor prevalence. This helps to avoid bias from common contemporaneous shocks that affect both the child and their neighbors (Manski, 1993). An alternative is to instrument *current* child labor shares, either with the leave-out-mean or with crop shares. Appendix Table A3 shows our results are similar for our key measure, drop out, using this approach.

Alternative Measures of Early Life Shocks. Our primary measure of early life shocks is an omnibus measure, the sum of all early life shocks. Appendix Figure A2 evaluates whether aggregating shocks sacrifices important heterogeneity for our key outcome of dropout. Panel (a) shows that there is limited heterogeneity in the effects of early shocks by age. Similar to Shah and Steinberg (2017), we find similar effects of rainfall shocks in the year before birth, the year of birth, and the year after birth within different district types. While this is consistent with the "first 1000 days" hypothesis, it is somewhat in contrast to previous work, which has found either the in utero period (Almond and Currie, 2011) or the year of infancy (Maccini and Yang, 2009) to be the most important. Critically for our paper, the interaction with child labor is also relatively stable across the three years of investment.

In Panel (b), we disaggregate further by showing the effects of positive and negative shocks separately. While there is some noise across the years, the effects of positive shocks look similar in size (though with the opposite sign) to the effects of negative shocks within district types. The biggest departure from symmetric effects occurs in the utero year, in which the effect of a negative shock is significantly larger than that of a positive shock. This could reflect a particular downside risk for pregnant women in lean times, but given the large number of comparisons in this figure, we do not want to overinterpet a single difference. Altogether, the figure gives us confidence that it is reasonable to combine both positive and negative shocks in yearly measures and sum over the yearly measures to produce an aggregate measure of "early life rainfall."

Working for a Wage and Consumption (NSS and IHDS). In the model, when positive early life shocks decrease education, it is because children are generating resources for their families. To shed light on this mechanism, we now estimate equation (1) with an indicator variable for working for a wage as the outcome. Appendix Table A4 reports the results using the NSS and the IHDS. In both data sets, we find that early life shocks reduce working for a wage among children in low child labor districts, consistent with the increases in education we observed in ASER and the NSS. On the other hand, this reduction in working attenuates as CL_{dt} increases and the total effects in high child labor districts are positive in most data sets and specifications. An important caveat to focusing on wage labor is that children do not only generate resources by working for a wage; we expect home production (including on the family farm) to be important as well. Hence, these estimates may underestimate the extent to which children work in response to positive early life shocks in high child labor districts.

In order to measure changes in the resources generated by children, including by home production, in Appendix Table A5, we report the effect of children's early life rainfall shocks on households' per capita consumption. Since consumption is measured at the household level, each observation is a household. Our aggregate rainfall shock measure now sums over the ELR_{dta} measures of all children aged 5-17 in the household, and we include controls for the number of children. Households with children who received more positive early life rainfall shocks have lower consumption per capita in low child labor districts (consistent with children attending more school). This attenuates with increases in child labor prevalence and becomes positive across all specifications (albeit not significantly so) in high child labor districts. These results are consistent with the model assumption that $\frac{\partial w^c_{2,high}}{\partial h} > \frac{\partial w^c_{2,low}}{\partial h}$, though we caution that household consumption will not capture the latent productivity of children who do not work.

Taken together, the results in this section show an extremely consistent pattern across a range of different outcomes from different data sets. Children who have experienced positive early life shocks in low child labor districts invest more in education, consistent with dynamic complementarities. However, when child labor is more prevalent, this positive effect attenuates, and in districts with high child labor, this relationship is reversed. In these districts, children with higher initial shocks to human capital have lower educational investment and instead provide resources to their family.

6 Long-Term Effects: Adult Consumption

While ELR_{dta} may decrease educational attainment in high child labor districts, it is not obvious that this reduction in education has negative long-term effects. If the schools are of low quality, or if increased experience counteracts the decline in education, then the shift from school to work may not matter in adulthood. In this section, we test for long-term effects on household consumption.

To measure the effects of early investment on long-run outcomes, we study the consumption of households whose heads faced different early life shocks. Since consumption is measured at the household-level, we calculate the exposure of a household to early life shocks using the shock of the male household head.[12] Table 6 reports the results. Columns 1–3 assume children's consumption is one half an adult's consumption, while columns 4–6 assume that children's consumption is one third an adult's consumption. Focusing on column 3, ELR_{dta} increases long-run per capita adult consumption in low child labor places by 0.59% due to a combination of the direct effects on early life investment, as well as subsequent increased schooling. The interaction of early life shocks with high child labor significantly decreases consumption (by 1.1%), enough that the net effect (0.47%) is significantly negative. Columns 4–6 show that the results are similar under the alternative measure of household consumption.

A potential concern is that this specification assumes household heads are still living in their district of birth: in the consumption modules, the NSS asks household heads about their current district but not their birthplace. If rainfall shocks lead to differential migration in high and low child labor districts, then this could induce bias in estimates of the effects

[12] In 7% of households which have an adult male present, the reported household head is a woman. For those households, we use the shock of her husband or the oldest working-age male in the household. Our approach implicitly assumes that child labor prevalence is persistent over time, even out-of-sample. That is, when looking at adults today, we assume that the districts that are high child labor today were also high child labor when those adults grew up (prior to the period for which we have data). Consistent with this assumption, the correlation of the share of children working across rounds is stable over time at about 0.5, with little decay for farther apart rounds. Idiosyncratic variation such as weather or economic shocks (Shah and Steinberg, 2017), as well as sampling bias, likely explain the time-variant portion of the variation.

of shocks on long-run consumption. To address this, we estimate the differential effects of ELR_{dta} on migration, exploiting a 2007 round of the NSS that asked households detailed questions about out-migrants.

Appendix Table A6 reports the average and differential effects of ELR_{dta} on an indicator variable coded as 1 for any male that the household reports has out-migrated to a separate district. One limitation of this measure is that it does not capture households where all household members have migrated out-of-district together. However, to the extent that parents are often left behind in rural India when young men migrate either seasonally or permanently for work (Chakraborti, 2004), we will capture much of the relevant migration. We focus on males because this is the relevant group for the long-run consumption regressions, which exploit variation in the male household head's shock. On average, ELR_{dta} has a very small and insignificant effect on the migration rate, and there is no evidence that ELR_{dta} has differential effects on migration in high child labor districts. Since the results suggest that the scope for bias is small, we conclude that it is unlikely to be driving our consumption results.

7 Robustness to Alternative Explanations

We now explore alternative explanations for our estimates of β_2. In the first subsection, we control for a variety of district-level covariates that may be associated with the prevalence of child labor and may also lead ELR_{dta} to have heterogeneous effects. In the second subsection, we consider whether saving, which we abstracted away from in our model, could be affecting our results. Finally, in the third subsection, we discuss whether differential returns to education across districts could drive the results.

7.1 Controlling for Differences Across Districts

Districts with high child labor prevalence likely differ from those with low child labor. For example, one might expect high child labor districts to be poorer, though this is not always the case in India, where cotton-growing regions like Gujarat are relatively wealthy and have high child labor. To understand the potential for omitted variables to be driving the results, in Appendix Table A7, we interact early life shocks with a variety of measures of local income, adult education, school quality, and socioeconomic factors.[13] In column 2, we control for

[13]Due to all of these interactions, it is difficult to interpret the coefficient for the (residual) direct effect of early life shocks, and so we omit it from the table, though it is included in the regression as a control. For all the district-level controls, for consistency with our child labor measure, we include when possible the leave-out share of rounds a district has above the median value. In cases where we only observe one round of data (e.g., DISE), we use an indicator for being above median. Using the mean values for these variables

measures of income: the average wage of adults, the share of adults and household heads who work for a wage, and the share of adults working in agriculture. In column 3, we control for measures of adult education: literacy and graduation rates. In column 4, we control for the available measures of school quality, described in Section A2. Column 5 controls for district level measures of the share of people from "scheduled" or "other-backwards" castes, the share who are Hindu, Muslim, or Christian, and a state-level measure of the share of people who practice purdah.[14] Column 6 includes measures of village-level development from ASER: indicators for if it has electricity, a tarred metal road leading to it, a post office, a pds ration shop, a bank, a government primary school, a government middle school, a government secondary school, and a private school. Because these questions were not asked in all ASER rounds, there is a substantial decline in the sample size. In column 7, we select from all of the above controls following the post-double selection lasso method (Urminsky et al., 2016); this machine learning procedure selects only the controls that are most predictive of the outcome and key endogenous variable, addressing potential concerns about the inclusion of too many interactions simultaneously. Across all specifications, the regression results for β_2 are nearly identical to the baseline results, shown in column 1.

Appendix Table A8 repeats this exercise for the IV strategy. In column 7, we use the same set of controls as chosen by the double lasso procedure in Table A7. While the estimate is less precise in column 7, consistent with the much smaller sample size, the point estimate of interest is insensitive to the inclusion of different sets of controls.

7.2 Household Saving and Investment

One concern with our primary empirical specification is that households who receive a positive shock in a child's critical period may save or invest that additional income in other ways than through the child's human capital.

To address this concern, first we note that the nature of our empirical strategy makes it unlikely that saving or capital investment could be driving our results. Our definition of a positive shock is rainfall above the 80th percentile in the district's historical distribution. That means that, by definition, the average household will experience three positive shocks, three negative shocks, and nine "normal" years in the first 15 years of a child's life. Thus, by the time we measure schooling or consumption of a school-aged child, their household has almost certainly experienced one or more positive and negative shock years at some point. Our empirical strategy relies on the exactly timing of those shocks relative to the

delivers very similar results (available on request).

[14] Practice of purdah is drawn from the IHDS, and the inclusion of this control reduces the sample since it is not available for every location.

birth of the child. We use the fact that shocks experienced between the in utero period and age 2 have a uniquely large and lasting impact on a child's human capital stock, relative to shocks experienced before or after. Thus, our regressions essentially compare children who experience a positive (or negative) shock during this critical period to those who experience the same shocks, but at, say, age 5 or 8, when they will have less impact on human capital.

Second, our primary specification includes district-year fixed effects, and compares children (or adults, in Table 6) from the same district and survey year of different birth years (and thus different early life shocks). Because shocks are measured at the district level, these households have, by definition, experienced the exact same history of positive and negative rainfall shocks. These fixed effects allow us to hold constant all district characteristics, including the exact sequence of droughts and positive rainfall shocks; identification only comes from the combination of the district and the child's year of birth.

Nonetheless, one might be concerned that household savings behavior responds differently when there is a child in the critical period relative to when there is not. To address this directly, we include household fixed effects in equation (1). Including household fixed effects means that the estimates are identified by the gap in the outcomes (measured at the same time) between two siblings who received different shocks in the same household. In this specification, we hold fixed the amount of capital, land or other assets held by the household, since our estimates are identified within, rather than across, households. Appendix Table A9 shows that the main results are robust to this more stringent specification, with the point estimates almost identical to those without the household fixed effects.

Lastly, we can directly test empirically whether households who have a child who experienced an early life shock appear to have increased saving/investment. First, Appendix Table A10 tests in the IHDS if early life shocks predict households' productive assets. We find no evidence that children who experienced different shocks early in life have different levels of household animals or assets. Second, if households are saving or investing in response to a child's early life shock, then we would expect the shock to affect all children in the household. Appendix Table A11 extends Table 3 to estimate how *siblings'* early life shocks affect dropout. We find that shocks to siblings do not affect dropout, either in general or in places with high child labor, consistent with our results being driven by individual-specific human capital, and not household-level changes. Finally, Appendix Table A12 extends Appendix Table A5 to estimate the effects of *toddlers'* early life shocks on current consumption. If early life shocks mainly affect children's households' consumption because children work, we should not see an effect since toddlers do not work. If, on the otherhand, our regressions pick up household savings behavior, we would expect to see a positive effect. Reassuringly, shocks to toddlers do not affect family consumption, consistent with our interpretation of our main results.

7.3 Returns to Education

Another possible concern is that the returns to education may differ across districts in India. We discuss the scope for differential returns to drive our results below.

First, we note that, in our model, differences in the returns to education will lead to differences in the educational response to early life rainfall specifically if increases in early life human capital have a larger positive effect on the the returns to education in low vs. high child labor areas. That is, the relevant question is not whether the returns to education are lower in levels in high child labor areas but whether the increase in the returns due to an increase in early life human capital is smaller. It is not entirely clear why this would be true, but even if it were, it would only attenuate the positive effects of early life rain in high child labor areas (that is, lead to $\beta_2 > 0$ when the outcome is dropout). To get a net negative effect on education in high child labor areas ($\beta_1 + \beta_2 > 0$ for dropout, as in Table 3), without the opportunity cost channel, increases in early life human capital/health would have to increase the returns to education in low child labor areas and *decrease* them in high child labor areas. We are not aware of any empirical or theoretical work consistent with the idea that early life health *decreases* the returns to educational investment. Put differently, even were the returns to education always *zero* in high child labor areas, this would lead positive early life shocks to have zero total effect on dropout rather than a positive net effect on dropout in high child labor districts.

Second, we note that our IV approach uses variation in the mix of crops grown, *conditional on overall agriculture share*, to instrument for child labor. That is, we are comparing districts with the same share of agriculture that grow relatively more cotton or flowers versus tree fruits or wheat. Our instrument picks up crops that have a specifically high return to *child labor* (typically because of differences in size and strength requirements), rather than *unskilled labor* more generally. Hence, the fact that our IV delivers if anything stronger results than the OLS reassures us that the results are not strongly biased by lower returns to education in high child labor districts (e.g., because of higher agricultural shares).

Finally, given the importance of this issue, we also empirically compare the relationships between consumption and education in high and low child labor districts. While this does not let us directly assess whether dynamic complementarities in the production function vary with child labor prevalence, if the relationship between education and consumption is similar across districts, this provides us with some reassurance that the functional form of $w_3^c(e, h)$ is the same in high and low child labor districts. We use the IHDS 2012 data to measure the effect of an additional year of schooling on log per capita consumption separately in districts that have above median shares of child labor more and less than 50% of the time.[15] An important caveat is that the observational Mincerian returns estimates may not be causal. Appendix Figure A3 shows nearly identical positive slopes on years of education in high and low child labor districts; the Mincerian returns to education appear to be similar regardless of child labor prevalence.

We can also use our estimated effects of the relationship between early life shocks, schooling, and later-life consumption to estimate our own measure of the return to schooling. Focusing on low child labor districts in column 3, Table 3 shows that one positive early life shock (relative to a neutral year) is associated with an increase of 0.07 (=0.0054×13) years of schooling, and Table 6 shows that it is associated with an increase in consumption of 0.59%, for an estimated return to schooling (plus the direct benefits of the early life shock) of 8%. In high child labor districts, early life investment is associated with 0.042 (=0.0032×13) fewer years of school and 0.47% less consumption, for a return to schooling (again gross of the direct benefits of the early life shock) of 11%. These results imply that the returns to schooling are similar in high and low child labor districts, and if anything might be higher in places with more child labor.[16] Reassuringly, these implied estimates of the return to schooling are in line with other estimates from low-income countries (Duflo, 2001; Patrinos and Psacharopoulos, 2020; Khanna, 2023).

8 Discussion: Are Parents Making Efficient Decisions?

We now consider whether households are inefficiently reducing children's human capital in high child labor districts in response to positive rainfall shocks. If this is the case, since ELR_{dta} strictly increases a child's early human capital, it may be that children and/or their parents are not making efficient decisions about the trade-off between schooling and work. In this section, we provide two tests for efficiency.

We first test whether the shocks have different effects on oldest sons, who are more likely to stay in the household as adults and care for parents in their old age, easing contracting frictions. If the outcomes of eldest sons in high child labor districts more closely resemble those of children in low child labor districts, this is evidence in favor of inefficient investment responses to shocks due to contracting frictions. Second, we back out the discount factor that

[15] We use the IHDS rather than the NSS because the NSS does not collect data on years of schooling.

[16] We can use the difference in the effects of an early life shock across high and low child labor districts in order to estimate the returns to schooling net of the direct benefits of early life shocks (assuming the direct benefits are the same across places). The difference in schooling is 0.11 years (=0.0086×13), and the difference in consumption is 1.1%, implying a return to schooling of 10%. This is well within the range of standard estimates, though the identification here is from children induced to drop out because of opportunity costs, instead of the more standard approach of using policy variation that encourages children to attend more school (Card, 2001; Duflo, 2001).

rationalizes the short-run increase in consumption with the long-run decrease in consumption that we observe in the data in high relative to low child labor places. Our estimate suggests that a quite low discount factor would be needed for behavior in high child labor areas to be efficient. We conclude that not only does child labor prevalence reduce and even reverse the educational benefits of early life shocks, but this reduction is likely welfare-reducing.

