

NBER WORKING PAPER SERIES

START WHAT YOU FINISH! *EX ANTE* RISK AND SCHOOLING INVESTMENTS
IN THE PRESENCE OF DYNAMIC COMPLEMENTARITIES

Andrew D. Foster
Esther Gehrke

Working Paper 24041
<http://www.nber.org/papers/w24041>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2017, Revised April 2022

The authors would like to thank the editor, James Heckman, three anonymous referees, Jishnu Das, Francisco Ferreira, Michael Grimm, Hanan Jacoby, Stephan Klasen, Holger Strulik, Sebastian Vollmer, and numerous seminar and conference participants for many helpful comments and suggestions. The REDS data used in this paper were collected by the National Council for Applied Economic Research (NCAER), India, in collaboration with Brown University, Yale University, and the World Bank. The SEPRI data were collected by the Institute for Rural Management, Anand (IRMA), India, in collaboration with the German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE) and the World Bank. Financial support by the Population Studies and Training Center at Brown University as well as the German Federal Ministry of Economic Cooperation and Development is gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2017 by Andrew D. Foster and Esther Gehrke. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Start What You Finish! Ex Ante Risk and Schooling Investments in the Presence of Dynamic Complementarities

Andrew D. Foster and Esther Gehrke

NBER Working Paper No. 24041

November 2017, Revised April 2022

JEL No. I25,O13,O15

ABSTRACT

We study the relationship between risk and schooling investment in a low-income setting. Our theoretical model reveals that parents respond to variance in future child schooling by reducing investments ex ante if the human capital production function exhibits dynamic complementarity and parental preferences for human capital are not too concave. We further explore the implications of the theory using multiple rounds of panel data from rural India that contain, in each round, three seasons of time allocation for each sampled child. In particular, we first use seasonally-measured study times to estimate the key parameters of the structural model. These estimates suggest that elimination of variance would result in a 15 percent increase in acquired human capital, 50% of which is attributable to an ex ante response. We then use cross-village variation in risk over time to estimate these effects, using the idea that the consequences of rainfall variation for income in a rural economy are importantly affected by the presence of irrigation. Using this variation, we estimate an ex ante elasticity of study time with respect to variance that is somewhat smaller, but of the same order of magnitude as those predicted by the structural model. Finally, we simulate the effects of an implicit social insurance program, modeled after the National Rural Employment Guarantee Scheme (NREGS). Our results suggest that the risk-reducing effect of the NREGS may offset adverse effects on child education that were evident during the NREGS phase-in due to rising wages.

Andrew D. Foster
Department of Economics and Community Health
Brown University
64 Waterman Street
Providence, RI 02912
and NBER
afoster@brown.edu

Esther Gehrke
Wageningen University and Research
Section Economics
Hollandseweg 1
6706KN Wageningen
Wageningen
Netherlands
esther.gehrke@wur.nl

1 Introduction

By the standards of most forms of household investment that have been studied in rural areas of low-income countries, the education of a household’s children has a very long time horizon. Successful completion of a particular credential involves a series of day to day decisions that accumulate and interact with each other over multiple years. These decisions, however, are complicated by the fact that rural life is inherently risky. Children at times may be withdrawn from school to support farm and household activities in the aftermath of adverse economic shocks. Given the complex nature of the learning process and the uncertainties of rural life, one might reasonably expect forward looking households to forgo such a long term project in favor of activities with less risky and more immediate returns.

This paper sets out to explore the role of uncertainties about future schooling in investments in current schooling. One specific mechanism that might drive such a relationship is the presence of dynamic complementarity in the production of human capital. Dynamic complementarity arises when schooling investments for a given child in different periods are complements in the production of human capital (Cunha and Heckman, 2007), and implies that early period investments in human capital have to be followed up by investments in later periods in order to make that early investment productive.

In the presence of dynamic complementarity, variance in the level of schooling investments over different time periods will affect the final stock of human capital via two distinct mechanisms. *First*, any mean-preserving increase in the variance in investments made at different points in time for a given child will reduce the final stock of human capital, as investments across periods are not perfectly substitutable. We call this the *ex post* effect of variance. *Second*, if future school investments are not known, households may be reluctant to invest in schooling in early periods, as they are uncertain about the ability to capture the returns to this investment at later stages. We call this the *ex ante* effect of variance.¹

In order to understand the *ex ante* response to risk that can be attributed to dynamic complementarities, we first develop a multi-period model of schooling investment and human capital accumulation, in which negative shocks to household income reduce schooling investments and *vice versa*, and parents anticipate variance in future incomes.² We analytically derive the *ex ante* effect

¹Note that we use the terms variance and risk interchangeably in this paper. As the literature usually refers to risk effects as *ex ante* effects, and we wish to highlight the difference between *ex ante* and *ex post* effects, we use the term variance where this distinction is particularly important.

²A reduction in study time in response to an economic shock might arise through a number of channels: Previous work has shown that children have to drop-out of school in order to work (Jacoby and Skoufias, 1997; Beegle et al., 2006; Duryea et al., 2007), and that households lack the financial resources to send children to school (see e.g. Jensen, 2000; Skoufias and Parker, 2006; Gubert and Robilliard, 2008; Bjorkman-Nyqvist, 2013). Finally, if adult household members increase their labor supply in response to adverse economic shocks (as has been shown by Kochar, 1999; Rose, 2001; Jayachandran, 2006) and home production takes up a substantial share in the time allocation of adults, particularly of women, any increase in parental work time is likely to increase the need for children to work at home (Skoufias, 1993; Ilahi, 2000; Shah and Steinberg, 2017). Note that any of these mechanisms is consistent with our characterization of schooling investment as market good in the theoretical model, which we impose for reasons of tractability.

of variance on schooling investments by restricting our attention to a two-period version of the model in which schooling investments in different periods are perfect complements and demand for human capital is linear. We then solve the two-period version of the model numerically to illustrate the *ex ante* effects of variance for more general assumptions about the elasticity of substitution in production and the curvature in the demand for human capital. We show that the *ex ante* effect of variance on school investments could be positive or negative, and that the sign of the effect depends on the presence of dynamic complementarity in the demand for child human capital, *i.e.* whether households invest more in schooling when the initial stock of human capital of a given child is exogenously higher.

For our empirical analysis we leverage the three rounds of the Rural Economic and Demographic Survey (REDS) collected in rural India between 1981 and 2008. These data are unique in that they include detailed information on time allocation of children and of mothers for three seasons of the year in each round. This allows us to construct the time a child allocates to studying (in school and at home) per day, which is arguably a key schooling investment in contexts in which children regularly support their families in household production. The REDS data are representative of rural India in 1967 and span multiple decades, allowing us to use the long-term variability of rainfall interacted with irrigation as a source of variation in risk within villages over time.

Our empirical strategy proceeds in two steps. First, to better understand the theoretical implications of dynamic complementarity for an *ex ante* response to risk, we structurally estimate a model of lifetime human capital accumulation using indirect inference (Gourieroux et al., 1993). In this analysis we allow for heterogeneity across households in the utility return to child human capital, for incomes to be correlated in the three seasons of the year, and for precautionary savings within the year but not across schools years. The analysis is identified using within village variation over over the course of the year and makes primary use of the child time-use data. The parameters of the auxiliary model are: the coefficients from a quadratic regression of third-period study time on first and second period study time and the interaction thereof, the coefficients from a quadratic regression of second-period study time on first period study time, the first two moments of study time in each season, and the correlation coefficients in household income from agriculture and labor between the three seasons. The auxiliary model parameters from the actual data match the parameters from the simulated data well.

Our estimated structural parameters point to substantial dynamic complementarity in the demand for child human capital, and imply that an increase in variance should reduce schooling investments due to the presence of dynamic complementarities. We use the estimated model parameters to simulate the lifetime effect of variance on child schooling, and decompose this effect into an *ex ante* and an *ex post* effect. We find that the *ex ante* effect explains a large share of the simulated response to risk. The elasticity of final stock of human capital with respect to variance is -0.15, 53-54% of which can be attributed to the *ex ante* response. We find an elasticity of school investments with respect to variance of -0.15 to -0.18.