8.1 Oldest Sons

Motivated by Propositions 4a and 4b, we examine whether the interaction of CL_{dt} and ELR_{dta} has heterogeneous effects for eldest sons. From these propositions, we know that if parents are imperfectly altruistic, increased early life investment may inefficiently reduce educational investment. This is because parents will value the earnings from a child working, which they can expropriate today, more than the gains to a child's future income, from which they may not benefit. If a child could contract to share his future earnings with parents, parents would be more likely to make efficient educational decisions.

As Proposition 4b shows, if parents are imperfectly altruistic, we expect the effects of early life shocks to be more positive in high child labor places for children for whom intergenerational incomplete contracting problems are likely to be small (e.g. when the altruism/contracting parameter γ is sufficiently high). Cultural traditions where specific children provide parents with old age support are one informal mechanism to solve this incomplete contracting problem (Bau, 2021) and can generate variation in incomplete contracting problems across children in the same household. In India, oldest sons are expected to care for parents in their old age (Dyson and Moore, 1983; Gupta, 1987). Jayachandran and Pande (2017) provide evidence that this is associated with son preference and greater investment in oldest sons. Since incomplete contracting problems with oldest sons are likely to be smaller, we can test whether the impact of early rainfall shocks depends on the strength of incomplete contracting problems by examining how these shocks interact with birth order among boys.[17]

Estimating heterogeneous effects for eldest sons in ASER is complicated by the fact that ASER does not collect data on birth order – only age – and only collects information on children aged 5-16. As a result, it is unclear whether the oldest surveyed son is truly the oldest son in the household, as opposed to the oldest of the sons for whom data were collected. Before proceeding to the main analysis, to assess the degree to which this is

[17] In terms of the model, we interpret γ as higher for oldest sons relative to other children. If the educational investment in response to greater early life human capital investment is inefficiently low, then raising γ also raises educational investment. Thus, if we observe that early life rainfall shocks have different effects on oldest sons versus other children in high child labor districts, this provides evidence that parents are inefficiently reducing educational investment in response to rainfall shocks for the other children.

likely to lead to mis-assignment of eldest sons, we use the NSS migration survey to estimate the likelihood that the oldest of the male children of the head aged 5-16 in the household roster is actually the head's oldest son. The "No Migrants" column of Appendix Table A13 reports the probability considering family members who still live at home, and the "Including Migrants" also includes out-migrant sons.[18] Younger boys who are assigned to be eldest are the true eldest with a high probability, while by age 16, the probability a son is the true eldest is below 50% when accounting for migration. If the oldest son surveyed is a 13 year old boy, the odds that he is the true oldest son are around two-thirds. This makes intuitive sense. For a very young child's elder brother to not be in the household, the two children must have an extremely large age gap or the elder brother must have migrated at a very young age, both of which are unlikely. Motivated by the results in Appendix Table A13, in our estimates of the heterogeneous effects of birth order in ASER, we report estimates for households where the oldest child is younger than either 13 or 14, though the results are robust to other cut-offs.

We estimate the heterogeneous effects of ELR_{dta} and CL_{dt} on eldest sons using a triple-differences specification, where we include the triple-interaction term $ELR_{dta} \times CL_{dt} \times eldest\ son_i$ in regression equation (1) along with a separate control for the $eldest\ son_i$ indicator variable for a son being coded as the oldest son in the household. We also control for the relevant double interactions.

Table 7 reports the triple-interaction estimates, restricting the sample to either households whose oldest child is 13 or younger (columns 1–3) or 14 or younger (columns 4–6). The effect of early life shocks on increased dropout is entirely concentrated among younger sons and daughters. Across specifications, the triple interaction term of interest ranges from -0.004 to -0.006. An effect of this size almost fully undoes the increase in dropout due to a positive shock in high child labor places (β_2), leading shocks to have close to zero effect on the dropout of oldest sons. On the other hand, now that the effects on other children are no longer pooled with older sons, they are more than 20% larger.[19] These findings may also help explain the earlier result that girls' education falls more in response to the positive early life shock than boys' in high child labor areas, as daughters are not typically expected to care for parents in their old age.

Altogether, the estimates in Table 7 illustrate a few things. First, consistent with how much parents value oldest son's human capital, across all specifications we find that oldest

[18] The NSS migration survey collects data on out-migrants but does not collect information on their relationship to the household head. We infer that male migrants who are between 15 and 40 years younger than the head are the head's sons.

[19] Appendix Table A14 shows that the results are robust to alternative cut-offs for the age of the oldest son in the household of 11 and 12.

sons are relatively less likely to work in places with high child labor (by around half of the overall effect of child labor prevalence). Second, we see that the pattern of investment we have documented throughout this paper – the surprising fact that in districts with high child labor, a positive early life shock to human capital actually reduces schooling later on –does not hold for oldest sons. Because we expect both altruism and intergenerational contracting to be better for oldest sons than other children, this indicates that dropout decisions may be inefficiently sacrificing long-run income for the child to gain increased consumption for the natal household today. Lastly, these results also suggest that the dropout is not entirely due to parental misperceptions about the returns to schooling, since it is unlikely that parents would underestimate the returns to schooling only for non-oldest sons.

8.2 Discount Factor Calibration

In our model, parents trade off the consumption value of children's work today with the discounted value of the return to education through future earnings. A natural question is whether our results on childhood dropout can be rationalized with perfectly altruistic parents and standard discounting. If the implied discount factor is very low, it provides further evidence that declines in education in higher child labor areas, as well as failing to increase education as much in response to positive early life shocks, are welfare-reducing.

In high child labor places, early life shocks lead children to dropout, which raises their family's consumption (Appendix Table A5), but leads to less consumption when they are adults (Table 6). To calculate discount factors, we estimate the discount factor ρ that would exactly offset the long-run consumption losses with the short-run consumption gains in the absence of incomplete contracting ($\gamma = 1$). This exercise does require some additional assumptions. In particular, by comparing present discounted Rupee pay-offs, we are implicitly assuming utility is linear in consumption. Furthermore, we assume that families anticipate the actual growth rate of India over the period (real PPP per capita GDP increased 2.8% a year in the Penn World Tables from 1950 to 2010).[20] Appendix B provides further details of how we calibrate the discount factor.

The discount factor needed to rationalize our preferred IV estimates is 0.88. While not impossibly low, this number is well below estimates of the social discount factor, which is thought to be between 0.95 (in low-income countries) and 0.97 (in high-income countries) (Haacker et al., 2020). The numbers are also below the discount factor of 0.93-0.95 implied by India's interest rate during the sample period. As a result, given both our estimated

[20]The implied discount factors are mechanically a little higher assuming no growth, or lower if instead we assume more growth. This is because growth increases the returns to education in the future, making forgoing educational investment more costly.

discount factors and the fact that the oldest sons are not induced to leave school, we think that the dropout is unlikely to reflect efficient choices.[21]

9 Conclusion

Interventions that increase early childhood investment may be a powerful tool for increasing educational attainment and ultimately setting children on a better life trajectory. However, such policies can also have perverse effects in low-income countries, where child labor is common. We provide new evidence that early life investments increase child wages, increasing the attractiveness of child labor. Furthermore, we document the fact that while early life investments positively affect educational outcomes in places where child labor is low, consistent with the existence of dynamic complementarities, this effect is attenuated and even reversed in places where child labor is high. We also provide evidence that the divergence in educational outcomes in high child labor areas relative to low child labor areas is welfare-reducing.

Our results speak to the need for targeting and designing policies based on local conditions. Many governments provide supplemental early life nutrition to pregnant mothers and young children. Our estimates suggest that a program that increased consumption by around the same amount as one additional positive rainfall shock in early life would have markedly different effects on the later educational investments and the adult consumption of those children. In one year in our data, 2014, if all children in India received this boost to their early life human capital, it would lead to a total of 180,000 additional dropouts in districts with consistently above-median child labor. Yet, among Indian districts with consistently below median child labor, such a policy would have net positive effects, reducing total dropouts by 310,000 children.[22]

This does not imply that low-income countries should not pursue policies that promote early childhood investment, even in areas with high child labor. Rather, the design of these policies must take into account the role of opportunity costs and incomplete contracting between parents and children. For example, policies such as conditional cash transfers for educational investment can offset the opportunity cost effects of increased early childhood investment.

Finally, our results have important implications for researchers interested in identifying the parameters of the human capital production function. Researchers, particularly those

[21]Note that even if our discount factors correctly reflect household impatience, given the social discount factor, our results suggest that parents fail to undertake socially efficient educational investments from the perspective of the less-myopic social planner.

[22]This calculation uses census data for total population by age in 2014, and our estimates from Table 3, column 3 to calculate dropout rates.

working on low-income countries, must take into account how the child human capital stock affects the opportunity cost of schooling, as well as the benefits of schooling.

References

Adhvaryu, Achyuta, Anant Nyshadham, Teresa Molina, and Jorge Tamayo, "Helping Children Catch Up: Early Life Shocks and the PROGRESA Experiment," *Economic Journal*, Forthcoming.

_ **and** _ , "Endowments at Birth and Parents' Investments in Children," *Economic Journal*, June 2016, *126* (593), 781–820.

Agarwal, Bina, "Gender Relations and Food Security: Coping with Seasonality, Drought and Famine in South Asia," in Lourdes Benería and Shelley Feldman, eds., *Unequal Burden: Economic Crises, Persistent Poverty and Women's Work*, Oxford and Colorado: Westview Press, 1992, pp. 181–218.

Agostinelli, Francesco and Matthew Wiswall, "Estimating the Technology of Children's Skill Formation," *Journal of Political Economy*, Forthcoming.

Aizer, Anna and Flavio Cunha, "The Production of Human Capital: Endowments, Investments and Fertility," 2012. NBER Working Paper No. 18429.

Alderman, Harold, John Hoddinott, and Bill Kinsey, "Long term consequences of early childhood malnutrition," *Oxford economic papers*, 2006, *58* (3), 450–474.

Almond, Douglas and Janet Currie, "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, 2011, *25* (3), 153–72.

_ **,** _ **, and Valentina Duque**, "Childhood circumstances and adult outcomes: Act II," *Journal of Economic Literature*, 2018, *56* (4), 1360–1446.

Aloni, Tslil and Hadar Avivi, "One Land, Many Promises: The Unequal Consequences of Childhood Location for Natives and Immigrants in Israel," *Working Paper*, 2024.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena, "Bride Price and Female Education," *Journal of Political Economy*, 2020, *128* (2), 591–641.

Atkin, David, "Endogenous Skill Acquisition and Export Manufacturing in Mexico," *American Economic Review*, August 2016, *106* (8), 2046–2085.

Attanasio, Orazio, Costas Meghir, and Emily Nix, "Human capital development and parental investment in India," *The Review of Economic Studies*, 2020, *87* (6), 2511–2541.

Azam, Mehtabul and Geeta Gandhi Kingdon, "Are Girls the Fairer Sex in India? Revisiting Intra-Household Allocation of Education Expenditure," *World development*, 2013, *42* (C), 143–164.

Baland, Jean-Marie and James A. Robinson, "Is Child Labor Inefficient?," *Journal of Political Economy*, 2000, *108* (4), 663–679.

Banerjee, Abhijit V, "Educational Policy and the Economics of the Family," *Journal of Development Economics*, 2004, *74* (1), 3–32.

Basu, Kaushik, "Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards," *Journal of Economic literature*, 1999, *37* (3), 1083–1119.

_ **and Van Pham**, "The Economics of Child Labor," *American Economic Review*, 1998, *88* (3), 412–27.

Bau, Natalie, "Can policy change culture? Government pension plans and traditional kinship practices," *American Economic Review*, 2021, *111* (6), 1880–1917.

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti, "Why Should We Care About Child Labor? The Education, Labor Market, and Health Consequences of Child Labor," *Journal of Human Resources*, 2009, *44* (4), 871–889.

Behrman, Jere R, Mark R Rosenzweig, and Paul Taubman, "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment," *Journal of Political Economy*, 1994, *102* (6), 1131–1174.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen, "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain," *Econometrica*, 2012, *80* (6), 2369–2429.

Bharadwaj, Prashant, Juan Pedro Eberhard, and Christopher A. Neilson, "Health at Birth, Parental Investments, and Academic Outcomes," *Journal of Labor Economics*, 2018, *36* (2), 349–394.

_ , Leah K Lakdawala, and Nicholas Li, "Perverse Consequences of Well Intentioned Regulation: Evidence from India's Child Labor Ban," *Journal of the European Economic Association*, 2013.

Bleakley, Hoyt, "Disease and development: evidence from hookworm eradication in the American South," *Quarterly Journal of Economics*, 2007, *122* (1), 73.

_ , "Health, human capital, and development," *Annu. Rev. Econ.*, 2010, *2* (1), 283–310.

_ , "Malaria eradication in the Americas: A retrospective analysis of childhood exposure," *American Economic Journal: Applied Economics*, 2010, *2* (2), 1–45.

Card, David, "Estimating the return to schooling: Progress on some persistent econometric problems," *Econometrica*, 2001, *69* (5), 1127–1160.

Carpena, Fenella, "How do droughts impact household food consumption and nutritional intake? A study of rural India," *World Development*, 2019, *122*, 349–369.

Cascio, Elizabeth U and Ayushi Narayan, "Who needs a fracking education? The educational response to low-skill-biased technological change," *ILR Review*, 2022, *75* (1), 56–89.

Chakraborti, Rajagopal D, *The greying of India: Population ageing in the context of Asia*, Sage, 2004.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo, "Housing Booms and Busts, Labor Market Opportunities, and College Attendance," *American Economic Review*, 2018, *108* (10), 2947–2994.

Chetty, Raj and Nathaniel Hendren, "The impacts of neighborhoods on intergenerational mobility II: County-level estimates," *The Quarterly Journal of Economics*, 2018, *133* (3), 1163–1228.

Chyn, Eric and Lawrence F Katz, "Neighborhoods matter: Assessing the evidence for place effects," *Journal of Economic Perspectives*, 2021, *35* (4), 197–222.

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

_ **and** _ , "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 2008, *43* (4), 738–82.

Currie, Janet and Maya Rossin-Slater, "Early-Life Origins of Life-Cycle Well-Being: Research and Policy Implications," *Journal of Policy Analysis and Management*, 2015, *34* (1), 208–242.

_ **and Tom Vogl**, "Early-Life Health and Adult Circumstance in Developing Countries," *Annual Review of Economics*, 2013, *5* (1), 1–36.

Deaton, Angus, *The Analysis of Household Surveys: a Microeconometric Approach to Development Policy*, The World Bank, 1997.

Dessy, Sylvain E., "A Defense of Compulsive Measures Against Child Labor," *Journal of Development Economics*, 2000, *62* (1), 261–275.

Deutscher, Nathan, "Place, peers, and the teenage years: long-run neighborhood effects in Australia," *American Economic Journal: Applied Economics*, 2020, *12* (2), 220–249.

Dizon-Ross, Rebecca, "Parents' Beliefs About Their Children's Academic Ability: Implications for Educational Investments," *American Economic Review*, 2019, *109* (8), 2728–65.

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

Duque, Valentina, Maria Rosales-Rueda, and Fabio Sanchez, "How Do Early-Life Shocks Interact with Subsequent Human Capital Investments? Evidence from Administrative Data," 2023. Working Paper.

Dyson, Tim and Mick Moore, "On Kinship Structure, Female Autonomy, and Demographic Behavior in India," *Population and Development Review*, 1983, pp. 35–60.

Edmonds, Eric V, "Child labor," *Handbook of Development Economics*, 2007, *4*, 3607–3709.

Edmonds, Eric V. and Nina Pavcnik, "Child Labor in the Global Economy," *Journal of Economic Perspectives*, 2005, *19* (1), 199–220.

Epstein, T. Scarlett, *Economic Development and Social Change in South India*, Manchester: Manchester University Press, 1962.

García, Jorge Luis, James J Heckman, Duncan Ermini Leaf, and María José Prados, "Quantifying the life-cycle benefits of an influential early-childhood program," *Journal of Political Economy*, 2020, *128* (7), 2502–2541.

Gilraine, Mike, "School Accountability and the Dynamics of Human Capital Formation," *Working Paper*, 2018.