We then test the predictions of our model at the estimated parameter values by exploiting

heterogeneity across villages in consumption variability, which we use as measure of risk. For this purpose we combine the REDS data with monthly precipitation data (ERA5 reanalysis data) obtained from the European Center for Medium-Range Weather Forecasts (ECMWF). We start by replicating results from previous studies on the role of rainfall shocks on agricultural incomes, consumption, and — given the focus of this paper — on the time use of children. We then explicitly test the empirical fit of different functional forms in the relationship between precipitation and consumption, and show that — at least in our data — consumption can be best explained by log rainfall interacted with the village level average of area under irrigation.

We use this specification to predict village-level risk. Specifically, we estimate the relationship between household consumption per capita, rainfall, village-level agricultural area under irrigation, and the interaction of these two variables. Using these estimates, we predict consumption outcomes at the village level for each observed rainfall outcome, given the current share of area that is irrigated. We then use the historical rainfall distribution to obtain the probability distribution of rainfall outcomes, and then calculate the probability distribution of predicted consumption for each village and round. Risk, finally, is defined as the variability in the distribution of predicted consumption. This concept of risk exploits the fact that labor markets in rural India are localized and largely dominated by agriculture over the time period we study. Hence, rainfall shocks not only affect farmers' consumption through on-farm production, they also affect consumption levels of landless households who mostly engage in casual agricultural employment.

We use variation in this variable within villages over time to estimate the effect of risk on study time, and find that risk significantly reduces the probability that children attend school. These results are extremely robust to various specifications, including specifications that control for household wealth and income, and state-specific shocks.

An important challenge in the reduced form estimation is to isolate the *ex ante* effect of variance from any *ex post* effects. Due to dynamic complementarity, past variation in study time will generally lead to lower levels of human capital than would be the case with lower variation but the same mean. But an *ex ante* response would do the same, as parents choose to reduce investments early on in order to avoid investing in child human capital without being able to capture the returns to these investments later. Conditioning on the stock of child human capital would thus underestimate the *ex ante* response to risk at any point in time. We isolate the *ex ante* effect by conditioning on the recent history of rainfall shocks, which arguably controls for the cumulative *ex post* effect of past shocks, but not for differences in the stock of child human capital that accrued from the household's *ex ante* responses in earlier grades. We also interact the history of the rainfall shocks with a dummy that equals one if the child was of school-age in that particular year, thereby allowing the effects of past shocks to be more pronounced for school-age children. Our results are robust to these controls, suggesting that we are indeed observing an *ex ante* effect of variance. Our findings imply that a 100% increase in our risk variable would reduce the probability that a child attends school by 4-5%, which is smaller than but of the same order of magnitude as those predicted from the structural model.

We also perform a battery of robustness checks. We show that our results are not driven by underlying differences in household or village characteristics, in the availability of schools, or in the history of weather shocks, nor driven by differential time trends. Finally we explore the potential for alternative explanations being the driver of our results. An important challenge for our analysis is that precautionary savings might generate another kind of *ex ante* relationship between variance and schooling investment. We explore this avenue, but cannot find any evidence that our results are driven by a savings motive. Our results also do not seem to be driven by differences in the returns to human capital, nor by differential fertility or gender-imbalances.

In order to assess the scope for public policy, we simulate the effects of an income-smoothing policy, modeled after the National Rural Employment Guarantee Scheme (NREGS) in India, on human capital investments. We estimate the extent to which NREGS reduces variability in consumption and use this reduction in risk to simulate the program's effect on child school attendance. We find that a similar program, that held the level of wages fixed, would increase school attendance by 1 percentage point.

The wage effects of such programs should not be ignored, however. Shah and Steinberg (2021) and Li and Sekhri (2020) find negative effects of the NREGS on schooling using the rolled phase-in of NREGS as a source of variation, and argue that the NREGS increased the opportunity costs of time of adolescents and therefore led them to drop out of school at younger ages. Arguably a risk mitigation effect could not be fully internalized by households at the early stages of the program when implementation was spotty and long-term viability was unclear. It is unclear how the direct effect of rising wages and the indirect effect of less variable incomes balance out in the longer term. However, our results suggest the negative wage consequences for schooling of the NREGS might be mitigated to the extent that the program is recognized as a reliable source of support during periods of adverse shocks.

By combining the structural estimation of key model parameters with a more reduced-form analysis of the effect of risk on child schooling, this paper is placed at the intersection of two important strands in the economics literature. We contribute to an exiting literature that explores the determinants of human capital investments in low and middle income countries (see *e.g.* Foster and Rosenzweig, 1996; Glewwe and Jacoby, 2004; Jensen, 2010, 2012; Oster and Steinberg, 2013; Atkin, 2016; Shah and Steinberg, 2017). Most closely related to our work are the papers by Fitzsimons (2007); Kazianga (2012) and Colmer (2021a). These papers explore the implications of weather variability on schooling in Indonesia, Burkina Faso and Ethiopia, respectively, and focus on precautionary savings motives in explaining the negative effect of risk on schooling (Fitzsimons, 2007; Kazianga, 2012), and the desire to diversify out of agriculture in explaining the positive effect of risk on schooling (Colmer, 2021a). Our paper, in contrast, highlights the existence of dynamic complementarities in the production of human capital, and investigates household investment decisions in response to these. There is other work that focuses on how income realizations affect schooling outcomes *ex post* in a setting where school investments at different points in time are complementary (in particular Jacoby and Skoufias, 1997; Shah and Steinberg, 2017). However, we

are not aware of any work that considers the question from an *ex ante* perspective.

We also contribute to an emerging literature that seeks to structurally estimate the parameters of the human capital production function, specifically the elasticity of inter-temporal substitution in human capital investments (Cunha et al., 2010; Attanasio et al., 2020; Agostinelli and Wiswall, 2020; Attanasio et al., 2020). Most of this work uses detailed data on skills and investments to estimate a fully dynamic model over the life-time of the child, and focuses on time-intervals of several years. Our paper identifies the substitutability in investments from variation in schooling investments over different time periods for a given child, which is arguably a useful alternative in contexts in which skills are measured with considerable error. Perhaps more importantly, in contrast to previous works, our structural parameters reveal what parents believe to be the true production function, which is arguably the relevant object if one wishes to understand the parental response to risk.³

The remainder of the paper proceeds as follows. Section 2 discusses the theoretical model. Section 3 estimates the structural parameters of the model and generates predictions about the effect of risk in rural India, and then empirically estimates the effect of risk on school attendance and study time. Section 4 simulates the effect of the NREGS on the outcome of interest, and Section 5 concludes.

2 A Model of Schooling Investment with Dynamic Complementarity in Human Capital Production and Risk

2.1 A general multi-period formulation

Our model emphasizes dynamic complementarity in the production function of child human capital. Following the literature (Cunha and Heckman, 2007; Cunha et al., 2010; Attanasio et al., 2020), we model the technology of human capital production to be constant elasticity of substitution, with the stock of human capital after the end of the normal schooling period T being a function of consecutive schooling investments s_κ over that interval.⁴

Because $T + 1$ denotes the beginning of the post-schooling interval, we can define completed human capital as

$$H_{T+1} = \sum_{\kappa=1}^T (\nu_\kappa s_\kappa^\rho)^{1/\rho}. \tag{1}$$

³Of course, if parental beliefs are incorrect and one wishes to estimate the consequences of risk for achieved human capital it is also necessary to know the true production function.

⁴To understand the *ex ante* effect of variance on schooling, it is helpful to treat child schooling investment as a market good. An analogous model in which the only cost of schooling is the opportunity cost of foregone time in home production is presented in appendix A.1. In that model, the parent responds to a negative income shocks by supplying more labor in the labor market *ex post* (as in Kochar, 1999; Rose, 2001; Jayachandran, 2006), while children are required to spend more time in home production (but not in the labor market). The two models are isomorphic as long as the income effect of falling wages due to shocks dominates the substitution effect, which is what we find empirically. We therefore work with the simpler model here.

with the $\nu_\kappa \in [0, 1]$ and $\rho \in (-\infty, 1]$. The degree of dynamic complementarity is entirely governed by ρ , with $\sigma = 1/(1 - \rho)$ being the elasticity of substitution between school investments in different periods κ . We further parametrize the ν_κ as a series of discount factors that sum to one reflecting the different contributions of investment at different periods of a child's life $\nu_\kappa = \frac{\delta^{T-\kappa}}{\sum_{\tau=0}^{T-1} \delta^\tau}$, where $\delta < 1$ implies that later contributions matter more than early contributions and conversely.