Goff, Leonard, Ofer Malamud, Cristian Pop-Eleches, and Miguel Urquiola, "Interactions Between Family and School Environments: Access to Abortion and Selective Schools," *Journal of Human Resources*, 2023.

Gupta, Monica Das, "Selective Discrimination Against Female Children in Rural Punjab, India," *Population and Development Review*, 1987, *13* (1), 77–100.

Haacker, Markus, Timothy B Hallett, and Rifat Atun, "On Discount Rates for Economic Evaluations in Global Health," *Health Policy and Planning*, 2020, *35* (1), 107–114.

Hazan, Moshe and Binyamin Berdugo, "Child Labor, Fertility, and Economic Growth," *Economic Journal*, 2002, *112* (482), 810–828.

Heckman, James J., "The Economics, Technology, and Neuroscience of Human Capability Formation," *Proceedings of the National Academy of Sciences*, 2007, *104* (33), 13250–13255.

Himaz, Rozana, *Is There a Boy Bias in Household Education Expenditure? The Case of Andhra Pradesh in India Based on Young Lives Data*, Young Lives, 2009.

Jayachandran, Seema, "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy*, 2006, *114* (3), 538–575.

_ , "The roots of gender inequality in developing countries," *Annual Reviews of Economics*, 2015, *7* (1), 63–88.

_ **and Rohini Pande**, "Why Are Indian Children So Short? The Role of Birth Order and Son Preference," *American Economic Review*, September 2017, *107* (9), 2600–2629.

Jensen, Robert, "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Johnson, Rucker C. and C. Kirabo Jackson, "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending," *American Economic Journal: Economic Policy*, 2019, *11* (4), 310–349.

Kaur, Supreet, "Nominal Wage Rigidity in Village Labor Markets," *American Economic Review*, October 2019, *109* (10), 3585–3616.

Khanna, Gaurav, "Large-scale education reform in general equilibrium: Regression discontinuity evidence from india," *Journal of Political Economy*, 2023, *131* (2), 549–591.

Knudsen, Eric I., James J. Heckman, Judy L. Cameron, and Jack P. Shonkoff, "Economic, Neurobiological and Behavioral Perspectives on Building America's Future Workforce," *Proceedings of the National Academy of Sciences*, 2006, *103* (27), 10155–10162.

Kone, Zovanga L, Maggie Y Liu, Aaditya Mattoo, Caglar Ozden, and Siddharth Sharma, "Internal Borders and Migration in India," *Journal of Economic Geography*, 2018, *18* (4), 729–759.

Krishna, Kowligi R, *Agroecosystems: Soils, Climate, Crops, Nutrient Dynamics and Productivity*, Apple Academic Press, 2014.

Kumar, Santosh, Ramona Molitor, and Sebastian Vollmer, "Drought and Early Child Health in Rural India," *Population and Development Review*, 2016, *42* (1), 53–68.

Laliberté, Jean-William, "Long-term contextual effects in education: Schools and neighborhoods," *American Economic Journal: Economic Policy*, 2021, *13* (2), 336–377.

Lancaster, Geoffrey, Pushkar Maitra, and Ranjan Ray, "Household Expenditure Patterns and Gender Bias: Evidence from Selected Indian States," *Oxford Development Studies*, 2008, *36* (2), 133–157.

Le, Kien and My Nguyen, "Droughts and child health in Bangladesh," *Plos one*, 2022, *17* (3), e0265617.

Levy, Victor, "Cropping Pattern, Mechanization, Child Labor, and Fertility Behavior in a Farming Economy: Rural Egypt," *Economic Development and Cultural Change*, 1985, *33* (4), 777–791.

Maccini, Sharon and Dean Yang, "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, June 2009, *99* (3), 1006–26.

Maluccio, John A, John Hoddinott, Jere R Behrman, Reynaldo Martorell, Agnes R Quisumbing, and Aryeh D Stein, "The impact of improving nutrition during early childhood on education among Guatemalan adults," *The Economic Journal*, 2009, *119* (537), 734–763.

Manski, Charles F, "Identification of endogenous social effects: The reflection problem," *The review of economic studies*, 1993, *60* (3), 531–542.

Merfeld, Joshua D and Jonathan Morduch, "Poverty at higher frequency," *KDI School of Pub Policy & Management Paper No. DP23-03*, 2023.

Munshi, Kaivan and Mark Rosenzweig, "Networks and Misallocation: Insurance, Migration, and the Rural-urban Wage Gap," *American Economic Review*, 2016, *106* (1), 46–98.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson, "The gift of moving: Intergenerational consequences of a mobility shock," *The Review of Economic Studies*, 2022, *89* (3), 1557–1592.

Ortiz-Ospina, Esteban and Max Roser, "Child Labor," *Our World in Data*, 2020. https://ourworldindata.org/child-labor.

Patrinos, Harry Anthony and George Psacharopoulos, "Returns to education in developing countries," in "The Economics of education," Elsevier, 2020, pp. 53–64.

Pritchett, Lant, *The Rebirth of Education: Schooling Ain't Learning*, CGD Books, 2013.

Ravallion, Martin and Quentin Wodon, "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy," *Economic Journal*, 2000, *110* (462), 158–175.

Rossin-Slater, Maya and Miriam Wüst, "What is the added value of preschool for poor children? Long-term and intergenerational impacts and interactions with an infant health intervention," *American Economic Journal: Applied Economics*, 2020, *12* (3), 255–286.

Santangelo, Gabriella, "Firms and Farms: The Local Effects of Farm Income on Firms' Demand," Cambridge Working Papers in Economics 1924, Faculty of Economics, University of

Cambridge January 2019.

Sayre, Rebecca K, Amanda E Devercelli, Michelle J Neuman, and Quentin Wodon, *Investing in early childhood development: review of the World Bank's recent experience*, The World Bank, 2015.

Schultz, Theodore W, "Capital formation by education," *Journal of Political Economy*, 1960, *68* (6), 571–583.

Shah, Manisha and Bryce Millett Steinberg, "Drought of Opportunities: Contemporaneous and Long Term Impacts of Rainfall Shocks on Human Capital," *Journal of Political Economy*, April 2017, *125* (2), 527–561.

_ **and** _ , "Workfare and human capital investment: Evidence from India," *Journal of Human Resources*, 2021, *56* (2), 380–405.

Singh, Abhijeet, Albert Park, and Stefan Dercon, "School Meals as a Safety Net: An Evaluation of the Midday Meal Scheme in India," *Economic Development and Cultural Change*, forthcoming.

Topalova, Petia, "Trade Liberalization, Poverty and Inequality: Evidence from Indian Districts," in "Globalization and Poverty," University of Chicago Press, 2007, pp. 291–336.

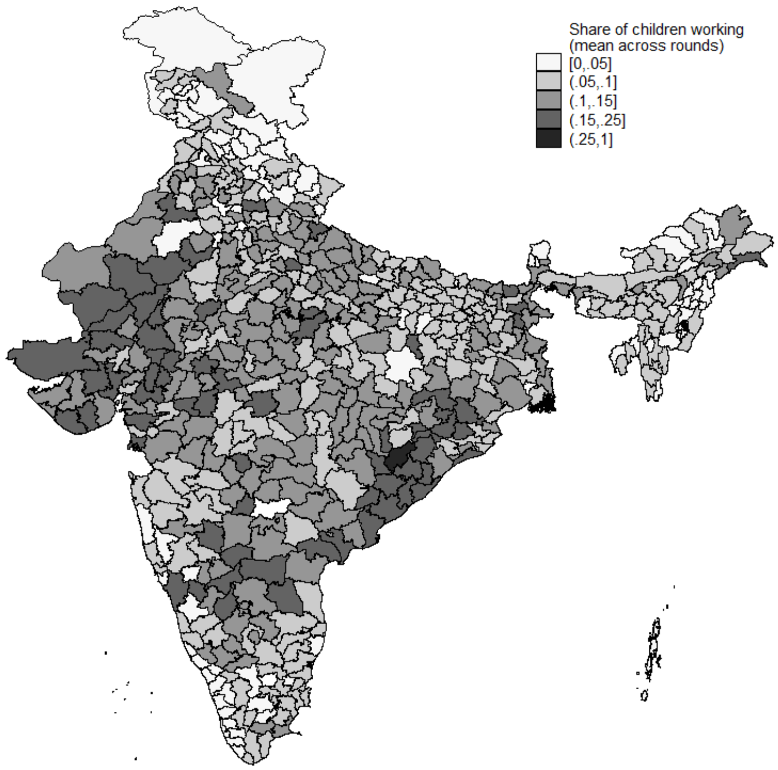
UNICEF, "The Situation of Children in India: A Profile," *New Delhi: UNICEF*, 2011.

Urminsky, Oleg, Christian Hansen, and Victor Chernozhukov, "Using double-lasso regression for principled variable selection," *Working Paper*, 2016.

Well, David N, "Accounting for the effect of health on economic growth," *The quarterly journal of economics*, 2007, *122* (3), 1265–1306.

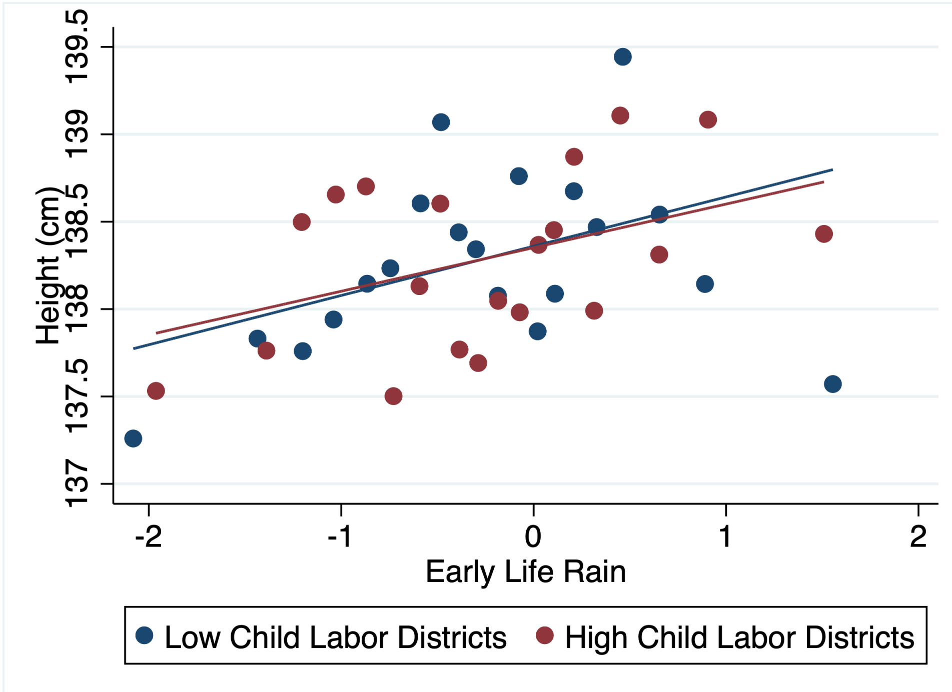
Figures

Figure 1: Share of Children Working by District



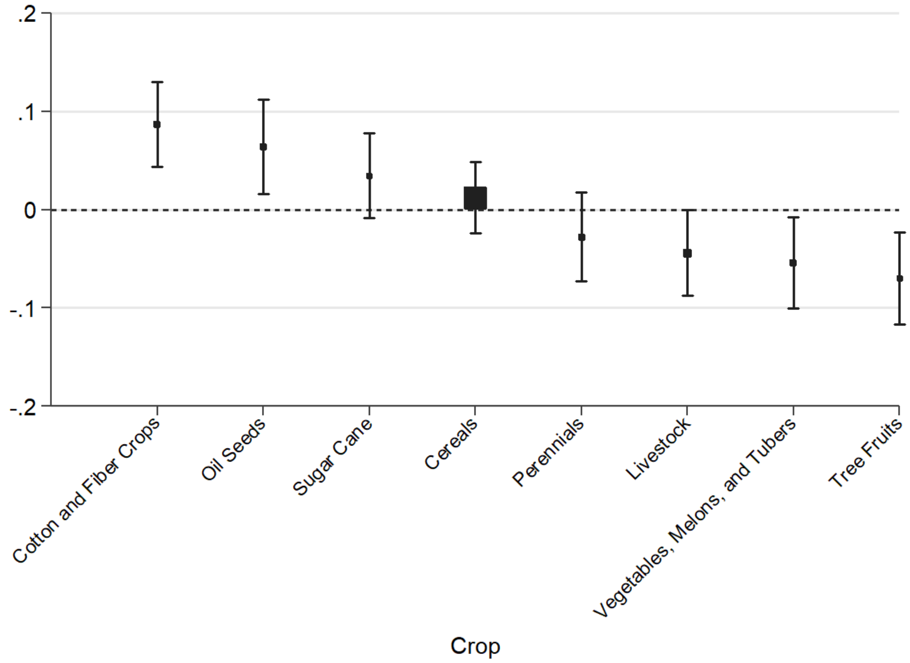
Notes: This figure shows a map of the districts of India, shaded by the prevalence of child labor, which is defined as the proportion of children aged 5-17 who report working in the market, in domestic work, or for a home enterprise as their primary activity. Source: NSS Schedule 10, 2004-2012.

Figure 2: Effect of Early Life Rainfall on Height by Child Labor Prevalence



Notes: This figure plots a binscatter of height (y-axis) on early life shocks (x-axis), separately by districts that have above median child labor shares less than 50% of the time (low child labor districts) and those with above median shares more than 50% of the time (high child labor districts), controlling for fixed effects for age and district. The coefficient estimate is 0.25 with a standard error of 0.11 (t = 2.26) in high child labor districts and 0.28 with a standard error of 0.14 (t = 1.98) in low child labor districts. Source: IHDS 2012.

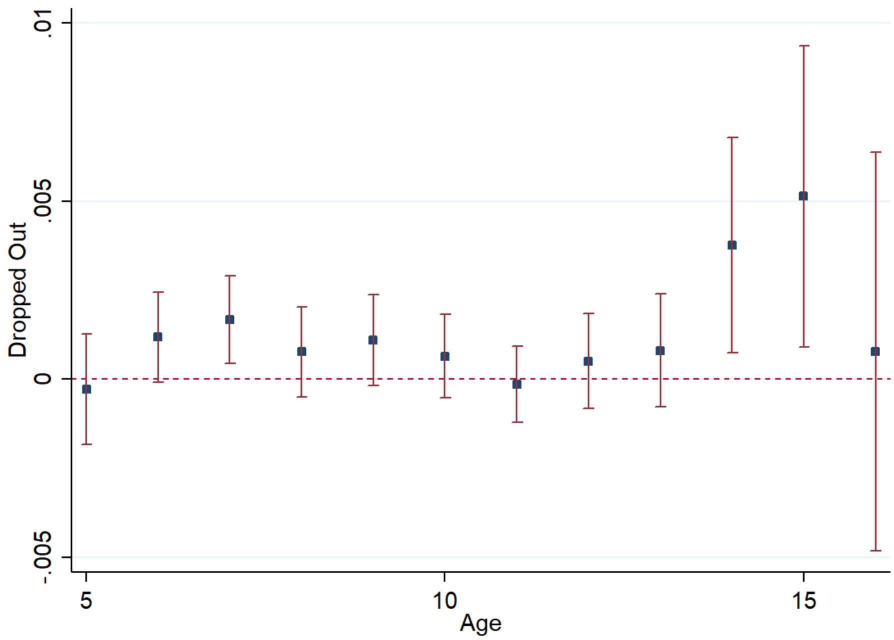
Figure 3: Association Between Crop Mix and Child Labor



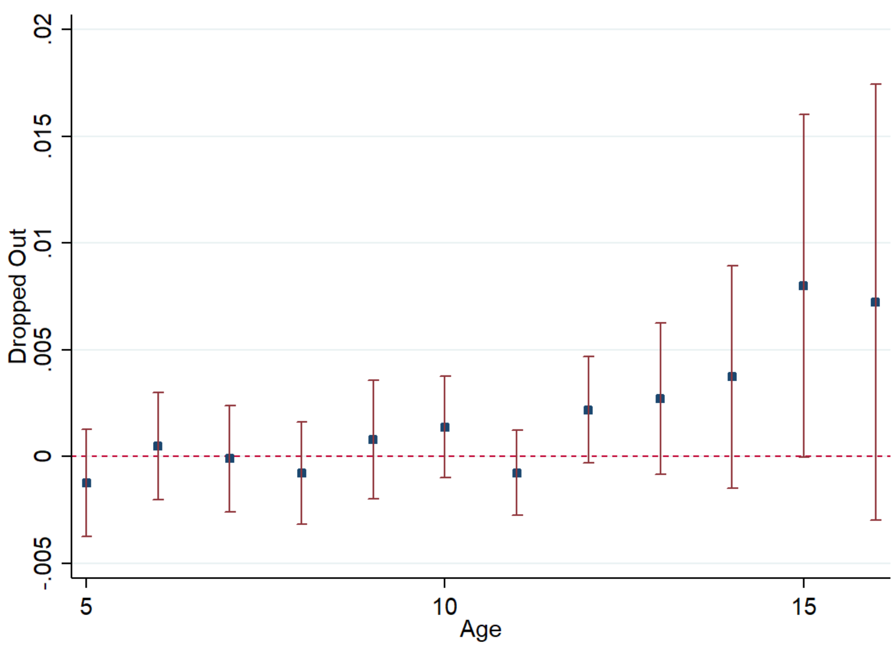
Notes: This figure plots the relationship between the share of adult agricultural employment in "large crops" (> 1 percent share of agricultural employment) and the main measure of child labor used in the paper (share of periods a district has above median child labor). All the coefficients are from a single regression. The size of the point estimates in the figure reflects the share of adult employment (in agriculture) for the crop. Source: NSS Schedule 10, 2004-2012.