Parental utility starting from any period t is assumed to depend on the discounted flow of current and future consumption as well as of the stock of human capital and of assets at $T + 1$. Households have stochastic incomes and may save. Parents maximize expected utility in each period by choosing school investment and net savings, with the remaining income being allocated to consumption. In particular, maximized expected discounted utility at time t is:

$$V_t = \max_{s_k, a_k \forall k=\{t, t+1, \dots, T\}} \mathbb{E}_t \left[\sum_{\kappa=t}^T \beta^{\kappa-1} u(c_\kappa) + \zeta g(H_{T+1}) + f(A_{T+1}) | \Omega_t \right]$$

s.t.

$$c_\kappa \leq y_\kappa + \theta_\kappa + A_\kappa - a_\kappa - s_\kappa \forall \kappa. \quad (2)$$

where c_κ denotes consumption, β is the discount rate, and ζ is a scalar governing the utility return to child human capital. The function $g(H)$ captures curvature in the demand for human capital and $f(A)$ the demand for assets. The variable y_κ denotes non-stochastic income and θ_κ is the stochastic component of income. Information on the joint distribution of future income shocks is captured by Ω_t . Net saving is a_κ and A_κ is the stock of savings at the beginning of κ . The stock is updated in each period according to $A_{\kappa+1} = A_\kappa + a_\kappa$. Note that the choice variables at time t , s_k and a_k for all $k \geq t$ are functions of future income realization over the period t to k as well as the information at the beginning of time t . We omit this dependence for notational convenience.

It is helpful to consider this problem as a recursive one, in which the stock of human capital H_t and the asset stock A_t at the start of each period t are state variables. To do this, however, we need to construct intermediate measures of the stock of human capital H_t that conform with eq. (1). In particular, we define

$$H_{t+1} = \left(\frac{1}{\sum_{\tau=0}^{t-1} \delta^\tau} (s_t^\rho + \sum_{\kappa=1}^{t-1} \delta^\kappa H_t^\rho) \right)^{1/\rho} \quad (3)$$

Note that $H_2 = s_1$ and that recursive substitution of eq. (3) into itself yields eq. (1). Also, the stock of human capital at the start of period $t + 1$ is a measure of human capital relative to a period-specific standard, rather than something that accumulates over time. For example, for constant $s_k = s^*$, $H_{t+1} = s^*$, not, for example, ts^* .

Substituting the budget constraint into the utility function, we can write this objective function

recursively as:

$$V_t(H_t, A_t) = \max_{s_k, a_k \forall k=\{t, t+1, \dots, T\}} u(y_t + \theta_t + A_t - a_t - s_t) + \beta \mathbb{E}_{t+1} [V_{t+1}(H_{t+1}, A_{t+1}) | \Omega_{t+1}]. \quad (4)$$

Note that the value function at time $T+1$ is the parametric function $\zeta g(H_{T+1})$.

2.2 Special Case: Perfect Complements and Linear Demand

To illustrate the *ex ante* effect of variance in incomes on schooling investments, we write down a simple two-period formulation of eq. (2) with $u(c) = \ln(c)$ and set aside discounting ($\beta = 1$). We also mute savings, set the $\theta_1 = 0$, and assume that θ_2 can take on two values $[r, -r]$ for $r > 0$ with equal probability. This simplification has the advantage that it allows us to clearly isolate the *ex ante* effect of variance by focusing on first-period schooling investments, and to abstract from any additional effect of risk on schooling via a precautionary savings motive. We then proceed in three steps: *First*, we present analytical solutions for the extreme case in which $\rho \rightarrow -\infty$ and $g(H) = H$. *Second*, we numerically illustrate the effect of risk for intermediate values of ρ and alternative specifications of $g(H)$. *Third*, we show numerical solutions for an extended model that incorporates precautionary savings in appendix A.2.⁵

A simplified two-period model has an objective function

$$V_1 = \max_{s_1, s_2} \ln(y_1 - s_1) + \mathbb{E}_1[\ln(y_2 + \theta_2 - s_2) + \zeta g(H_3)], \quad (5)$$

and final stock of human capital given by

$$H_3 = (\nu s_1^\rho + (1 - \nu) s_2^\rho)^{1/\rho}. \quad (6)$$

with $\nu = \delta/(1 + \delta)$. When first and second period school times are perfect complements $H_3 = \min\{s_1, s_2\}$, second period schooling is never chosen to be higher than first-period investment. We can distinguish two cases. In the first case, the realized income in the second period is sufficiently high that the household chooses second-period schooling to be equal to first-period schooling but would choose higher investment in period 2 if period 1 investment were exogenously higher. We call this the binding case. In the second case, second-period income is sufficiently low that $s_2 < s_1$. In this non-binding case second-period schooling would not be affected by an exogenous increase in first period investment. For $g(H) = H$, we have the binding case if $y_2 + \theta_2 > s_1 + \frac{1}{\zeta}$, in which case $s_2 = s_1$. Otherwise, $s_2 = y_2 + \theta_2 - \frac{1}{\zeta}$. Note that in the first (binding) case, child schooling does not depend on the income shock, but in the second (non-binding) case, schooling increases with the income shock (θ_2).

By substituting optimal s_2 into V_1 and taking expectations over the assumed distribution of

⁵Allowing for savings essentially yields a negative predicted effect of risk throughout, which is consistent with the notion of deterring investments in the presence of risk. Yet, in a model with savings, the expected value of second-period (*ex post*) investment actually increases with variance, which is not what we find empirically. We also find no evidence that risk increases or decreases savings across years in our context. We return to this point in Section 3.3.

shocks, we can then write the maximand as a function of child school time in the first period only. This expression is straightforward in the case that the variance of the shock is small enough so the binding case is achieved — that is as long as things cannot be not so bad in the second period that the household chooses to provide less schooling in period 2 than was chosen in the first period. In particular, assuming we have an interior solution, we can differentiate with respect to s_1 . While the resulting expression cannot be solved in closed form, we can write a series approximation (to second order) in the size of the shock for small values of r :

$$s_1 = y_1 - \frac{2}{\zeta} - \frac{\zeta}{4}r^2 + O(r^3) \quad (7)$$

where we assume that $y_2 = y_1$ to keep the solution simple. This result shows clearly that first period schooling is decreasing in the variance of the shock in the binding case.

On the other hand, if r is larger, such that $s_2 < s_1$ in the bad state of the world, s_1 can be approximated by

$$s_1 = y_1 - \frac{3}{\zeta} + \frac{1}{3}r - \frac{2\zeta}{27}r^2 + O(r^3). \quad (8)$$

This expression suggests that, at least for moderate values of r , first-period schooling could rise with r .⁶ In the bad state, when $s_2 < s_1$, human capital depends only on second period investment given the assumption of perfect complements. This result implies that the *ex post* return to first-period schooling is zero and does not decline further with increasing variance. On the other hand, in the good state, human capital depends on both first and second period schooling. As the variance increases, the utility gain of additional schooling in the second period under the good state increases, as does the *ex post* return to first-period investment. Thus, in expectation over both states, the return to first-period investment rises in r .⁷

But interpreting these expressions is challenging because, *one*, they are approximations in r for small r , and *two*, which equation applies depends itself on the magnitude of r . A numerical simulation helps to clarify these issues. In figure 1, we graph s_1 , s_2 by state, and $\mathbb{E}[H_2]$ by the standard deviation of the shock ($\sigma_\theta = r$), assuming parameter values of $y_1 = 10$, $y_2 = 10$, and $\zeta = 1$.

For low values of r , second period investment equals first period investment, which then also equals total human capital, regardless of the state of the world, and all three measures decline in variance. At $r \approx 1.6$ the measures divide. Now the household chooses $s_2 = s_1$ in the good state of the world, but in the bad state s_2 is lower than s_1 . Note that initial period investment is now rising in r over this range. However, second period schooling falls with variance (and faster than

⁶However, note that in the bad state, $s_2 = y_1 - r - 1/\zeta$ such that it declines faster in r than s_1 (and s_2) increases in r in the good state. Therefore, the expected value of second-period investment and thus expected human capital is always declining in the variance the shock.

⁷This result has an analogy with respect to borrowing in the case of investment risk: increasing return variance can increase the willingness to take a loan at a given interest rate because it raises the return in the good state and does not decrease it in the bad state, because defaulting is the only option in the bad state throughout.

s_1 rises), such that the expected value of human capital declines with variance throughout. We can think of this as a combination of an *ex ante* and *ex post* response in the presence of dynamic complementarity. Clearly, for given first period investment, a given mean preserving increase in the gap between school time in the good and bad states results in lower human capital accumulation as would be the case in an *ex post* model. But the fact that first-period schooling is lower throughout than in the no-variance case, and thus caps good-state investment, reflects the *ex ante* response.

Note that this pattern is the opposite of what would happen under a precautionary savings motive. In a typical precautionary savings model, period one investment is a substitute for income in the second period, and thus one avoids the downside risk by reducing consumption in the first period as variance increases. For human capital with a complementary production function, first period investment is a complement with second period investment. Thus one invests less in the asset (human capital) when faced with greater risk.

2.3 The Effect of Risk for General Values of ρ and ϕ

While the case of perfect complements nicely illustrates the *ex ante* effect of risk, it imposes strong assumptions of the production function of human capital. We therefore present numerical solutions for more general values of ρ and curvature in the demand for child human capital.

Before we do this, we introduce a new concept, namely dynamic complementarity in the demand for human capital. The demand for child human capital exhibits dynamic complementarity if $d^2g(H_3(s_1, s_2))/ds_1ds_2 > 0$, so that an increase in first period schooling leads to an increase in the marginal utility of second period schooling.

In particular, let $\zeta = 1$ and the demand for child human capital being constant relative risk aversion (CRRA):

$$g(H) = \frac{1}{1-\phi}(H)^{1-\phi}, \quad (9)$$

with $\phi \geq 0$.⁸ With $g(H)$ as described by eq. (9) and a child human capital production function as described by eq. (6), the demand for child human capital will exhibit dynamic complementarity if and only if $1 - \phi - \rho > 0$. To see this, substitute eq. (6) into (9), and consider the cross-derivative with respect to s_1 and s_2 :

$$\frac{d^2g(H_3(s_1, s_2))}{ds_1ds_2} = (1 - \phi - \rho)H_3^{(1-\phi-2\rho)}(s_1s_2)^{\rho-1}\nu(1-\nu). \quad (10)$$

For any $1 > \nu > 0$, $s_1 > 0$ and $s_2 > 0$, this expression is strictly positive as long as $1 - \phi - \rho > 0$.

A few numerical examples illustrate the relationship between $d^2g(H_3(s_1, s_2))/ds_1ds_2$, ϕ and ρ . For $\phi = 0$, *i.e.* linear utility with respect to child human capital, there will always be dynamic complementarity in demand, except in the extreme case where the human capital production function exhibits perfect substitutability ($\rho = 1$). On the other hand for logarithmic utility ($\phi = 1$), school

⁸At $\phi = 1$, $g(H) = \ln(H)$.

investments at different points in time will be dynamic complements if and only if the production function exhibits dynamic complementarity (*i.e.* $\rho < 0$). Finally, with $\phi = 2$ school investments at different points in time will be dynamic complements for any $\rho < -1$.

Equipped with this insight, we numerically solve for optimal s_1 at three representative values of ϕ and at various values of ρ . In particular, we calculate optimal investment for three functional forms of $g(H)$: $g(H) = H$, $g(H) = \ln(H)$ and $g(H) = -1/H$, which correspond to the three values of ϕ that were discussed above (*i.e.* $\phi = 0$, $\phi = 1$, $\phi = 2$).⁹

Figure 2 shows the derivative of s_1 with respect to r for these three functional forms and $-3.5 \leq \rho < 0.8$. Since we are plotting against $\ln(\sigma)$, the y-axis captures the Cobb-Douglas case. As $\ln(\sigma) \rightarrow 1.5$, the production function approaches a simple additive function in school times, *i.e.* the case of perfect substitutes. And as $\ln(\sigma) \rightarrow -1.5$, the human capital production function exhibits increasing dynamic complementarity.

As can be seen, the numerical solutions suggest that the sign of the effect of an increase in r on first-period human capital investment maps 1:1 with the presence of dynamic complementarity in demand for human capital as described above. That is, with linear $g(H)$ the predicted effect of risk remains negative throughout, and approaches zero only as the human capital production function exhibits perfect substitutability ($\rho = 1$). With concave $g(H)$, this is no longer the case. Second period risk will then affect the expected marginal utility of the human capital stock, and households choose to invest more in first-period schooling in order to avoid very low stocks of human capital for their children. If $g(H) = \ln(H)$, for example, first period schooling will increase in risk for all $\rho > 0$. If $g(H) = -1/H$, the predicted effect of risk on human capital investments turns positive at values of $\rho > -1$.¹⁰

In contrast, child human capital declines for almost all specifications in these simulations. Figure 3 presents the effect of an increase in r on the expected value of human capital for different values of ϕ and ρ . Not surprisingly the largest effects are observed for linear preferences and highly complementary inputs in the production function and the smallest effects (that turn weakly positive) arise with a strong degree of risk aversion and very substitutable production. This latter result arises because, with substitutability, human capital becomes a mechanism for precautionary savings. In any case, there are moderate negative effects over much of the range. For example, when the elasticity of substitution in human capital production is one, the effect of 0.1 increase in r on $\mathbb{E}[H_3]$ is -0.009 when $\phi = 1$ and -0.026 when $\phi = 0$.

Whether variance increases or decreases first-period investment, at least in a two-period model, thus depends on the presence of dynamic complementarity in the demand for human capital. Dynamic complementarity in the demand for human capital is essentially a composite of the dynamic complementarity in human capital production and the curvature in parental demand for child hu-

⁹Here we set $y_1 = y_2 = 10$, $\zeta = 1$, $\nu = 1/2$ and $r = 1$.

¹⁰Note, that the effect of an increase in risk is not constant in r . Importantly, with linear g , and values of $r \geq \frac{y_1}{2}$, the effect of risk on first-period human capital investment turns positive at very low values of ρ , which is analogous to the case of perfect complements described above. The effect of risk on first-period school maintains its sign for values of $\phi > 0$. Numeric approximations for $y_1 = y_2 = 10$, $\zeta = 1$, $\nu = 1/2$ and $r = 5$ are depicted in figure C.1 (figures C.1 to C.7 are available in appendix C).

man capital. While there is emerging consensus about the complementarity in the production, this does not necessarily map perfectly with what parents believe to be the true parameter. Furthermore, the shape of human capital preferences is unclear, and — to the best of our knowledge — has not been assessed empirically so far. Yet, deriving clear predictions on the effect of risk on human capital investment remains impossible without an intuition about the values of ϕ and ρ . Below we propose a novel approach for estimating these values.

3 Empirical Strategy and Results

Our empirical analysis proceeds in two steps. First, we structurally estimate the key model parameters. This allows us to generate predictions about the effect of risk on schooling investments in the context we study. Second, we test the predictions empirically by exploiting the variation across survey rounds in village-level irrigation, and income risk.