Figure 4: Total Effect of Early Life Shocks on Dropout in High Child Labor Districts by Age

(**a**) OLS



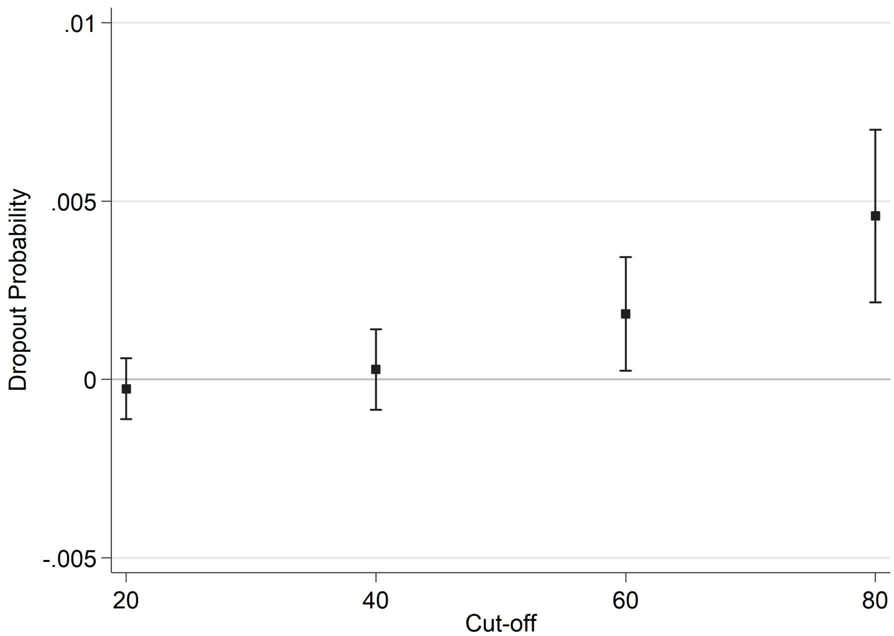
(**b**) LASSO IV



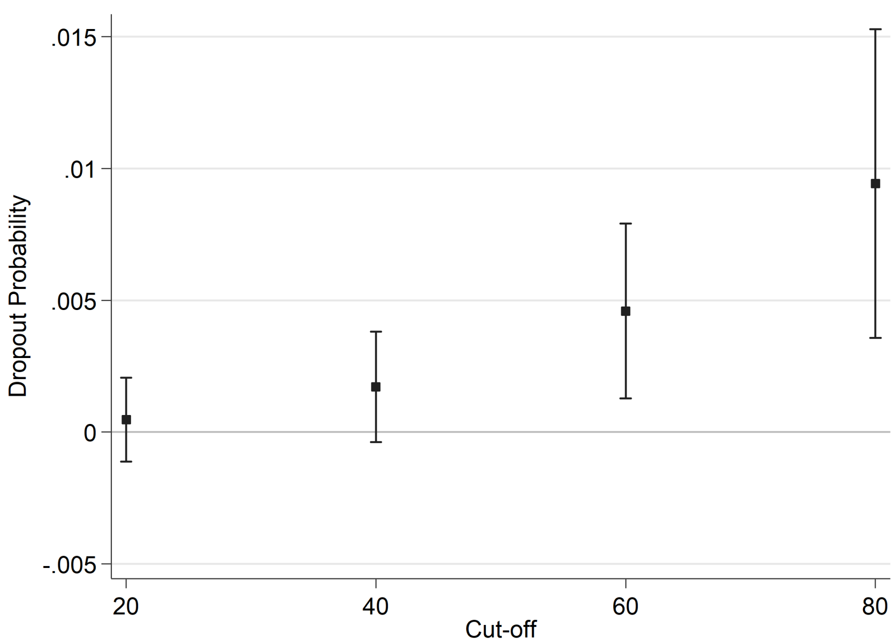
Notes: This figure reports estimates of the total effect of a 1 unit increase in ELR_{dta} by age in districts that always have above median child labor. The point estimates are dots, and bars show a 95% confidence interval. Panel (a) shows OLS results, and Panel (b) shows results using crops as an IV for child labor intensity. Standard errors are clustered by district. Source: ASER 2005-2014.

Figure 5: Total Effect of Early Life Shocks on Dropout in High Child Labor Districts, Varying the Cut-off Used to Generate CL_{dt}

(**a**) OLS



(**b**) LASSO IV



Notes: This figure varies the definition of CL_{dt} and re-estimates equation (1) for dropout using the ASER data. For each value of the x-axis, CL_{dt} is defined as the share of rounds of the NSS, leaving the current round out, that a district's child labor prevalence is above the cut-off denoted by the x-axis. The coefficients are calculated from different regressions, each of which controls for the baseline fixed effects (gender, age, and district by time period), and show the total estimated effects of early life shocks on dropout in districts above the cutoff. Panel (a) shows OLS results, and Panel (b) shows results using crops as an IV for child labor intensity. Standard errors are clustered by district. Source: ASER 2005-2014.

Tables

Table 1: Summary of Data Sources

Data Source	**Type**	**Years**	**Key Variables Used**
Annual Status of Education Report (ASER) - Household	Repeated Cross-Section	2005-2014	Dropout
Annual Status of Education Report (ASER) - Schools	Repeated Cross-Section	2005, 2007, 2009-2014	Attendance
National Sample Survey (NSS) Schedule 1	Repeated Cross-Section	2004-2012	Consumption Meals at School
National Sample Survey (NSS) Schedule 10	Repeated Cross-Section	2004-2012	Primary Activity Consumption Attends School Sector Within Agriculture
National Sample Survey (NSS) Schedule 10.2	Single Cross-Section	2007	Migration
India Human Development Survey (IHDS)	HH Panel	2005 & 2012	Wages Anthropometrics Math Scores Years of Schooling
Unified District Information System (DISE)	District Cross-Section	2005	Education Quality Measures
University of Delaware Gridded Rainfall Data	District Panel	1957-2014	Rainfall

Notes: This table reports the datasets (and the key variables) used in the analysis. More details on the specific rounds of the NSS used in the paper are reported in the text.

Table 2: Summary Statistics for Outcome and Explanatory Variables

	Mean	SD	N
ASER for Children 5 to 16, Household level			
Dropped Out	.035	.184	5,283,537
ASER Classroom Observations			
Attendance	21.3	19.3	640,915
NSS Schedule 1, Children 5 to 17			
Ate at Least 1 Meal in School	.244	.429	540,122
Ate at Least 20 Meals in School	.153	.36	540,122
Panel B: Secondary Outcomes and Controls			
NSS Schedule 10, Children 5 to 17			
Share Children Working as Primary Activity	.095	.294	486,295
Attends School	.817	.387	486,295
Works for Wage	.022	.147	486,295
IHDS for Children 5 to 17			
ln(wage)	2.53	.541	948
Any Wage	.037	.188	20,650
Height (cm)	138	18.2	22,007
NSS Schedule 1 and 10, Household level			
ln(Consumption per adult + 1/3 kids)	7.04	.592	544,629
ln(Consumption per adult + 1/2 kids)	6.95	.595	544,629
IHDS, Adults 24 to 55			
ln(Consumption per adult + 1/3 kids)	10.2	.631	37,553
ln(Consumption per adult + 1/2 kids)	10.1	.637	37,553
Years of Schooling	6.36	5.05	37,553
NSS Schedule 10, District Characteristics			
Share Adults in Agriculture	.613	.109	568
Share Adults in Manufacturing	.086	.054	568
Share Of Agriculture in Cereals	.698	.29	568
in Livestock	.065	.115	568
in Cotton and Fiber Crops	.041	.131	568
in Oil Seeds	.042	.101	568
in Vegetables, Melons, and Tubers	.04	.095	568
in Perennials	.034	.107	568
in Sugar Cane	.03	.105	568
in Tree Fruits	.024	.094	568

Notes: This table reports summary statistics for our main outcomes, explanatory variables, and key district characteristics.

Table 3: Effect of Early Life Shocks on Individual-Level Dropout and Classroom-Level Attendance

	Dropped Out (Individual)			Attendance (Classroom)		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV
Early Life Rain	-0.00063*	-0.0029***	-0.0054***	-0.12**	0.31***	0.35
	(0.00038)	(0.00071)	(0.0016)	(0.060)	(0.12)	(0.26)
Early Life Rain		0.0041***	0.0086***		-0.75***	-0.81*
× (Above Median) Child Labor		(0.0012)	(0.0029)		(0.16)	(0.43)
Mean Outcome	.035	.035	.035	21.3	21.3	21.3
Total Effect		0.0012*	0.0032**		-0.43***	-0.46**
SE of Total Effect		(0.0007)	(0.0014)		(0.08)	(0.19)
Kleibergan-Papp Robust F Stat			17			16.8
Number Districts	567	567	567	566	566	566
Number Observations	5283537	5283537	5283537	640915	640915	640915

Notes: This table reports estimates of the effects of early life shocks on schooling. District-level child labor classifications use the leave-out share of rounds a district has above median child labor. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for the adult share working in specific crops. The outcome for columns 1–3 is "dropped out" at the individual level, and the outcome for columns 4–6 is "number of kids in the classroom" with the early life shock calculated using the statutory age for the grade. The analysis includes all children between the ages of 5 and 16 in columns 1–3 and includes all surveyed grades in columns 4–6. Regressions include fixed effects for age, gender (only for columns 1–3), and district by time. Standard errors are clustered by district. Source: ASER 2005-2014.

Table 4: Effect of Early Life Shocks on Attending School & Meals at School

	Attends School			Any Meals			Over 20 Meals		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV
Early Life Rain	0.0044***	0.011***	0.011***	0.011***	0.021***	0.034***	0.0022*	0.0079***	0.014***
	(0.0012)	(0.0023)	(0.0041)	(0.0017)	(0.0032)	(0.0047)	(0.0012)	(0.0022)	(0.0037)
Early Life Rain		-0.014***	-0.013*		-0.019***	-0.046***		-0.011***	-0.024***
× (Above Median) Child Labor		(0.0034)	(0.0074)		(0.0048)	(0.0086)		(0.0036)	(0.0072)
Mean Outcome		.817	.817	.244	.244	.244	.153	.153	.153
Total Effect		-0.0026	-0.002		0.0012	-0.012***		-0.0035	-0.01**
SE of Total Effect		(0.0018)	(0.0036)		(0.0025)	(0.005)		(0.0021)	(0.0039)
Kleibergen-Papp Robust F Stat			17.1			16.5			16.5
Number Districts	571	568	568	568	568	568	568	568	568
Number Observations	486536	486295	486295	540122	540122	540122	540122	540122	540122

Notes: This table uses the NSS data to estimate the effect of early life shocks on different individual-level measures of school attendance. District child labor classifications use the leave-own round out share of rounds a district has above median child labor. In columns 3, 6, and 9, child labor prevalence is instrumented using a lasso-selected set of variables for the adult share working in specific crops. The outcome for columns 1–3 is equal to one if a child reports "attends school" as their primary activity, and the outcome for columns 4–9 is calculated from reported number of meals at school in the past month. The analysis includes all children between the ages of 5 and 17. Regressions include fixed effects for age, gender, and district by time. Standard errors are clustered by district. Data: NSS Schedules 1 and 10, 2004-2012.

Table 5: Effects with Average Share of Children Working as Measure of Child Labor

	Dropped Out (Individual)			Attendance (Classroom)		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV
Early Life Rain	-0.00063*	-0.0040***	-0.0067***	-0.12**	0.31**	0.30
	(0.00038)	(0.00097)	(0.0020)	(0.060)	(0.15)	(0.30)
Early Life Rain		0.031***	0.057***		-3.77***	-3.65
× (Mean) Child Labor		(0.0089)	(0.019)		(1.16)	(2.52)
Mean Outcome	.035	.035	.035	21.3	21.3	21.3
Total Effect		0.0012*	0.0027**		-0.31***	-0.31**
SE of Total Effect		(0.0007)	(0.0012)		(0.08)	(0.14)
Kleibergan-Papp Robust F Stat			21.2			20.3
Number Districts	567	567	567	566	566	566
Number Observations	5283537	5283537	5283537	640915	640915	640915

Notes: This table reports estimates of the effect of early life shocks on schooling. District-level child labor classifications use the average share of children who work in a district, leaving out the most recent round in the NSS. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for the adult share working in specific crops. The outcome for columns 1–3 is "dropped out" at the individual level, and the outcome for columns 4–6 is "number of kids in the classroom," with the early life shock calculated using the statutory age for the grade. The analysis includes all children between the ages of 5 and 16 in columns 1–3 and includes all surveyed grades in columns 4–6. For consistency, "Total Effect" reports the implied effect for a district that always is observed with above median child labor. Regressions include fixed effects for age, gender (only for columns 1-3), and district-by-time. Standard errors are clustered by district. Source: ASER 2005-2014.

Table 6: Effect of Early Life Shocks on Adult Consumption

	ln(Consumption Per Adult + 1/2 * kids)			ln(Consumption Per Adult + 1/3 * kids)		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV
Early Life Rain	0.0012**	0.0032***	0.0059***	0.0013**	0.0028***	0.0053***
	(0.00058)	(0.00099)	(0.0019)	(0.00061)	(0.0011)	(0.0019)
Early Life Rain		-0.0046***	-0.011***		-0.0033**	-0.0090**
× (Above Median) Child Labor		(0.0016)	(0.0038)		(0.0016)	(0.0039)
Mean Outcome	6.94	6.94	6.94	7.02	7.02	7.02
Total Effect		-0.0014	-0.0047**		-0.0005	-0.0037*
SE of Total Effect		(0.0009)	(0.0021)		(0.00096)	(0.0021)
Kleibergen-Papp Robust F Stat			29.7			29.7
Number Districts	568	568	568	568	568	568
Number Observations	544629	544629	544629	544629	544629	544629

Notes: This table reports the effect of early life shocks on adult consumption. District child labor classifications use the leave-own year out share of rounds a district has above median child labor. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for the adult share working in specific crops. Consumption is measured per capita, with children counting as 1/3 (columns 1–3) or 1/2 an adult (columns 4–6). Each household is in the data once, and a household's shock is coded as the male household head's shock. Household heads are either self-reported household heads, married to the reported household head (if the head is female), or if no member is coded as the household head, the oldest male under the age of 55. Regressions include controls for an indicator for being the oldest son and being the oldest son interacted with share of rounds a district has above median child labor, as well as fixed effects for age, gender, and district by time. Standard errors are clustered by district. Source: NSS Schedules 1 and 10, 2004-2012.

Table 7: Effect of Early Life Shocks on Dropout For Oldest Sons

	Dropped Out (Individual)					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV
Early Life Rain	0.00031	-0.0024***	-0.0038*	0.00031	-0.0023***	-0.0038**
	(0.00039)	(0.00065)	(0.0020)	(0.00038)	(0.00064)	(0.0019)
Early Life Rain	-0.00065**	0.0018***	0.0020	-0.0010***	0.0011***	0.0013
× Oldest Son	(0.00031)	(0.00041)	(0.0016)	(0.00029)	(0.00039)	(0.0015)
Oldest Son ×	-0.022***	-0.023***	-0.023***	-0.016***	-0.017***	-0.017***
(Above Median) Child Labor	(0.0012)	(0.0012)	(0.0014)	(0.0011)	(0.0012)	(0.0013)
Early Life Rain		0.0049***	0.0076**		0.0048***	0.0076**
× (Above Median) Child Labor		(0.0012)	(0.0037)		(0.0012)	(0.0035)
Early Life Rain		-0.0045***	-0.0050		-0.0039***	-0.0044
× (Above Median) Child Labor × Oldest Son		(0.00094)	(0.0032)		(0.00090)	(0.0030)
Mean Outcome (Not-Oldest Sons)	.03	.03	.03	.03	.03	.03
Mean Outcome (Oldest Sons)	.014	.014	.014	.021	.021	.021
Total Effect (Not-Oldest Sons)		0.0025***	0.0037**		0.0025***	0.0037**
SE of Total Effect		(0.0008)	(0.0018)		(0.0007)	(0.0017)
Total Effect (Oldest Sons)		-0.00021	0.00074		-0.00035	0.00067
SE of Total Effect		(0.00042)	(0.00076)		(0.00047)	(0.00083)
Kleibergan-Papp Robust F Stat (Not-Oldest Sons)			16.9			16.9
Kleibergan-Papp Robust F Stat (Oldest Sons)			17			17.1
Number Districts	567	567	567	567	567	567
Number Observations	3923446	3923446	3923446	4412448	4412448	4412448
Age Cutoff	13	13	13	14	14	14

Notes: This table reports the effect of early life shocks on dropout, allowing the effects to vary by child labor prevalence and whether a child is an oldest son. District child labor classifications use the leave-own year out share of rounds a district has above median child labor. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for the adult share working in specific crops. Regressions include fixed effects for age, gender, district by time, as well as an indicator for oldest son and its interaction with the measure of district child labor prevalence. Standard errors are clustered by district. Source: ASER 2005-2014.

Appendix A: Mathematical Appendix

A1 Additional Propositions

Proposition 3b shows that even if opportunity cost effects are not large enough to fully reverse the positive effects of early life human capital investment on education, they can still dampen those positive effects. To introduce Proposition 3b, we first note that for a given value of h, the parent will educate a child if $U_2^p(1,h) \geq U_2^p(0,h)$. Since $\frac{\partial U^p(h,1)}{\partial \alpha} > 0$ and $\frac{\partial U^p(h,0)}{\partial \alpha} = 0$, this relationship exhibits single-crossing. Thus, for any combination of h and d, there exists a cutoff value $\alpha_d^*(h)$ for α where $e = 1$ for all children with $\alpha \geq \alpha_d^*(h)$. Appendix Figure A4 illustrates this by plotting the ability distribution and showing that $e = 1$ if $\alpha > \alpha_d^*(h)$.

Proposition 3b. *If* $\frac{f(\alpha_{high}^*(h_{high}^*(y_1)))}{f(\alpha_{low}^*(h_{low}^*(y_1)))} < \Phi$, *then* $\frac{\partial \lambda_{high}(y_1)}{\partial y_1} < \frac{\partial \lambda_{low}(y_1)}{\partial y_1}$.

Proof. See subsection A2.

This proposition indicates that a positive income shock increases education (and adult wages) more in low child labor districts than high child labor districts, as long as an assumption is satisfied that the increased returns to child labor dominate two other, second order effects with ambiguous directions (the value of Φ is given below in Section A2). The effect we expect to dominate is that an increase in h increases the relative returns to education more in low child labor areas because, in high child labor areas, increasing h also increases the outside option, $w_{2,high}^c$. The additional ambiguous effects come from the fact that (1) the density of children on the margin of being educated is different in high and low child labor regions since enrollment rates are different, and (2) the derivative of adult wages with respect to early childhood investment may be different in high and low child labor regions if underlying investment in h is different in these regions. If underlying early life human capital investment rates are similar and the densities of the distribution at $\alpha_d^*(h_d(y_1))$ are similar across these regions, these additional, second order effects will be small.[23]

Appendix Figure A4 illustrates the intuition for Proposition 3b. In both high and low child labor districts, the increase in y_1 increases the relative returns to schooling, causing $\alpha_d^*(h_d^*))$ to fall. But α_{low}^* falls more than α_{high}^* because the relative returns to schooling increase more in low child labor districts. The share of children whose educational outcomes are changed is captured by the gray areas, which integrate over the ability distribution from the old to the new values of α_{low}^* and α_{high}^*. Even though the density at the cutoff is different

[23]The assumption that $\frac{f(\alpha_{high}^*(h(y_1)))}{f(\alpha_{low}^*(h(y_1)))} < \Phi$ bounds how much greater the density at α_{low}^* can be relatively to the density at α_{high}^*. That is, if the density at α_{high}^* is sufficiently high, it can lead the response to shocks to be greater in high child labor places even though the change in the ability cutoff is smaller.

in high and low child labor districts, as long as it is not too much greater in high child labor districts, more children will be affected in low child labor districts, where the integral is taken over a larger set of values of α.

A2 Proofs

Proof of Proposition 1.

Define $V = E\left[\max_e u(y_2 - c_e e + w^c_{2,d}(h)(1-e)) + \delta(U^c(w^c_3(e,h)) + \alpha e)\right]$, where the expectation is taken over realizations of α. Then, in period 1, the parent solves

$$\max_h u(y_1 - c_h h) + \rho V(h),$$

where ρ is the discount rate. From the first order condition, h^* must satisfy

$$-c_h u'(y_1 - c_h h^*) + \rho \frac{\partial V(h^*)}{\partial h} = 0,$$

To sign $\frac{\partial h^*}{\partial y_1}$, differentiate this expression with respect to y_1 and re-arrange to get

$$\frac{\partial h^*}{\partial y_1} = \frac{c_h u''(y_1 - c_h h^*)}{c_h^2 u''(y_1 - c_h h^*) + \rho \frac{\partial^2 V(h^*)}{\partial h}}.$$

To sign $\frac{\partial h^*}{\partial y_1}$, note that $c_h u''(y_1 - c_h h^*) < 0$ and $c_h^2 u''(y_1 - c_h h^*) < 0$ since $c_h > 0$ and $u'' < 0$. Then, the only term that remains to sign is $\frac{\partial^2 V(h^*)}{\partial h}$. To sign $\frac{\partial^2 V(h^*)}{\partial h^2}$, observe that

$$\begin{aligned}\frac{\partial^2 V(h^*)}{\partial h^2} =& E\Big[u''(y_2 - c_e e^* + w^c_{2,d}(h^*)(1-e^*))\Big(\frac{w^c_{2,d}(h^*)}{\partial h}\Big)^2(1-e^*) \\ &+ u'(y_2 - c_e e^* + w^c_{2,d}(h^*)(1-e^*))\frac{\partial^2 w^c_{2,d}(h^*)}{\partial h^2}(1-e^*) \\ &+ \delta\Big(U^{c\prime\prime}(w^c_3(h^*,e^*) + \alpha e^*)\Big(\frac{\partial w^c_3(e^*,h^*)}{\partial h}\Big)^2 + (U^{c\prime}(w^c_3(e^*,h^*) + \alpha e^*)\frac{\partial^2 w^c_3(e^*,h^*)}{\partial h^2}\Big)\Big],\end{aligned}$$

where e^* is the equilibrium choice of e. This expression is < 0 if $\frac{\partial^2 w^c_3(h)}{\partial h^2} \leq 0$ and $\frac{\partial^2 w^c_2(h)}{\partial h^2} \leq 0$. Therefore, $\frac{\partial h^*}{\partial y_1} > 0$.

Proof of Proposition 2. For a given h, a child drops out if $U^p_2(0,h) \geq U^p_2(1,h)$. Substituting in the values for consumption, this expression can be rewritten as

$$u(y_2 + w^c_{2,d}(h)) - u(y_2 - c_e) \geq \delta(U^c(w^c_3(h,1) + \alpha) - U^c(w^c_3(h,0))). \tag{2}$$

The derivative of the LHS with respect to y_1 is $\frac{\partial LHS}{\partial y_1} = u'(y_2 + w^c_2(h^*))\frac{\partial w^c_{2,d}(h^*)}{\partial h}\frac{\partial h^*}{\partial y_1}$, which is equal to 0 in low child labor places by assumption. The derivative of the RHS is $\frac{\partial RHS}{\partial y_1} = \delta\Big(U^{c\prime}(w_3^c(h^*,1)+\alpha)\frac{\partial w_3^c(h^*,1)}{\partial h}\frac{\partial h^*}{\partial y_1} - U^{c\prime}(w_3^c(h^*,0))\frac{\partial w_3^c(h^*,0)}{\partial h}\frac{\partial h^*}{\partial y_1}\Big)$. From diminishing marginal returns, $U^{c\prime}(w_3^c(h,1)+\alpha) < U^{c\prime}(w_3^c(h,0))$, so for the RHS to be increasing, we need that $\frac{\partial w_3^c(h,1)}{\partial h} > \frac{\partial w_3^c(h,0)}{\partial h}$. This expression implies that, for an early life shock to increase education rates in low child labor areas, there are dynamic complementarities between e and h.

Proof of Proposition 3a. Observe that $\lambda_d(h_d^*(y_1)) = 1 - F(\alpha_d^*(h_d^*(y_1)))$. Therefore, $\frac{\partial \lambda_d(h_d^*(y_1))}{\partial y_1} = -f(\alpha_d^*(h_d^*(y_1))\frac{\partial \alpha_d^*(h_d^*(y_1))}{\partial y_1}$. Therefore, $\frac{\partial \lambda_d(h_d^*(y_1))}{\partial y_1} = -f(\alpha_d^*(h_d^*(y_1))\frac{\partial \alpha_d^*(h_d^*(y_1))}{\partial y_1} \Rightarrow \frac{\partial \lambda_{high}(h_{high}^*(y_1))}{\partial y_1} = -f(\alpha_{high}^*(h_{high}^*(y_1))\frac{\partial \alpha_{high}^*(h_{high}^*(y_1))}{\partial y_1}$, where $f(\alpha_{high}^*) > 0$. To solve for $\frac{\partial \alpha_d^*(h_d^*(y_1))}{\partial y_1}$, note that $\alpha_d^*(h_d^*(y_1))$ is characterized by $U_2^p(0, h_d^*(y_1)) = U_2^p(1, h_d^*(y_1))$, which can be rewritten as

$$u(y_2 + w_{2,d}^c(h_d^*)) - u(y_2 - c_e) - \delta U^c(w_3^c(1, h_d^*) + \alpha_d^*) + \delta U^c(w_3^c(0, h_d^*)) = 0$$

Applying the implicit function theorem to this expression, we arrive at an expression for $\frac{\partial \alpha_d^*}{\partial y_1}$:

$$\frac{\partial \alpha_d^*}{\partial y_1} = -\frac{\partial w_3^c(1, h_d^*)}{\partial y_1} + \frac{u'(y_2 + w_{2,d}^c(h_d^*))\frac{\partial w_{2,d}^c(h_d^*)}{\partial y_1} + \delta\frac{\partial w_3^c(0,h_d^*)}{\partial y_1}U^{c\prime}(w_3^c(0, h_d^*))}{\delta U^{c\prime}(w_3^c(1, h_d^*) + \alpha_d^*)}.$$

Then,

$$\frac{\partial \alpha_{high}^*}{\partial y_1} = -\frac{\partial w_3^c(1, h_{high}^*)}{\partial y_1} + \frac{u'(y_2 + w_{2,high}^c(h_{high}^*))\frac{\partial w_{2,high}^c}{\partial y_1} + \delta\frac{\partial w_3^c(0,h_{high}^*)}{\partial y_1}u'(w_3^c(0, h_{high}^*))}{\delta U^{c\prime}(w_3^c(1, h_{high}^*) + \alpha_{high}^*)}.$$

Then, $\frac{\partial \lambda_{high}(h_{high}^*(y_1))}{\partial y_1} < 0$ if $\frac{\partial \alpha_{high}^*}{\partial y_1} > 0$. Rearranging $\frac{\partial \alpha_{high}^*}{\partial y_1} > 0$ shows that this satisfied if

$$\delta\left(\frac{\frac{\partial w_3^c(1,h_{high}^*)}{\partial y_1}U^{c\prime}(w_3^c(1, h_{high}^*) + \alpha_{high}^*) - \frac{\partial w_3^c(0,h_{high}^*)}{\partial y_1}U^{c\prime}(w_3^c(0, h_{high}^*))}{u'(y_2 + w_2^c(h_{high}^*))}\right)\left(\frac{\partial h}{\partial y_1}\right)^{-1} < \frac{\partial w_2^c(h_{high}^*)}{\partial h}.$$

Before proving Proposition 3b, we define Assumption A1.

Assumption A1.

$$\Phi > \frac{f(\alpha_{high}^*(h_{high}^*(y_1)))}{f(\alpha_{low}^*(h_{low}^*(y_1)))},$$

where

$$\Phi = \frac{\frac{\partial w_3^c(1,h^*_{low})}{\partial y_1} - \frac{\frac{\partial w_3^c(0,h^*_{low})}{\partial y_1} U^{c\prime}(w_3^c(0,h^*_{low}))}{U^{c\prime}(w_3^c(1,h^*_{low})+\alpha^*_{low})}}{\frac{\partial w_3^c(1,h^*_{high})}{\partial y_1} - \frac{u'(y_2+w^c_{2,high}(h^*_{high}))\frac{\partial w^c_{2,high}(h^*_{high})}{\partial y_1}+\delta\frac{\partial w_3^c(0,h^*_{high})}{\partial y_1}U^{c\prime}(w_3^c(h^*_{high},0))}{\delta U^{c\prime}(w_3^c(1,h^*_{high})+\alpha^*_{high})}}$$

Proof of Proposition 3b.
Recall that $\lambda_d(h_d^*(y_1)) = 1 - F(\alpha_d^*(h_d^*(y_1)))$. Therefore, $\frac{\partial \lambda_d(h_d^*(y_1))}{\partial y_1} = -f(\alpha_d^*(h_d^*(y_1))\frac{\partial \alpha_d^*(h_d^*(y_1))}{\partial y_1}$. Using the expression for $\frac{\partial \alpha_d^*(h_d^*(y_1))}{\partial y_1}$ from the proof of Proposition 3a and substituting this expression into $\frac{\partial \lambda_d(h^*(y_1))}{\partial y_1} = -f(\alpha_d^*(h_d^*(y_1))\frac{\partial \alpha_d^*(h^*(y_1))}{\partial y_1}$, we find that

$$\frac{\partial \lambda_{low}(h^*_{low}(y_1))}{\partial y_1} = \left(\frac{\partial w_3^c(1,h^*_{low})}{\partial y_1} - \frac{\frac{\partial w_3^c(0,h^*_{low})}{\partial y_1}U^{c\prime}(w_3^c(h^*_{low},0))}{U^{c\prime}(w_3^c(1,h^*_{low})+\alpha^*_{low})}\right) f(\alpha^*_{low})$$

$$\frac{\partial \lambda_{high}(h^*_{high}(y_1))}{\partial y_1} = \left(\frac{\partial w_3^c(1,h^*_{high})}{\partial y_1} - \frac{u'(y_2+w^c_{2,high}(h^*_{high}))\frac{\partial w^c_{2,high}(h^*_{high})}{\partial y_1}+\delta\frac{\partial w_3^c(0,h^*_{high})}{\partial y_1}U^{c\prime}(w_3^c(h^*_{high},0))}{\delta U^{c\prime}(w_3^c(1,h^*_{high})+\alpha^*_{high})}\right) \times f(\alpha^*_{high}).$$

Thus, $\frac{\partial \lambda_{low}(h^*(y_1))}{\partial y_1} > \frac{\partial \lambda_{high}(h^*(y_1))}{\partial y_1}$ under Assumption A1. To provide intuition for when Assumption A1 is satisfied, when h_d^* and α_d^* are sufficiently similar across the two types of districts, $\Phi > 1$. This is because the additional term in the denominator, $u'(y_2 + w^c_{2,high}(h^*_{high}))\frac{\partial w^c_{2,d}(h^*_{high})}{\partial y_1} > 0$, indicating that the denominator is smaller than the numerator. If α^*_{low} and α^*_{high} are sufficiently similar, $\frac{f(\alpha^*_{high}(h(y_1)))}{f(\alpha^*_{low}(h(y_1)))} \approx 1$ and Assumption A1 will be satisfied.