In both steps, we use the Rural Economic and Demographic Survey (REDS) data which were collected in rural India in 1981-82, 1998-99 and 2007-08, which we combine with high-resolution precipitation data (Copernicus Climate Change Service, 2017). Two unique features of the data stand out relative to other datasets: *First*, the REDS covers an extensive time period and a variety of geographic locations, allowing us to investigate the effects of changes in risk profiles over a long time period. *Second*, the survey contains detailed time-use information, particularly the time allocated to studying on a typical day, which is a key educational input in contexts in which children regularly work on the farm or in the household. Information on time use is collected for all women in the household and their children, and is available for three seasons of the year, which are also marked by very different levels of agricultural activity. The reference months are: October/November (season one), February (season two), and April/May (season three). We provide a detailed description of the data in appendix B.¹¹

3.1 Estimating Model Parameters

It is clear from the above analysis that the *ex ante* effect of income variance on, for example, study time could be either positive or negative depending on the structural parameters of the model. Moreover, it is not immediately clear that the insights from the two-period structure carry over to a more realistic setting, where past shocks will have an effect on current levels of child human capital even in the absence of an *ex ante* response. We therefore build on our theoretical analysis to estimate the key parameters of a fully-dynamic structural model of human capital accumulation with a particular emphasis on backing out the intertemporal elasticity of substitution in the production function as perceived by the parents and the curvature in the demand for child human capital, which are arguably the relevant objects if one wishes to understand the parental response to risk.

¹¹Summary statistics at the household level and the individual level are reported in Tables D.1 and D.2 (Tables D.1 - D.36 are available in appendix D).

3.1.1 Indirect Inference

We estimate the structural model parameters using an indirect inference procedure (Gourieroux et al., 1993). Indirect inference is a minimum distance estimator. But instead of matching the moments of the data directly, indirect inference identifies the parameters of the structural model by matching the parameters of an auxiliary model generated from the theory to comparable parameters generated from the data.

This approach to the structural analysis is importantly governed by the strengths and weaknesses of the REDS data. As noted, we have multiple observations of study time over a calendar year for each child but the survey rounds themselves are widely spaced. While we thus can observe how study time evolves over time for a slice of each child’s life, we do not observe a full trajectory of study times over a full childhood. It is thus necessary to use the within-year information on a cross-section of children from different grades to gain insight into the longer-run process of human capital accumulation. In addition, we do not observe comprehensive measures of the stock of human capital. As with most surveys we have years of schooling and whether someone is in school, but these are far from perfect measures of acquired human capital, particularly in a context in which social promotion may be important. What we do have is measures of inputs and, as our theory suggests, inputs are indicative of the *ex ante* response to variance. Finally, while we have asset data at the beginning of the survey year and an estimate of annual expenditures, we do not have within-year measures of saving or consumption. We do, however, have seasonal information on crop incomes and labor earnings that can be used to identify the pattern of income shock correlations within the year.

3.1.2 Three-period Model of Schooling Investments

The structure of our general model presented in Section 2.1 provides a way to link the within-year information across years. In particular, the CES production function implies that the intertemporal elasticity of substitution between any two periods is the same.¹² In addition, the assumed structure of the ν parameters gives a way of projecting the relationship of study times observed across different seasons within one grade to the relationship within another grade or across grades. It seems plausible that the process of accumulation of child human capital over seasons (within the school year) is informative about the process of human capital accumulation over a child’s lifetime (across school years), as school curricula build on the content of previous grades, such that children who fell behind in the past grade will also have more difficulties to keep up with the content in the current grade, and throughout the entire schooling lifetime more generally (Das et al., 2022; von Hippel and Hamrock, 2019).¹³ While we will use this longer-term structure in our projections of the effects of risk on schooling over the child’s entire lifetime, we base our estimation of the

¹²Of course, this is an important restriction in that nearby periods may be more substitutable or more complementary than those at a longer interval.

¹³This structure also implies that the model ignores the effects of long school breaks or transitions to different teachers and/or schools, which would typically happen between grades.

key parameters on the problem faced in a single grade. We focus on the second grade, accounting for optimal first-grade choices, because our model suggests it is in the early grades where *ex ante* choices are most salient.

Our first step in linking the model to the data is to rewrite the recursive human capital equation, without further loss of generality, into a function over three consecutive periods instead of a single period as we did in eq. (3). The three periods may be thought of as consecutive seasons within the same grade:

$$H_{t+3} = \left(\frac{1}{\sum_{\tau=0}^{t+1} \delta^\tau} (s_{t+2}^\rho + \delta s_{t+1}^\rho + \delta^2 s_t^\rho + \sum_{\tau=3}^{t+1} \delta^\tau H_t^\rho) \right)^{\frac{1}{\rho}}. \quad (11)$$

We also make four additional assumptions to address particular aspects of our setting and our data. *First*, coincident with the idea that grain stocks are a primary source of liquid savings for agricultural households and that such assets are generally depleted at the end of the crop-year (Saha and Stroud, 1994; Waldman et al., 2020), we assume the stock of assets to be non-negative and to fall to zero at the end of the crop-year, which coincides with the end of the school-year. Within the year and subject to the zero constraint, however, we allow savings to be optimized given income realizations and correlations within the year.¹⁴ Limiting attention to three seasons of a single school year, where $A_t = 0$ at the beginning of the school year, we can write the objective function recursively as follows:

$$V_t(H_t) = \max_{s_k, a_k \forall k = \{t, t+1, t+2\}} \mathbb{E}_t[u(y_t + \theta_t - s_t - a_t) + \beta u(y_{t+1} + \theta_{t+1} - s_{t+1} + a_t - a_{t+1}) + \beta^2 u(y_{t+2} + \theta_{t+2} - s_{t+2} + a_{t+1}) + \beta^3 V_{t+3}(H_{t+3}) | \Omega_t]. \quad (12)$$

Second, we assume there are two types of households with varying utility return to child human capital ζ^i for $i \in \{L, H\}$, that are observed with equal probability in the population. This approach, in effect, approximates the distribution of heterogeneity with its first two moments.

Third, for given initial age t from the set \mathbb{Z}_3^{0+} and stock of human capital H_t , we approximate the value function $V_{t+3}(H_{t+3})$, with the parametric function $\zeta^i \frac{1}{1-\phi} (H_{t+3})^{1-\phi}$. We recognize, of course, that the assumed functional form for the value function cannot hold exactly at every grade, as it could, for example, in a linear quadratic system. But as we will show in our simulation exercise where we recursively construct the value functions of earlier grades, this assumption is a reasonable approximation.

Fourth, because we wish to capture in a simple way the joint distribution of incomes in the three seasons of the year, we restrict ourselves to the first two moments of the joint distribution and assume the marginal distributions are the same by season. This implies that we need to fit five features of the data (the mean and variance of the marginal distribution and the correlations between each of the three incomes). We assume that income has three points of support with the

¹⁴Formally, $A_t > 0 \forall t \in \mathbb{Z}^{0+}$, $A_t = 0 \forall t \in \mathbb{Z}_3^{0+}$, with \mathbb{Z}^{0+} being the set of non-negative integers and \mathbb{Z}_3^{0+} the set of integers that are divisible by 3: $\mathbb{Z}_3^{0+} = \{\mathbb{Z}^{0+} | t \bmod 3 = 0\}$.

middle point being the mean of the high and low and set the marginal probabilities to specific non-zero values that add to one, imposing that the marginal distribution is symmetric. We specify fixed probabilities for high, medium and low incomes (3/8, 1/4, 3/8, respectively), but to ensure robustness explore alternative probabilities as discussed in Section 3.1.5. The five parameters capturing the joint distribution of incomes over the year are the high and low incomes $y^H = y + \theta$ and $y^L = y - \theta$ and three parameters ω_1 , ω_2 , and ω_3 that reflect the conditional probabilities of high and low states in period 2 conditional on period 1, and in period 3 conditional on periods 2 and 1, respectively.

Setting $\beta = 1$ for now, this yields 10 structural parameters to estimate, described by $\Psi = [\zeta^L, \Delta\zeta, \phi, \sigma, \delta, y^L, \Delta y, \omega_1, \omega_2, \omega_3]$, where we translated ρ into the elasticity of substitution $\sigma = 1/(1 - \rho)$ for ease of interpretation. Note that $\Delta\zeta = \zeta^H - \zeta^L$ and $\Delta y = y^H - y^L$.