Proof of Proposition 4a. Returning to the proof of Proposition 2, an increase in y_1 will cause child labor to increase if the derivative of the LHS of equation (2) is greater than the derivative of the RHS for the marginal child whose ability is $\alpha_d^*(h_d(y_1))$. This is true if

$$u'(y_2 + w^c_{2,d}(h))\frac{\partial w^c_{2,d}(h^*)}{\partial h}\frac{\partial h^*}{\partial y_1} \geq \delta\Big(U^{c\prime}(w_3^c(h^*,1)+\alpha^*)\frac{\partial w_3^c(h,1)}{\partial h}\frac{\partial h^*}{\partial y_1} - U^{c\prime}(w_3^c(h,0))\frac{\partial w_3^c(h,0)}{\partial h}\frac{\partial h^*}{\partial y_1}\Big). \tag{3}$$

Substituting ρ for δ and $\tilde{w}_3^c$ for w_3^c, this is efficient if

$$u'(y_2 + w^c_{2,d}(h))\frac{\partial w^c_{2,d}(h^*)}{\partial h}\frac{\partial h^*}{\partial y_1} \geq \rho\Big(U^{c\prime}(\tilde{w}_3^c(h^*,1)+\alpha^*)\frac{\partial \tilde{w}_3^c(h,1)}{\partial h}\frac{\partial h^*}{\partial y_1} - U^{c\prime}(\tilde{w}_3^c(h,0))\frac{\partial \tilde{w}_3^c(h,0)}{\partial h}\frac{\partial h^*}{\partial y_1}\Big). \tag{4}$$

Now consider each of our two cases. If $\gamma < 1$ and $\frac{\partial w_3^c(h,1)}{\partial h} = \frac{\partial \tilde{w}_3^c(h,1)}{\partial h}$, $\rho > \delta$ and the RHS of equation (4) is greater than that of equation (3). This implies that there is a range of values over which equation (3) is satisfied while equation (4) is not and therefore, changes in educational investment are inefficient. If $\rho = \delta$, inefficiency will occur for a given h^* if the left-side of equation (4) is greater than the left-side of equation (3) (since the right sides of the equations are the same). With some algebra, we can see this will be the case if

$$\frac{\partial \tilde{w}^c(h^*,1)/\partial h}{\partial w^c(h^*,1)/\partial h} > \frac{U^{c\prime}(w_3^c(h^*,1)+\alpha^*)}{U^{c\prime}(\tilde{w}_3^c(h^*,1)+\alpha^*)}.$$

Thus, as long as this condition is satisfied, inefficiency will occur. This condition is intuitive: a larger increase in wages due to an increase in h pushes parents toward educating their children (left-side), but this is offset by the fact that the higher wage decreases the marginal value of more income (right side). That is, it is satisfied as long as the substitution effect dominates the income effect. If there is no diminishing marginal utility of consumption (utility is linear), this expression is always satisfied.

Proof of Proposition 4b. Note that $\frac{\partial \alpha^*}{\partial h} < 0$ if

$$u'(y_2 + w_2^c(h))\frac{\partial w_{2,d}^c(h^*)}{\partial h}\frac{\partial h^*}{\partial y_1} < \delta\Big(U^{c\prime}(w_3^c(h^*,1)+\alpha^*)\frac{\partial w_3^c(h,1)}{\partial h}\frac{\partial h^*}{\partial y_1} - U^{c\prime}(w_3^c(h,0))\frac{\partial w_3^c(h,0)}{\partial h}\frac{\partial h^*}{\partial y_1}\Big). \tag{5}$$

By assumption,

$$u'(y_2 + w_{2,d}^c(h))\frac{\partial w_{2,d}^c(h^*)}{\partial h}\frac{\partial h^*}{\partial y_1} \leq \rho\Big(U^{c\prime}(\tilde{w}_3^c(h^*,1)+\alpha^*)\frac{\partial \tilde{w}_3^c(h,1)}{\partial h}\frac{\partial h^*}{\partial y_1} - U^{c\prime}(\tilde{w}_3^c(h,0))\frac{\partial \tilde{w}_3^c(h,0)}{\partial h}\frac{\partial h^*}{\partial y_1}\Big),$$

so equation (5) is satisfied if $\gamma = 1$. Additionally, the RHS of equation (5) is strictly increasing in γ, while the LHS does not depend on γ. Thus, there is single-crossing in γ, indicating there exists a $\bar{\gamma}$ above which $\frac{\partial \alpha^*}{\partial h} < 0$.

Appendix B: Data Sources and Measurement

Annual Status of Education Report: Dropout and Attendance. Our primary measure of school dropout – our key outcome – comes from the Annual Status of Education Report (ASER), which surveys households about children's education from almost every rural district in India, including those who are out of school, from 2005–2014. Data are collected from approximately 500,000 children in each round, and children are surveyed at home in order to observe both those who have and have not dropped out. Starting in 2008, ASER also includes additional questions about the economic development of the village, which we include as controls in robustness checks.

Furthermore, ASER surveys all of the children 5–16 in a household. This allows us to conduct two separate sources of robustness analysis. First, we can include household fixed effects. Second, we calculate the early life shocks of siblings (excluding those within one year of age, who have overlapping shocks).

We supplement individual-level dropout measures with school-level attendance measures. ASER surveyors visit local schools and count the number of children in school that day. In order to show that our results are not driven by issues related to the measurement of self-reported dropout and to incorporate additional intensive margin variation in the attendance of enrolled children, we also use these classroom observations of the number of children in attendance at a school as a supplemental outcome.[24] We focus on attendance numbers rather than rates because our framework suggests that enrollment (the denominator in the attendance rate measure) will be affected by both early life shocks and their interaction with child labor.

National Sample Survey, Schedule 10: Child Labor and Consumption. We use the National Sample Survey (NSS) to create our district-level measure of child labor. The NSS is a repeated cross section of an average of 100,000 Indian households a year, conducted by the Indian government. We use rounds 60, 61, 62, 64, 66, and 68 (2004, 2004-5, 2005-6, 2007-8, 2009-10, 2011-12) in our analysis. These rounds roughly overlap with our ASER sample and have a relatively consistent set of district geographies. The Schedule 10 asks for the "primary" activity of each member in the household and includes categories for school attendance, wage labor, salaried work, domestic work, and so on. We define a child as "working" if her primary activity is any form of wage/salary labor, work with or without

[24]Since surveyors did not ask the age of every child in each class, we impute the early life shock using the statutory age-for-grade. While this aligns with the modal age-for-grade, the match is not perfect. For this reason, we consider the classroom-level effects a secondary outcome. We use school-level data from the same period as the individual-level data, with the exceptions of 2006 and 2008, since no school surveys were conducted during these years.

pay at a "home enterprise" (usually a farm, but the data also includes other small family businesses), or domestic chores.[25]

We use this primary activity variable to generate a district-level measure of child labor, our key explanatory variable. To create that measure for a round t, we calculate the share of survey years a district is above the median for share of children reporting working (leaving out own-survey year t). We leave-out own year since, when we estimate regressions using outcomes from the NSS, we want to avoid putting the same information on the left and right-side of the regression. When we merge the measure with the ASER data, we merge in the child labor prevalence measure for the nearest NSS round. This means we also leave-out the nearest round of the NSS when we use the measure with outcomes from ASER. While the ASER data include different children, and therefore, there is no risk of mechanically putting the same information on the left and right-hand side of a regression, the leave own-year out strategy still avoids a case where current child labor and a child's own outcomes are influenced by the same idiosyncratic time-varying district-level shocks.

Using an above/below-median cutoff for high child labor aligns with the model, where districts are either high or low child labor, and provides us a straightforward benchmark for calculating the total effect of a shock in a high child labor district. We operationalize this with the share of rounds a district is above median (which can be thought of as the probability of being a high child labor district) to help account for the fact that measurement error could lead a district to be just above or below the median cut-off in a given year. In practice, 17% of districts are never above median, while 19% are always above median, so these districts roughly correspond to the top and bottom quantiles of the child labor distribution. To ensure our results are not an artifact of this specific measure of child labor prevalence, we also estimate the main results using the leave one-out average share of children working across all rounds as the measure of child labor. Furthermore, we calculate additional measures of child labor using the share of rounds a district is above the 20th, 40th, 60th, or 80th percentiles to explore whether the effects of shocks change monotonically with the prevalence of child labor.

To corroborate the results generated from the ASER dropout measure, we generate "attends school" as an alternative measure of educational investment using the NSS data. This variable is generated when the reported primary activity of an individual is "attends school." We additionally generate an indicator if the child reports working for a wage.

We also use the NSS Schedule 10 data to measure the contemporaneous and long-run

[25] These categories comprise most of the non-schooling primary activities of children under 18, though there are other categories that are omitted, such as too young/infirm for work (typically the very old and very young), and "other," which includes begging and prostitution.

effects of shocks on consumption. The NSS Schedule 10 captures consumption by asking households to provide a consumption diary over the past week. We sum over expenditures in the different categories to arrive at a household consumption measure. Households with more children on average have lower per capita nominal consumption, both because children earn fewer resources and because they have lower demands. Since overall total consumption is therefore not a reliable measure of household welfare, we follow Deaton (1997) and calculate per capita consumption by weighting children as one-half or one-third of an adult.

Unlike ASER, the NSS provides a household roster which includes all of the children in a household, even babies. In robustness analysis, we calculate the early life shocks of toddlers (between the ages of 1 and 3), who are too young to work or go to school.

To control for potential differences between high and low child labor districts, we also use the NSS Schedule 10 to generate additional district- and household-level control variables. Our district-level controls from these data consist of the average wage for adult wage-earners in the district, the share of adults who work for wages, the share of male and female adults who are literate, the share who graduated primary school, and the share who have graduated from secondary school.

National Sample Survey, Schedule 10: Local Crop Mix Finally, we use the NSS Schedule 10 to create a pool of potential instruments for the prevalence of child labor. NSS round 68 asks respondents for their "principal industry" at a dissagregated level (we use 4 digit NIC 2008 codes). Round 68 allows us to measure much more variation than the other rounds (which, for instance, combine all food grain crops). Our pool of potential instruments is then the share of adults working in agriculture who report working in each disaggregated agricultural activity.

For each district, we calculate the share of adults (in agriculture) working in each agricultural sub-sector, merging together the very small subcategories. This leaves us with 37 categories. In order from most to least common, the categories are rice, wheat, other cereals, cattle and buffaloes, cotton (including merged categories), vegetables and melons, jowar/bajra/millets, sugarcane, mixed farming, soya bean oil seed, oleaginous fruits, beverage crops, support activities for crop production, other perennial crops, pulses, groundnut oil seed, spices and aromatic crops, pome fruits and stone fruits, other tree and bush fruits and nuts, mustard oil seed, tropical and subtropical fruits, sheep and goats, other oil seeds, other non-perennial crops, tobacco, poultry, grapes, sunflower oil seed, citrus fruits, other animals, plant propagation, post-harvest crop activities, pigs, support activities for animal production, hunting and trapping, and seed processing for propagation.[26]

[26]Slightly over half of adults work in either rice or wheat. There is fair amount of variation in agricultural

We then include all of the shares as potential instruments in the first stage. We follow the IV-lasso methodology proposed by Belloni et al. (2012) and use lasso to choose the set of instruments that best predict CL_{dt}. To maintain consistency, we select crops once using our main outcome, dropout in ASER (the first stage of the IV regression in column 3 of Table 3) and then use the same set of instruments throughout the analysis. The selected instruments are cotton, sugar cane, rice, cattle and buffaloes, sheep and goats, wheat, jowar/bajra/millets, other cereals, and other oil seeds. For all of our IV regressions, we include the share of adults working in these 9 sectors as instruments.

National Sample Survey, Schedule 1: Meals at School and Additional Measures of Consumption. The NSS Schedule 1 (Household Consumer Expenditure) surveys a (different) cross-section of households for the same rounds as Schedule 10 (60, 61, 62, 64, 66, and 68). This dataset gives us an additional source of information on total household consumption. In addition, the survey asks children how many meals they had at school in the past month. In general, it is difficult to measure intensive margin school attendance at scale since many census-style surveys only ask binary yes/no questions about enrollment, and time-use surveys tend to be small. Since many schools in India provide cooked meals (Singh et al., forthcoming), we use the number of meals consumed at school to measure if children go to school at all (any meals at school in the month) and if they go full time (over 20 meals).[27]

National Sample Survey, Migration Survey. In addition to the NSS Schedule 10 and 1, we also draw on a special migration survey conducted as part of the NSS in 2006-2007. Unlike other rounds of the NSS, this survey asked households about all members who had ever left the household, including information on those members' age. This allows us to create a dataset that includes both migrants and non-migrants and evaluate the effect of early life shocks on migration.

Yearly Gridded Rainfall: Variation in Human Capital. The data on rainfall shocks is from the University of Delaware Gridded Rainfall Data (version 5) for 1957-2008. Following the literature (Jayachandran, 2006; Shah and Steinberg, 2017), we define a "rainfall shock" as equal to one if rain is in the top 20th percentile for the district, -1 if it is in the bottom 20th percentile, and 0 otherwise.[28] We form an aggregate early life rainfall shock measure

sectoral shares across districts: for all crops, the standard deviation across districts is larger than the average.

[27]On aggregate, reported meals at school in the NSS is likely under-reporting total attendance, as the share who report any meals at school is lower than even the most pessimistic estimates of school attendance.

[28]In India, though flooding occasionally occurs in rural areas, more rain is essentially always better for crop yields. See Jayachandran (2006), Kaur (2019) and Santangelo (2019) for more discussion of the direct relationship between rainfall and crop yields.

denoted ELR by summing over the shocks when the child is in utero (age$= -1$), in her first year of life (year of birth), and in her second year of life. Thus, the aggregate shock variable ranges from -3 to $+3$. We match these data to individuals in all the other datasets using their birth year and district. We assume that people are born in their district of residence since cross-district migration in India is low (Topalova, 2007; Munshi and Rosenzweig, 2016), particularly for children (Kone et al., 2018). Furthermore, we verify that migration is not differential by ELR using the migration supplement to the NSS described above.

India Human Development Survey. We use the India Human Development Survey (IHDS), a panel dataset that was implemented in 2005 and 2012, for additional data on child wages and their correlation with measures of child human capital. The IHDS is a nationally representative, multi-topic panel survey whose rural section took place in 1503 villages across India. This survey measures child height, weight, and cognitive abilities, and these data allow us to provide evidence for the model's key assumption that children with higher early life human capital are more productive at child labor (as proxied by the wages they earn in the market). In the case of test scores, since cognitive skills may depreciate once children dropout and tests were only administered to children 8–11, we do so by estimating the association between test scores in Round 1 of the panel and wages in Round 2.[29] For consistency, we use second round outcomes in all our regressions since we use children's cognitive scores in the first round as controls in some specifications.

Unified District Information System for Education: Educational Quality. To obtain measures of educational quality at the district-level, we draw on the 2005 round of the Unified District Information System for Education (DISE), which was developed by India's National University for Educational Planning and Administration. We draw on this round to align the measures of school quality with the first year of ASER data (the dataset of our main outcome variable). These data allow us to observe the percent of schools with single classrooms and teachers, the percent with student-teacher ratios greater than 60, the percent of primary schools with boys and girls toilets, the percent with blackboards, the percent without buildings, and the average number of textbooks per school at the district-level, all of which we use as controls for school quality.

[29] All children aged 8–11 in the IHDS completed short writing, reading, and arithmetic tests as part of the survey regardless of whether they were in school.