3.1.3 Identification of Structural Model

Before turning to estimation, it is helpful to consider the issue of identification. While we incorporate a number of different bits of information from our data, a primary source of identification is the relationship between the three endogenous study-time measures within the year. A central concern is whether it is plausible that this relationship can help distinguish heterogeneity in preferences from dynamic complementarity. With just two study-observations per child, this would seem to be a fraught task. Heterogeneity in the utility return to child human capital across households is going to lead to a positive correlation between study time at different points in time just as would dynamic complementarity in the absence of heterogeneity. However, the presence of a third observations makes a significant difference. Intuitively, dynamic complementarity yields different implications for s_{t+2} of variation in $s_t - s_{t+1}$ given $s_t + s_{t+1}$ than does heterogeneity.

To see this, consider the implications of the optimization problem in the third season of a given grade t with $\delta = 1$:

$$V_{t+2} = \ln(y_{t+2} + \theta_{t+2} - s_{t+2} + a_{t+1}) + \frac{\zeta}{(1 - \phi)} (H_t^\rho + s_t^\rho + s_{t+1}^\rho + s_{t+2}^\rho)^{\frac{(1-\phi)}{\rho}}. \quad (13)$$

If we then reparametrize so that $sT = s_t + s_{t+1}$ and $sD = s_t - s_{t+1}$ and implicitly differentiate the first order condition for s_{t+2} with respect to sT and sD we get the following:

$$\frac{\partial s_{t+2} / \partial sD}{\partial s_{t+2} / \partial sT} = \frac{(s_t / s_{t+1})^\rho - (s_t / s_{t+1})}{(s_t / s_{t+1})^\rho + (s_t / s_{t+1})}. \quad (14)$$

The denominator of eq. (14) must be positive but the numerator will be positive or negative depending on the sign of ρ and the ratio of the first and second schooling investments. In particular if $s_t > s_{t+1}$, eq. (14) will be positive if $\rho > 0$ and negative otherwise. More generally, conditional on the parameters, H_t , and y_t , this ratio should depend only on the ratio of first and second period school time within the grade. In short with dynamic complementarity unequal study time in seasons one and two for the same average study time will yield lower human capital at the beginning of

season three and thus reduce season three study time.

If there is heterogeneity in preferences, however, H_t will not be the same for all households initially. This implies that the observed distribution of study times will be a mixture of multiple type-specific distributions. While one cannot, therefore, identify ρ with information on the distribution of s_3 conditional on s_1 and s_2 alone, dynamic complementarity of this form imposes substantial structure on how simple forms of heterogeneity can impact the joint distribution of s_1 , s_2 , and s_3 .

Figure 4 presents a non-parametric smoothed version of season three school time as function of s_1 and s_2 data, where we set $t = 1$ so the labels correspond to seasons. The two horizontal axes are school time in the first two seasons and the vertical axis is school time in the third season. Note that the surface slopes upward toward the point where s_1 and s_2 are highest — higher inputs in the first two seasons predict higher inputs in the third season. But also note that along a cross-section where $s_1 + s_2$ is a constant, s_3 rises when s_1 and s_2 are more different. In effect, for $s_1 > s_2$ the slope of s_3 is positive with respect to sT and with respect to sD . This suggests given eq. (14) that $\rho > 0$. But if study time in different parts of the year are substitutes then it is not clear why s_3 should be increasing convexly in sT , at least in the absence of heterogeneity. As we will see below this discrepancy can be resolved by our estimated model.

3.1.4 Estimation of Structural Model

Denote by $B(\Psi)$ the vector of parameters of the auxiliary model obtained by numerically optimizing over the choice of study time and savings over three seasons given the vector of structural parameters Ψ and then summarizing the joint distribution of the resulting endogenous variables with a series of regression coefficients and other combinations of moments. Also let \hat{B} be the corresponding combinations of moments from the actual data. The minimum distance estimator chooses the parameters Ψ that minimize the squared difference between the two vectors, weighted by the inverse of the variance-covariance matrix V of the estimated \hat{B} ,

$$(B(\Psi) - \hat{B})'V^{-1}(B(\Psi) - \hat{B}). \tag{15}$$

The first set of parameters of the auxiliary model arise, as noted, from the relationship between season 3 study time and that in seasons 1 and 2. We focus, in particular, on a quadratic approximation to this surface of the form

$$s_{3i} = \alpha_0 + \alpha_1 s_{1i} + \alpha_2 s_{2i} + \alpha_3 s_{1i}^2 + \alpha_4 s_{2i}^2 + \alpha_5 s_{1i} s_{2i} + \alpha_6' x_i + \epsilon_{3i}, \tag{16}$$

where x_i denotes a vector of individual, household, and village characteristics. Empirical estimates reveal a similar pattern to figure 4, which is very stable across various specifications (see table 1).¹⁵

¹⁵Note that we restrict this analysis to the 1998-99 and 2007-08 rounds of data. As we show in table D.3, the results are identical when using all rounds of data. The results are also unaffected by controlling for household fixed effects, or household-specific time trends, in the subset of panel households, see table D.4. Because season three could

In particular, we see that third-period schooling is increasing and convex in the schooling of each previous period but that the coefficient on the interaction between s_1 and s_2 is negative.

In addition to α_1 - α_5 from eq. (16), our auxiliary model captures the relationship between first and second period schooling, γ_1 and γ_2 as estimated from

$$s_{2i} = \gamma_0 + \gamma_1 s_{1i} + \gamma_2 s_{1i}^2 + \gamma_3' x_i + \epsilon_{2i}, \quad (17)$$

as well as the first two moments of s_1 , s_2 and s_3 .¹⁶

As noted above, the pattern of correlations in income shocks across seasons within a year may importantly influence the relationships between study times in different seasons. Thus the model, in which incomes are determined by draws from a joint distribution of seasonal incomes as well as endogenous study time choices on the part of the students should be able to match observed seasonal income correlations. We could use actual income correlations for this purpose but given possible measurement error in income and the fact that rainfall importantly drives income variation in a rural monsoon economy, it seemed prudent to use correlations in predicted income based on rainfall data. In particular, we note that rains not only affect the productivity of rainy season crops but can influence soil moisture and aquifer depths that may be relevant in subsequent seasons. Farmers also may respond to deficits in previous seasons by farming more or less extensively in subsequent seasons, and this also may affect earnings of agricultural laborers. Thus, we first regress household incomes from agriculture and labor by season on 16 months of lagged rainfall data. We then predict income levels, standardize, and compute correlations between each of the three seasons' incomes.¹⁷ Because these correlations will also incorporate any variation in child labor that is driven by rainfall inclusive of that arising from dynamic complementarity, we match these moments with correlations from the simulated data in income net of study hours, that is $y_t - s_t$.

We construct the 13 parameters of the auxiliary model \hat{B} using the study times and measures of seasonal income from the 1998-99 and 2007-08 rounds of data, as comparable empirical estimates of incomes are not available in 1981-82. While estimating equations (16) and (17) in the actual data, we also control for variables such as age and village characteristics that are relevant to the actual data but are not explicitly modeled in the data generating process, which focuses on the allocation process within a year for a child of given age. We calculate the variance-covariance matrix using the jackknife resampling method, clustered at the level of village-by-round. We differentially weight the estimates associated with eq. (16) by a factor of two, given its central role in identification as discussed above, though results are broadly similar when we use equal weights.

Three points warrant attention when optimizing the theoretical model. *First*, while the esti-

overlap with summer vacations in southern states of India, but not in the northern states, we also test if our results are robust to restricting our data to the northern states. We show in table D.5 that this is indeed the case. Finally, our results are unaffected by slicing the sample along age groups or gender (see table D.6).

¹⁶Results of estimating eq. (17) are displayed in table D.7.