Appendix C: Details of Discount Factor Calibration

We model the parent as choosing between the high and low child labor stream of consumption when a child is 5, in line with the sample we use to estimate the consumption benefits of early life shocks in high child labor places during childhood. The increase in consumption from an early life unit increase in aggregate rainfall in a high child labor place is given by

$$\sum_{t=0}^{13}(g\rho)^t \Delta c^h,$$

where ρ is the discount factor, g is the growth rate, and Δc^h is the change in consumption per capita for a household in a high child labor district relative to a low child labor district. In a low child labor district, the relative payoff from the rainfall shock occurs due to increased consumption in adulthood (starting at 18), which is represented by

$$\sum_{t=14}^{T}(g\rho)^t \Delta c^l,$$

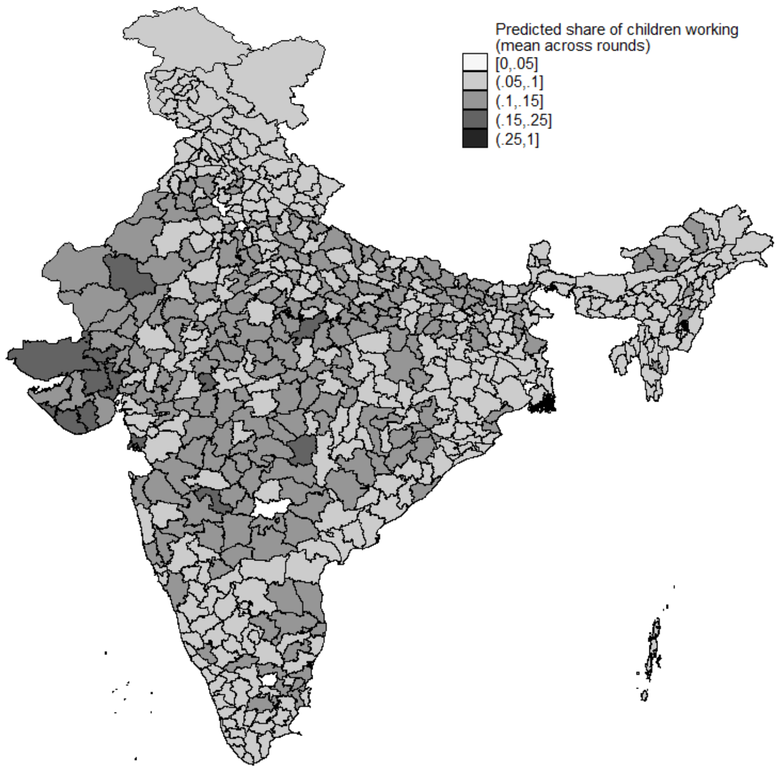
where Δc^l is the relative change in per capita consumption of the child in his adulthood in a low child labor district and T is the last year that the adult experiences consumption gains. We set $T = 65$ to be consistent with NSS measures of what ages individuals work, and $g = 1.028$ to match India's growth rate of 2.8% over our sample period.

The results in Table 6, which estimate the long-run effects of rainfall shocks on adult males' consumption, can be used to estimate Δc^l. The level value of Δc^l is just given by converting the log per capita effect of a unit increase in rainfall in a low relative to a high child labor district into a level effect using average consumption.

To calibrate Δc^h, we use estimates of the effect of rainfall shocks on per capita consumption by high and low child labor districts during the affected individual's childhood. The results of these regressions are reported in Appendix Table A5. Using these estimates, we calculate Δc^h the same way as we calculated Δc^l. With these estimates in hand, we can now solve for the maximum ρ for which $\sum_{t=0}^{13}(g\rho)^t \Delta c^h \geq \sum_{t=14}^{T}(g\rho)^t \Delta c^l$. Since geometric sums have a closed-form solution, setting the left and right side of this equation equal results in one equation with one unknown, which can be solved with Matlab as a non-linear optimization problem.

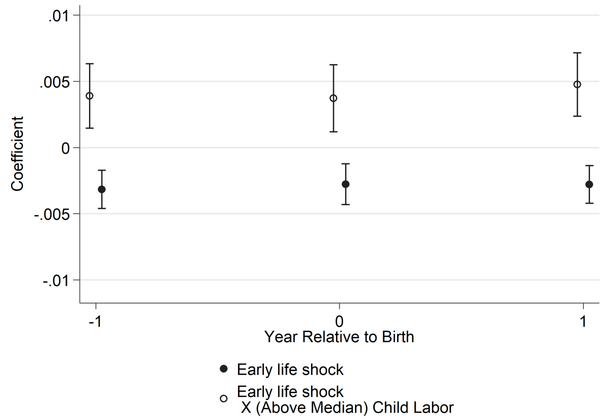
Appendix Figures

Figure A1: Predicted Share of Children Working by District (Crop IV)

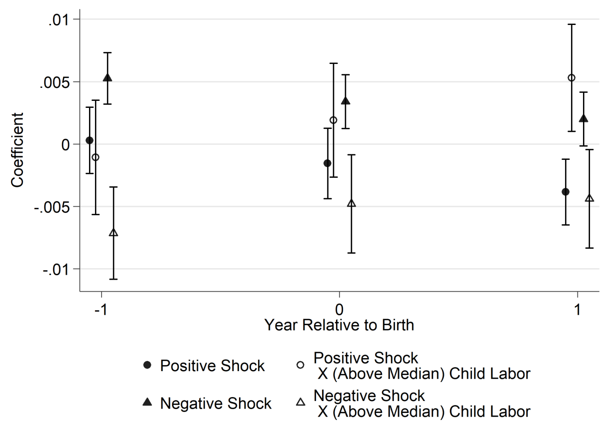


Notes: This figure shows a map of the districts of India, shaded by the prevalence of *predicted* share of children working (actual share of children working is reported in Figure 1). We predict child labor using the instruments selected by lasso IV from the pool of adult crop shares. We code a child aged 5-17 as working if she reports working in the market, in domestic work, or for a home enterprise as her primary activity. Source: NSS Schedule 10, 2004-2012.

Figure A2: Robustness of Effects on Dropout to Disaggregated Measures of Early Life Shocks



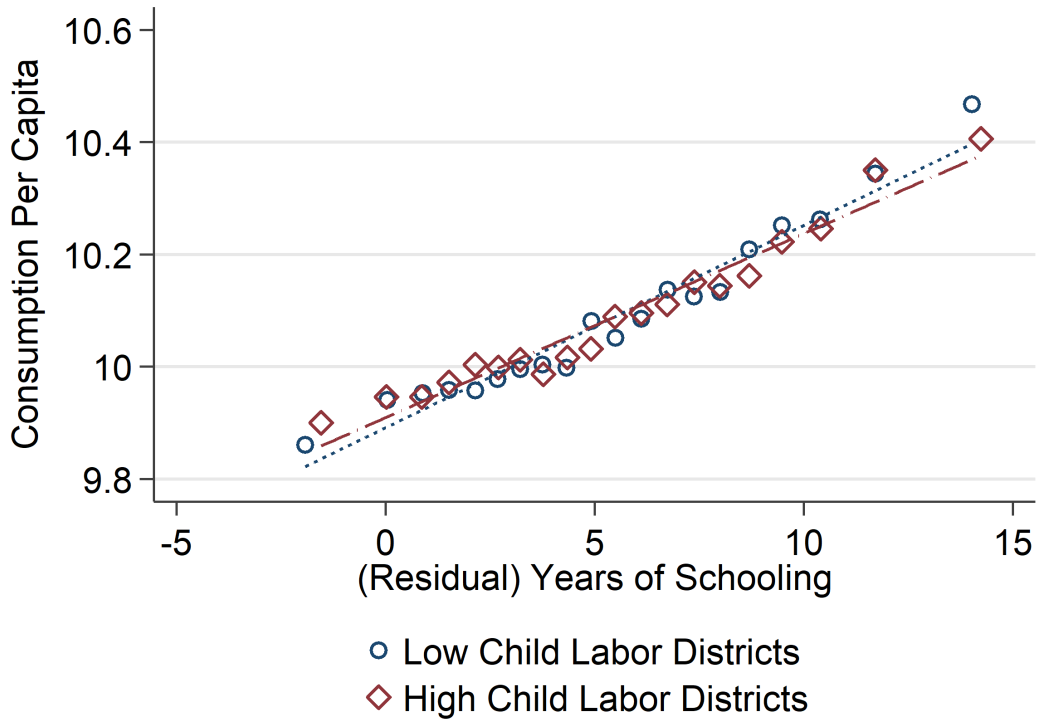
(**a**) Effects Disaggregated by Age



(**b**) Effects Disaggregated by Age and Direction of Shock

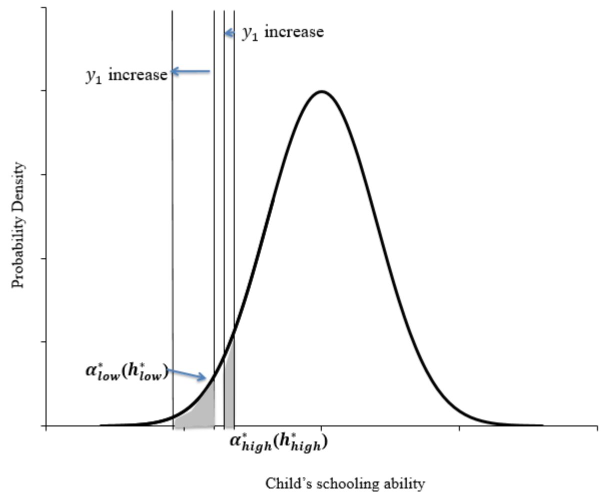
Notes: This figure reports the effect on dropout of early life shocks. District-level child labor classifications use the leave-out share of rounds a district has above median child labor. This figure shows coefficients from regressions similar to those in Table 3 Column 2, but additionally disaggregates early life rain by the timing of the shock and direction. Panel A shows the effect separately for the year before birth (-1), year of birth (0) and year after birth (1). Panel B shows the effects separately for positive and negative shocks. Both panels plot the coefficients from a regression showing (a) the direct effect of the shock and (b) its interacted effect with our measure of the location having above median child labor share. The analysis includes all children between the ages of 5 and 16. Regressions include fixed effects for age, gender, and district by time. Standard errors are clustered by district. Source: ASER 2005-2014.

Figure A3: Minceriam Consumption Returns to Education by Child Labor Prevalence



Notes: This figure reports the descriptive Mincerian relationship between years of schooling and log household consumption, testing for heterogeneity by child labor prevalence. Consumption is measured per capita, with children counting as 1/3 of an adult. The explanatory variable is the education (in years) of the household head, and relationships are estimated separately by districts that have above median child labor shares less than 50% of the time (low child labor districts) and those with above median shares more than 50% of the time (high child labor districts). The coefficient for high child labor districts is 0.035 with a standard error of 0.00099 (t=35.6), and the coefficient for low child labor districts is 0.038 with a standard error of 0.0012 (t=31.2). Source: IHDS 2012.

Figure A4: Illustration of Proposition 3b



Notes: This figure illustrates the intuition for Proposition 3b. $a^*_{low}(h^*_{low})$ denotes the cutoff exogenous returns to schooling above which a child is educated in a low child labor district for a given first period human capital investment h^*_{low}, and $a^*_{high}(h^*_{high})$ denotes the cutoff for high child labor districts. The graph illustrates how these cutoffs change as a function of shocks to first period income y_1. The gray shaded areas represent the children who were not educated before and become educated as a result of the change in y_1.

Table A1: Hedonic Predictors of Child Wages

	ln(wage)		
	(1)	(2)	(3)
Height (cm)	0.0060***		0.0053*
	(0.0021)		(0.0032)
Math Score		0.051**	0.034
		(0.022)	(0.026)
Mean Outcome	2.52	2.55	2.51
Number Districts	247	227	200
Number Observations	949	676	518

Notes: This table reports the descriptive relationship between height, cognitive skill, and ln(wage) conditional on working, controlling for age and gender fixed effects for children 5-17. Standard errors, clustered at the district level, are reported in parentheses. Wages and height are from the IHDS II (2012), while lagged test score data is from the IHDS I (2005). Height is measured in centimeters, while the math score is the number of math problems answered correctly.

Table A2: Effect of Early Life Shocks on Dropout By Gender

	Dropped Out (Individual)		
	(1)	(2)	(3)
	OLS	OLS	LASSO IV
Early Life Rain	-0.0017***	-0.0036***	-0.0073***
(Boys)	(0.00038)	(0.00070)	(0.0015)
Early Life Rain	0.00067	-0.0022***	-0.0034
(Girls)	(0.00045)	(0.00080)	(0.0022)
Early Life Rain		0.0034***	0.010***
× (Above Median) Child Labor (Boys)		(0.0011)	(0.0025)
Early Life Rain		0.0052***	0.0074*
× (Above Median) Child Labor (Girls)		(0.0014)	(0.0041)
Mean Outcome	.035	.035	.035
Total Effect (Boys)		-0.00021	0.0028**
SE of Total Effect (Boys)		(0.00063)	(0.0012)
Total Effect (Girls)		0.003***	0.004**
SE of Total Effect (Girls)		(0.0009)	(0.002)
Differential Effect Of Above Median Child Labor (Boys Minus Girls)		-0.0018**	0.0026
SE of Difference		(0.0009)	(0.003)
Kleibergen-Papp Robust F Stat (Boys)			17
Kleibergen-Papp Robust F Stat (Girls)			16.9
Number Districts	567	567	567
Number Observations	5283537	5283537	5283537

Notes: This table reports the effect of early life shocks on dropout, separately by gender. District child labor classifications use the leave-own-survey out share of rounds a district has above median child labor. In column 3, child labor prevalence is instrumented using a lasso-selected set of variables for adult crop share. The analysis includes all children between the ages of 5 and 16. Regressions include fixed effects for age by gender and district by time by gender. Standard errors are clustered by district. Source: ASER 2005-2014.

Table A3: Effect of Early Life Shocks, Instrumenting for Current Child Labor Prevalence

	Dropped Out (Individual)		
	(1)	(2)	(3)
	OLS	OLS	LASSO IV
Early Life Rain	-0.00060	-0.0037***	-0.0041***
	(0.00038)	(0.00095)	(0.0014)
Early Life Rain		0.0056***	0.0063**
× (Above Median) Child Labor		(0.0017)	(0.0025)
Mean Outcome			
Total Effect		0.0019**	0.0022*
SE of Total Effect		(0.0009)	(0.0012)
Kleibergan-Papp Robust F Stat		698	17.7
Number Districts	567	564	567
Number Observations	5276424	5265901	5276424

Notes: This table reports the effect of early life shocks on schooling, instrumenting for current child labor prevalence. In column 2, current child labor prevalence is instrumented using the leave-own-survey out share of rounds a district has above median child labor, as described in the text. In column 3, current child labor prevalence is instrumented using a lasso-selected set of crops, as described in the text. The outcome is "dropped out" at the individual level. The analysis includes all children between the ages of 5 and 16. Regressions include fixed effects for age, gender, and district by time. Standard errors are clustered by district. Source: ASER 2005-2014.

Table A4: Effect of Early Life Shocks on Working For a Wage

	Any Wage (NSS)			Any Wage (IHDS)		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV
Early Life Rain	-0.0026***	-0.0051***	-0.0071***	0.00025	-0.012***	-0.012
	(0.00043)	(0.00069)	(0.0016)	(0.0024)	(0.0033)	(0.0076)
Early Life Rain		0.0050***	0.0091***		0.021***	0.021*
× (Above Median) Child Labor		(0.0013)	(0.0033)		(0.0057)	(0.012)
Mean Outcome		.022	.022	.037	.037	.037
Total Effect		-0.000031	0.002		0.0089**	0.0089
SE of Total Effect		(0.00085)	(0.0018)		(0.0039)	(0.0058)
Kleibergan-Papp Robust F Stat			17.1			5.94
Number Districts	571	568	568	256	256	256
Number Observations	486536	486295	486295	20650	20650	20650

Notes: This table reports the effect of early life shocks on working for a wage. District child labor classifications use the leave-out share of rounds a district has above median child labor. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for adult share working in specific crops. The analysis includes all children between the ages of 5 and 17. Regressions include fixed effects for age, gender, and district by time. Standard errors are clustered by district. Source: NSS (2004-2012) for columns 1–3, and IHDS (2012) for columns 4–6.

Table A5: Effect of Early Life Shocks on Consumption During Childhood

	ln(Consumption Per Adult + 1/2 * kids)			ln(Consumption Per Adult + 1/3 * kids)		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV
Total HH Early Life Rain	-0.0047***	-0.0090***	-0.010***	-0.0046***	-0.0092***	-0.011***
	(0.00077)	(0.0015)	(0.0021)	(0.00078)	(0.0015)	(0.0021)
Total HH Early Life Rain $\times$ (Above Median) Child Labor		0.0093***	0.013***		0.0098***	0.013***
		(0.0022)	(0.0040)		(0.0023)	(0.0040)
Mean Outcome	6.94	6.94	6.94	7.02	7.02	7.02
Total Effect		0.00022	0.0022		0.0006	0.0026
SE of Total Effect		(0.0011)	(0.0021)		(0.0011)	(0.0021)
Kleibergen-Papp Robust F Stat			38.3			38.3
Number Districts	571	571	571	571	571	571
Number Observations	510435	510435	510435	510435	510435	510435

Notes: This table reports the effect of early life shocks of the children in the household on current consumption. District child labor classifications use the leave-out share of rounds a district has above median child labor. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for adult share working in specific crops. Consumption is measured per capita, with children counting as 1/2 (columns 1–3) or 1/3 of an adult (columns 4–6). Each household is in the data once, and a household's shock is coded as the total shock of all of children between the ages of 5 and 17. Regressions include fixed effects for the gender-by-age makeup of the children, district by time, and the early life shock of the household head. Standard errors are clustered by district. Source: NSS Schedules 1 and 10, 2004-2012.