¹⁷Note that income from agriculture and labor are the only income sources available by season. Jointly these two income sources make up for 90% (65%) of total household income in 1998-99 (2007-08). The results of regressing seasonal income on monthly precipitation are displayed in table D.8 and the resulting correlation coefficients in table 2.

mated data effectively have a continuous distribution, the simulated data will take on 54 values: 27 combinations of realized states and two types. For each type of household and set of realizations of y in seasons 1, 2, and 3, we numerically solve for the optimal values of the state contingent and type-specific s^i and a^i given the parameter vector Ψ . We construct the vector of parameters $B(\Psi)$ by weighting observations in the simulated data for the endogenously determined probability of observing each state. *Second*, because the ν parameters vary with grade, *i.e.* the stock of human capital plays a larger role relative to investments in higher grades, it is necessary to estimate the model assuming a particular grade. In our preferred specification, we use second grade as the grade of choice. Referring to eq. (12), that means that $t = 4$ and the parametric value function is $V_7(H_7) = \zeta^i \frac{1}{1-\phi} (H_7)^{1-\phi}$. *Third*, at any given grade a child will have accumulated a certain stock of human capital. This stock may be systematically different by grade, by parental type, and by the history of shocks. To solve for optimal s^i , we need to make some assumption about the initial stock. Building on the decision to estimate off of the second grade choice, we can recover H_4^i by recursively solving for the value function at the beginning of period four $V_4(H_4)$ given the estimated structural parameters Ψ .¹⁸ With that value function at hand, we solve for state contingent s_1^i , s_2^i , s_3^i and $E[H_4^i]$. The initial stock of human capital for each type H_4^i is then found by iterating over starting values H_4^H and H_4^L and solving for Ψ until the predicted stock (in expectation) for each type given Ψ matches with the starting values.

3.1.5 Structural Results

Of particular interest is the extent to which the model captures the basic structure of the data over the relevant ranges. Table 2 presents the targeted parameters and the simulated parameters of the auxiliary model. The coefficients are appropriately signed and largely of the correct magnitude. More important than the specific coefficients is the ability of the model to capture the overall pattern of the relationship between the three school times. Figure 5 provides a picture of the quadratic surfaces for the regressions based on the actual (lighter) and simulated data (darker). The horizontal axes are s_1 and s_2 and the vertical axis is s_3 . The basic shapes of the curve are similar, with the simulated data lying above the actual data for high values of s_2 but closely coinciding elsewhere.

Table 3 presents the estimated parameter vector Ψ with standard errors, along with the sensitivity matrix Λ (Andrews et al., 2017) that describes how the matched moments (α and γ estimates) map into the parameters of interest and thus illustrate the primary sources of identification for each parameter of the model. Standard errors are calculated based on the variance-covariance matrix of the estimates from the actual data.¹⁹

The estimates of the structural parameters (presented in Panel A) are relatively precise in percentage terms. The elasticity of substitution (σ) is 0.290, which corresponds to a value of

¹⁸The procedure we use to recover the value function V_4 is described in more detail in Section 3.1.6.

¹⁹If $S = dB(\Psi)/d\Psi'$ and because our weighting matrix is V^{-1} sensitivity $\Lambda = (S'V^{-1}S)^{-1}S'V^{-1}$ and the parameter variance-covariance matrix is $\Sigma = \Lambda V \Lambda'$.

$\rho = -2.45$. There is thus a strong degree of complementarity in the production function. But preferences over human capital are also quite concave. The important difference $1 - \phi - \rho = 1.951$ indicating that demand is overall complementary. We also see substantial heterogeneity in demand as the gap in the preference parameter ζ is 6.992, a difference of 132% over the lower value.

This substantial heterogeneity is responsible for the convex shape of the relationship between study times across three seasons. To illustrate this point we shut down savings and income correlations and plot the relationship between s_3 and the earlier study times conditional on type in figure 6. As expected given the presence of dynamic complementarity the surfaces for each type is concave. However when we plot the surface that results from combining the two types into one data set we recover a convex surface comparable to that observed in figure 4. The reason is, of course, that the low type is over-represented at low values of s_1 and s_2 , while the high type is over-represented at higher values of s_1 and s_2 .²⁰

There is also considerable uncertainty in income with the high income level being being 186% above the low income level. Finally, the δ of 1.357 indicates that ν falls with age. Study time in the early periods hence contributes relatively more to the final stock of child human capital.²¹

Panel B of table 3 presents the sensitivity matrix Λ as described above. The pattern suggests a complex interrelationship between the coefficients from the auxiliary model and the structural parameter estimates. There is little sense that one or more parameters is particularly "identified by" any one of the coefficients. It is notable that the utility return to human capital estimates (both ζ^L and $\Delta\zeta = \zeta^H - \zeta^L$) are most strongly influenced by the moments (mean and standard deviation) of study time in seasons two and three. Also, changes in all the coefficient values (α s and γ s) lead to decreases in the ζ terms. Likewise, increases in the coefficient estimates lead to decreases in the estimate of the elasticity of substitution. The negative estimate of -0.031 for α_5 and σ indicates that a one standard deviation increase (.005) in the negative season 1 \times season 2 interaction would yield a 3.1% decrease in the elasticity of substitution. We also show in table D.9 that our results are unaffected by selecting alternative probabilities for the high, medium and low incomes, or by using the inverse of the original variance-covariance matrix in eq. (15).

3.1.6 *Ex ante* and *ex post* Effects of Variance over the Child's Lifetime

We can now turn to the central questions of interest, whether, *one* given the estimated parameters of the model, an expansion in the variance of the effective income distribution would result in lower investment in each period as well as in a lower stock of child human capital, and *two* how much of any observed effects is attributable to *ex ante* risk.

We solve the longer-run problem recursively. We first use parameter estimates for ζ and ϕ to construct the parametric value function $V_{T+1}(H_{T+1}) = \zeta^i g(H_{T+1})$ at the end of normal schooling

²⁰The fact that s_3 for the combined curve falls below the low curve is a consequence of extrapolation using a quadratic surface

²¹Note that this structure is likely heavily driven by the within-year nature of our data, and needs not extrapolate beyond school years. It seems worthwhile pointing out that the estimated $1 - \phi - \rho$ does not change substantially if we constrain $\delta \leq 1$ as shown in table D.9.

years which we take as $T = 27$ seasons (9 grades x 3 seasons). Then using our estimated vector Ψ and adjusting for age in the production function, we calculate $V_{T-2}(H_{T-2})$, that is expected utility at the beginning of the last year of schooling for different values of H_{T-2} at the start of the grade $T - 2$, and create a smooth function from the discrete values using a cubic spline. This estimate of the empirical value function $V_{T-2}(H_{T-2})$ is then used given the existing parameters to find $V_{T-5}(H_{T-5})$ at different values of H_{T-5} and then these values are also splined. This process is then repeated for each grade down to the starting grade at school, $t = 1$. Note that V_1 does not depend on H_1 because there is no previous human capital accumulation at that point by assumption.

The derivatives of these value functions with respect to initial H_t over the range of observed human capital stocks are depicted in figure 7 for the low type (panel a) and for the high type (panel b). Note that there are small level shifts between the derivatives of $V_{28} = \zeta^i g(H_{28})$ (parametric value function) and of V_7 (empirical value function at the end of grade 2), but that the broad trend is well aligned. While this lends further credibility to our estimation strategy in which we approximate the empirical value function $V_7(H_7)$ with the parametric function $\zeta^i g(H_7)$, it also highlights the challenges involved in selecting a particular grade for the structural estimation. To further refine the match between what the empirical value function V_7 should be (from solving the model backwards) and what we actually estimate in the data ($\zeta^i g(H_7)$), we iterate over ϕ until the first derivatives of both correspond at the predicted end-of-second-grade stock with variance (6.9 for the low type and 10.3 for the high type, see table 4). We find that the first derivative of the empirical value function V_7 maps most closely with that of the estimated parametric value function for a ϕ in V_{28} that is 0.394 and 0.297 units lower for the low and for the high type, respectively. Note that these modifications imply even greater dynamic complementarity in demand: $1 - \phi - \rho \approx 2.345$ for the low type, and 2.248 for the high type. In figure 7, the derivative of the modified $\zeta^i g(H_{28})$ is shown by the black dash-dot line. The dashed and dotted curves are the marginal empirical value functions from grades two through nine without variance and with variance, respectively. The solid black line shows the derivative of the estimated parametric value function $\zeta^i g(H_7)$.