Table A6: Effect of Early Life Rainfall on Male Migration by Child Labor Prevalence

	Migrated Away From District		
	(1)	(2)	(3)
	OLS	OLS	LASSO IV
Early Life Rain	-0.00019	0.0022	0.0070
	(0.0014)	(0.0021)	(0.0048)
Early Life Rain		-0.0055	-0.016
× (Above Median) Child Labor		(0.0036)	(0.011)
Mean Outcome	.218	.218	.218
Total Effect		-0.0033	-0.0093
SE of Total Effect		(0.0025)	(0.0062)
Kleibergan-Papp Robust F Stat			12.6
Number Districts	568	568	568
Number Observations	86548	86548	86548

Notes: This table reports the effects of early life rainfall on migration for men, controlling for fixed effects for age, gender, and district by time. District classifications use the leave-out share of rounds a district has above median child labor. Standard errors are clustered by district. The analysis includes all adults 25-54. Source: NSS 2007 migration supplement.

Table A7: Robustness of Interaction Between Child Labor Prevalence and Early Life Shocks

	Dropped Out						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Early Life Rain	0.0041***	0.0033**	0.0042***	0.0031***	0.0041***	0.0043***	0.0039**
× (Above Median) Child Labor	(0.0012)	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0015)	(0.0017)
Controls	Baseline	Income	Adult Education	School Quality	Caste and Religion	Village Characteristics	PDS Lasso
Mean Outcome	.035	.035	.035	.035	.035	.035	.035
Number Districts	567	564	554	557	492	565	486
Number Observations	5283537	5266455	5189644	5206726	4725486	3124335	2836947

Notes: This table reports the effect of the interaction between child labor and the early life shock on schooling, with additional controls relative to the baseline analysis. District child labor classifications use the leave-out share of rounds a district has above median child labor. All controls are interacted with the early life rainfall measure. Column 1 reports our baseline estimates from Table 3. Column 2 includes controls for local income, column 3 includes controls for the educational attainment of local adults, column 4 includes controls for local school quality, column 5 includes controls for local socioeconomic and religious characteristics, and column 6 controls for measures of village-level development. Column 7 selects from the full set of controls in the previous columns following the post-double selection lasso method (Urminsky et al., 2016). The specific variables are listed below, where the district-level controls are, for the variables we have in every round, the leave-out share of rounds a district has above the median value, and otherwise an indicator for being above median. The analysis includes all children between the ages of 5 and 16. Regressions include fixed effects for age by gender and district by time by gender. Standard errors are clustered by district. Source: ASER 2004-2012.

Income Controls (District-Level): the share of adults who work for a wage, their average wage, the share of household heads who work for a wage, and the share of adults who report working in agriculture.

Education Controls (District-Level): the average literacy of household heads, of women, and the overall graduation rate.

School Quality Controls (District-Level): schools per capita, the number of government and total primary schools, the number of schools with one classroom or one teacher, the share of schools with over 60 pupils per teacher, if the school has a separate toilet for girls, if the school has a blackboard, a building, or a textbook.

Caste and Religion (District-Level): share of the district in a Scheduled Caste, Scheduled Tribe or Other Backward Class, the share who are Hindu, the share who are Muslim, the share who are Christian, and the share who practice Purdah.

Village Characteristics (Village-Level): indicators for if the village has electricity, a tarred metal road leading to it, a post office, a pds ration shop, a bank, a government primary school, a government middle school, a government secondary school, and a private school.

Table A8: Robustness of the IV Estimates of the Interaction Between Child Labor Prevalence and Early Life Shocks

	Dropped Out						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Lasso IV	Lasso IV	Lasso IV	Lasso IV	Lasso IV	Lasso IV	Lasso IV
Early Life Rain	0.0086***	0.0093**	0.011***	0.0090**	0.0080**	0.0090***	0.0093
× (Above Median) Child Labor	(0.0029)	(0.0040)	(0.0033)	(0.0042)	(0.0031)	(0.0035)	(0.0057)
Controls	Baseline	Income	Adult Education	School Quality	Caste and Religion	Village Characteristics	PDS Lasso
Mean Outcome	.035	.035	.035	.035	.035	.035	.035
Kleibergan-Papp Robust F Stat	17	13.2	11.7	9.35	13.6	16.2	6.55
Number Districts	567	564	554	557	492	565	486
Number Observations	5283537	5266455	5189644	5206726	4725486	3124335	2836947

Notes: This table reports the effect of the interaction between child labor and the early life shock on schooling, instrumenting local child labor with agricultural production and including additional controls relative to the baseline analysis. District child labor classifications use the leave-out share of rounds a district has above median child labor. All controls are interacted with the early life rainfall measure. Column 1 reports our baseline estimates from Table 3. Column 2 includes controls for local income, column 3 includes controls for the educational attainment of local adults, column 4 includes controls for local school quality, column 5 includes controls for local socioeconomic and religious characteristics, and column 6 controls for measures of village-level development. Column 7 uses the covariates selected from the full set of controls in Table A7 following the post-double selection lasso method (Urminsky et al., 2016). The specific variables are listed below, where the district-level controls are, for the variables we have in every round, the leave-out share of rounds a district has above the median value, and otherwise an indicator for being above median. The analysis includes all children between the ages of 5 and 16. Regressions include fixed effects for age by gender and district by time by gender. Standard errors are clustered by district. Source: ASER 2004-2012.

Income Controls (District-Level): the share of adults who work for a wage, their average wage, the share of household heads who work for a wage, and the share of adults who report working in agriculture.

Education Controls (District-Level): the average literacy of household heads, of women, and the overall graduation rate.

School Quality Controls (District-Level): schools per capita, the number of government and total primary schools, the number of schools with one classroom or one teacher, the share of schools with over 60 pupils per teacher, if the school has a separate toilet for girls, if the school has a blackboard, a building, or a textbook.

Caste and Religion (District-Level: share of the district in a Scheduled Caste, Scheduled Tribe or Other Backward Class, the share who are Hindu, the share who are Muslim, the share who are Christian, and the share who practice Purdah.

Village Characteristics (Village-Level): indicators for if the village has electricity, a tarred metal road leading to it, a post office, a pds ration shop, a bank, a government primary school, a government middle school, a government secondary school, and a private school.

Table A9: Robustness of Effects on Dropout to Inclusion of Household Fixed Effects

	Dropped Out (Individual)		
	(1)	(2)	(3)
	OLS	OLS	LASSO IV
Early Life Rain	-0.00031	-0.0020***	-0.0049***
	(0.00036)	(0.00064)	(0.0014)
Early Life Rain		0.0031***	0.0083***
× (Above Median) Child Labor		(0.0011)	(0.0024)
Mean Outcome	.035	.035	.035
Total Effect		0.0011*	0.0034***
SE of Total Effect		(0.0007)	(0.0011)
Kleibergan-Papp Robust F Stat			16.7
Number Districts	567	567	567
Number Observations	4397457	4397457	4397457

Notes: This table reports the effect of early life shocks on schooling by child labor prevalence, controlling for household fixed effects. District child labor classifications use the leave-out share of rounds a district has above median child labor. In column 3, child labor prevalence is instrumented using a lasso-selected set of measures of adult employment share in different crops. The outcome is "dropped out" and is measured at the individual level. The analysis includes all children between the ages of 5 and 16. Regressions include controls for an indicator for being the oldest son and being the oldest son interacted with share of rounds a district has above median child labor, as well as fixed effects for age, gender, and district by time. Standard errors are clustered by district. Source: ASER 2005-2014.

Table A10: Effect of Early Life Rain on Household Assets

	Owns Animal			Household Assets		
	(1) OLS	(2) OLS	(3) LASSO IV	(4) OLS	(5) OLS	(6) LASSO IV
Early Life Rain	0.0055*	0.0091	0.0020	-0.014	-0.034	-0.13
	(0.0032)	(0.0058)	(0.014)	(0.033)	(0.065)	(0.14)
Early Life Rain		-0.0061	0.0061		0.034	0.20
× (Above Median) Child Labor		(0.0085)	(0.022)		(0.097)	(0.23)
Mean Outcome	.657	.657	.657	13.1	13.1	13.1
Total Effect		0.003	0.008		-0.00042	0.066
SE of Total Effect		(0.0049)	(0.0096)		(0.052)	(0.1)
Kleibergan-Papp Robust F Stat			5.4			5.4
Number Districts	256	256	256	256	256	256
Number Observations	20650	20650	20650	20643	20643	20643

Notes: This table reports the effect of early life shocks on household assets. District child labor classifications use the leave-out share of rounds a district has above median child labor. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for adult share working in specific crops. The analysis includes all children between the ages of 5 and 17. The outcome is if the household reports owning an animal (columns 1–3), or the number of assets out of a given list of 34, including bicycle, sewing machine, TV, etc. (columns 4–6). This variable ranges from 0-34. Regressions include fixed effects for age, gender, and district by time. Standard errors are clustered by district. Source: IHDS

Table A11: Effect of Siblings' Early Life Shocks on Dropout

	Dropped Out (Individual)		
	(1)	(2)	(3)
	OLS	OLS	LASSO IV
Sibling Early Life Rain	0.000038	-0.00012	0.0018
	(0.00014)	(0.00025)	(0.0018)
Sibling Early Life Rain × (Above Median) Child Labor		0.00029	-0.0033
		(0.00043)	(0.0034)
Mean Outcome	.035	.035	.035
Total Effect		0.00017	-0.0015
SE of Total Effect		(0.00026)	(0.0016)
Kleibergan-Papp Robust F Stat			6.58
Number Districts	567	567	567
Number Observations	4141219	4141219	4141219

Notes: This table reports the effect of siblings' early life shocks on schooling. District-level child labor classifications use the leave-out share of rounds a district has above median child labor. In column 3, child labor prevalence is instrumented using a lasso-selected set of variables for the adult share working in specific crops. The outcome "dropped out" is at the individual level. The analysis includes all children between the ages of 5 and 16. Sibling Early Life Rain is the total number of sibling early life shocks. We exclude from the calculation siblings with an age gap of a year or less, as they have overlapping shocks. Regressions include fixed effects for age and gender, the number of siblings, and district by time. The specification and sample is the same as that in Table 3, columns 1–3, except with three additional sibling-related covariates, and dropping children with no siblings. Standard errors are clustered by district. Source: ASER 2005-2014.

Table A12: Effect of Toddlers' Early Life Shocks on Consumption During Childhood

	ln(Consumption Per Adult + 1/3 * kids)			ln(Consumption Per Adult + 1/2 * kids)		
	(1) OLS	(2) OLS	(3) LASSO IV	(4) OLS	(5) OLS	(6) LASSO IV
HH Toddler Early Life Rain	0.0011	-0.00054	0.0031	0.0017	0.00030	0.0042
	(0.0013)	(0.0023)	(0.0043)	(0.0012)	(0.0022)	(0.0041)
HH Toddler Early Life Rain × (Above Median) Child Labor		0.0029	-0.0039		0.0025	-0.0047
		(0.0033)	(0.0075)		(0.0033)	(0.0072)
Mean Outcome	7.01	7.01	7.01	6.91	6.91	6.91
Total Effect		0.0024	-0.00073		0.0028	-0.00052
SE of Total Effect		(0.0019)	(0.0036)		(0.0019)	(0.0035)
Kleibergen-Papp Robust F Stat			25.1			25.1
Number Districts	571	571	571	571	571	571
Number Observations	397243	397243	397243	397243	397243	397243

Notes: This table reports the effect of toddlers' early life shocks on the household's current consumption. District child labor classifications use the leave-out share of rounds a district has above median child labor. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for adult share working in specific crops. Consumption is measured per capita, with children counting as 1/2 (columns 1–3) or 1/3 of an adult (columns 4–6). Each household is in the data once, and a household's shock is coded as the total shock of all of children between the ages of 1 and 3. Regressions include fixed effects for the gender-by-age makeup of the children and toddlers, district by time, and the early life shock of the household head. This is the exact same sample and specification as Appendix Table A5, expect with additional controls for the sum of toddler early life shocks and fixed effects for the gender-by-age makeup of the toddlers. Standard errors are clustered by district. Source: NSS Schedules 1 and 10, 2004-2012.

Table A13: Estimates of Likelihood the Assigned Oldest Son is the True Oldest Son in ASER

Age	No Migrants	Including Migrants
5	0.969	0.935
6	0.955	0.921
7	0.940	0.908
8	0.926	0.884
9	0.898	0.846
10	0.871	0.816
11	0.841	0.767
12	0.779	0.691
13	0.746	0.657
14	0.658	0.572
15	0.572	0.475
16	0.532	0.440

Notes: For each household in the NSS 2007, this table calculates the probability that the oldest son of the household head between the ages of 5 and 16 (the only children observed in ASER) is actually the oldest son of the head. The "No Migrants" column only includes children of the household head living in the household as sons of the head. This provides an upperbound measure of the probability that the assigned eldest is the true eldest since it ignores out-migration. The "Including Migrants" column includes male out-migrants as sons of the household head at the cost of inferring that a migrant is a son if he is 15-40 years younger than the head. Source: NSS 2007 migration supplement.

Table A14: Effect of Early Life Shocks on Dropout For Oldest Sons, Alternative Age Cut-offs

	Dropped Out (Individual)					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	LASSO IV	OLS	OLS	LASSO IV
Early Life Rain	0.00049	-0.0020***	-0.0029	0.00034	-0.0022***	-0.0035*
	(0.00039)	(0.00065)	(0.0018)	(0.00039)	(0.00065)	(0.0019)
Early Life Rain	-0.00015	0.0027***	0.0035**	-0.00037	0.0023***	0.0030*
× Oldest Son	(0.00031)	(0.00044)	(0.0014)	(0.00031)	(0.00043)	(0.0015)
Oldest Son ×	-0.029***	-0.030***	-0.030***	-0.025***	-0.026***	-0.027***
(Above Median) Child Labor	(0.0014)	(0.0014)	(0.0015)	(0.0012)	(0.0013)	(0.0014)
Early Life Rain		0.0045***	0.0062*		0.0047***	0.0070*
× (Above Median) Child Labor		(0.0012)	(0.0034)		(0.0012)	(0.0036)
Early Life Rain		-0.0052***	-0.0069**		-0.0051***	-0.0062**
× (Above Median) Child Labor × Oldest Son		(0.00091)	(0.0027)		(0.00094)	(0.0030)
Mean Outcome (Not-Oldest Sons)	.029	.029	.029	.03	.03	.03
Mean Outcome (Oldest Sons)	.007	.007	.007	.011	.011	.011
Total Effect (Not-Oldest Sons)		0.0025***	0.0033**		0.0025***	0.0035**
SE of Total Effect		(0.0007)	(0.0017)		(0.0008)	(0.0018)
Total Effect (Oldest Sons)		-0.000021	-0.00004		-0.00024	0.00025
SE of Total Effect		(0.00042)	(0.00072)		(0.00042)	(0.00077)
Kleibergan-Papp Robust F Stat (Not-Oldest Sons)			16.7			16.8
Kleibergan-Papp Robust F Stat (Oldest Sons)			16.5			16.8
Number Districts	567	567	567	567	567	567
Number Observations	2980773	2980773	2980773	3511408	3511408	3511408
Age Cutoff	11	11	11	12	12	12

Notes: This table reports the effect of early life shocks on dropout, allowing the effects to vary by child labor prevalence and whether a child is an oldest son. District child labor classifications use the leave-own year out share of rounds a district has above median child labor. In columns 3 and 6, child labor prevalence is instrumented using a lasso-selected set of variables for adult share for specific crops. Regressions include fixed effects for age by gender and district by time by gender, as well as an indicator for oldest son and the relevant double-interactions. Standard errors are clustered by district. Source: ASER 2005-2014.