Equipped with these value functions, we then solve for optimal schooling investments in each grade and the expected stock of human capital at the end of each grade. The expected final stock of human capital is then obtained by recursively applying the grade-specific functions

$$H_{T+1} = H_{T+1}(H_{T-2}(H_{T-5}(\dots H_4(H_1)\dots))) \quad (18)$$

where the $H_{t+3}(H_t)$ is the expected value of human capital at the end of the grade as a function of the beginning-of-the-grade human capital given that study time is optimally chosen. A similar exercise can then be carried out with a counterfactual distribution of incomes and a comparison of these two measures describes the effect of variance on total human capital accumulation. This difference combines the *ex ante* and *ex post* effects accumulated over the entire schooling lifetime.

To construct a measure of the *ex ante* effect across the school-going ages it is helpful to consider the process of human capital accumulation for a child whose parent is forward-looking and responds to income shocks by adjusting study time but who is naive in the sense that the parent does not

anticipate future uncertainty. By construction this parent will not adjust schooling to variance *ex ante*, and consequently any reductions in child human capital can be entirely attributed to *ex post* effects. The *ex post* effects arise because variance in study times, keeping mean study time constant, will lead to lower accumulated human capital in the presence of dynamic complementarity. If we know the total decrease in human capital that arises from variance as well as how much of this decrease is due to the *ex post* effect, then we can ascribe the difference to the *ex ante* effect.

Let $H_{t+3}^k(H_t^k)$ for $k \in \{v, nv, m\}$ represent the expected end-of-grade human capital for parents facing variance, facing no variance, and who are facing variance but are naive as described above, respectively. At the end of any grade, the percentage difference in acquired human capital between the variance and no variance cases can be described by $(H_{t+3}^{nv}(H_t^{nv}) - H_{t+3}^v(H_t^v))/H_{t+3}^v(H_t^v)$. Substituting the first-order Taylor expansion of $H_{t+3}^{nv}(H_t^{nv})$ around H_t^v into this formulation, gives:

$$\frac{H_{t+3}^{nv}(H_t^{nv}) - H_{t+3}^v(H_t^v)}{H_{t+3}^v(H_t^v)} \approx \frac{H_{t+3}^{nv}(H_t^v) - H_{t+3}^v(H_t^v)}{H_{t+3}^v(H_t^v)} + \epsilon_h^{nv} \frac{H_{t+3}^{nv}}{H_{t+3}^v} \frac{H_t^{nv} - H_t^v}{H_t^v} \quad (19)$$

where

$$\epsilon_h^{nv} = \frac{\partial H_{t+3}^{nv}}{\partial H_t^v} \frac{H_t^v}{H_{t+3}^{nv}}. \quad (20)$$

A similar expression compares the human capital accumulation for the naive parent to that of the sophisticated parent with variance

$$\frac{H_{t+3}^m(H_t^m) - H_{t+3}^v(H_t^v)}{H_{t+3}^v(H_t^v)} \approx \frac{H_{t+3}^m(H_t^v) - H_{t+3}^v(H_t^v)}{H_{t+3}^v(H_t^v)} + \epsilon_h^m \frac{H_{t+3}^m}{H_{t+3}^v} \frac{H_t^m - H_t^v}{H_t^v}. \quad (21)$$

These expressions reflect the idea that the end-of-grade difference may be decomposed into two components. The first term arises from the difference in optimal study times given the initial level of child human capital and the second from the initial differences in child human capital (which might grow smaller or larger over time). The ratio of eq. (21) to eq. (19) provides a measure of the fraction of human capital loss due to the *ex ante* effect.

Recovering the *ex ante* effect requires an estimate of $H_{t+3}^m(H_t)$: the end-of-grade stock of the naive parent conditional on any H_t . One approach to estimating the $H_{t+3}^m(H_t)$ would be to build a behavioral model of the naive parent given the estimated parameters that would then generate values of the stock of human capital for each grade. A convenient alternative is to use the fact that a naive parent subject to variance will make approximately the same study-time choices in expectation as the no-variance parent conditional on the initial stock of child human capital. To see this, consider a first-order approximation to the study-time decision rule derived from the

maximization of eq. (12) around $\bar{\theta} = 0$ and $\bar{\Omega} = 0$:²²

$$s_t^*(\theta, \Omega, H_t^v) \approx s_t^*(0, 0, H_t^v) + \frac{\partial s_t^*}{\partial \theta} \theta + \frac{\partial s_t^*}{\partial \Omega} \Omega. \quad (22)$$

Evaluating this approximation at $\theta = 0$ and $\Omega = 0$ yields the no-variance study time: income does not vary with the state of the world. Evaluating at the actual realizations of the shock θ and $\Omega = 0$ yields naive study time: income does vary but parents assume that it does not. Finally, evaluating at the actual θ and Ω yields the study time when income varies and this variation is anticipated. Taking expectations yields:²³

$$\mathbb{E}_t(s_{t+\tau}^m) \approx s_{t+\tau}^{nv}. \quad (23)$$

Differencing a first-order Taylor expansion of the production function around the study times of the no-variance child evaluated at the study times of the naive child from the same approximation evaluated at the study times of the variance child, approximates the difference in human capital between the naive and variance cases:

$$H_{t+3}^*(H_t^v, s_t^m, s_{t+1}^m, s_{t+2}^m) - H_{t+3}^*(H_t^v, s_t^v, s_{t+1}^v, s_{t+2}^v) \approx \sum_{\tau=0}^2 \frac{\partial H_{t+3}^*}{\partial s_{t+\tau}^{nv}} (s_{t+\tau}^m - s_{t+\tau}^v), \quad (24)$$

where H_{t+3}^* denotes the production function. Taking expectations, substituting (23) and dividing by $H_{t+3}^v(H_t^v)$ yields:

$$\frac{H_{t+3}^m(H_t^v) - H_{t+3}^v(H_t^v)}{H_{t+3}^v(H_t^v)} \approx \sum_{\tau=0}^2 \eta_\tau \frac{s_{t+\tau}^{nv} - \mathbb{E}_t s_{t+\tau}^v}{\mathbb{E}_t s_{t+\tau}^v}, \quad (25)$$

where

$$\eta_\tau = \frac{\delta^{2-\tau} s_{t+\tau}^{nv \rho}}{\sum_{k=3}^{t+1} \delta^k H_t^{nv \rho} + \delta^2 s_t^{nv \rho} + \delta s_{t+1}^{nv \rho} + s_{t+2}^{nv \rho}} \frac{H_{t+3}^{nv}(H_t^v) \mathbb{E}_t s_{t+\tau}^v}{H_{t+3}^v(H_t^v) s_{t+\tau}^{nv}}. \quad (26)$$

Finally, we can construct analogous expression to eq. (21) that illustrates the *ex ante* effect on expected study time:

$$\frac{s_t^m(H_t^m) - s_t^v(H_t^v)}{s_t^v(H_t^v)} \approx \frac{s_t^{nv}(H_t^v) - s_t^v(H_t^v)}{s_t^v(H_t^v)} + \epsilon_s^{nv} \frac{s_t^{nv}}{s_t^v} \frac{H_t^m - H_t^v}{H_t^v}, \quad (27)$$

where

$$\epsilon_s^{nv} = \frac{\partial s_t^{nv} H_t^v}{\partial H_t^v s_t^{nv}}. \quad (28)$$

²²Note that in slightly abuse of notation we treat Ω as a scalar here. We can think of this scalar as a parameter that scales the underlying covariance matrix proportionately. For more detail on this decision rule, see Section 3.2.3.

²³Analogously, it may be shown that $\partial H_{t+3}^m / \partial H_t^v \approx \partial H_{t+3}^{nv} / \partial H_t^v$ and $\epsilon_h^m \frac{H_{t+3}^m}{H_{t+3}^v} \approx \epsilon_h^{nv} \frac{H_{t+3}^{nv}}{H_{t+3}^v}$.

This expressions says the total *ex ante* difference in study time in a given grade reflects the grade-specific difference in study times for a given initial stock of human capital plus an effect that arises from the difference in the initial stock of child human capital between the naive and sophisticated parents faced with variance.

Each of the differences can be constructed by solving for optimal s_t, s_{t+1}, s_{t+2} in each grade and for the expected stock H_{t+3} at the end of the grade (with variance and without variance) conditional on the beginning-of-the-grade stock of human capital, and knowing δ and ρ (for which we used the estimated values). In Table 4 we show these computations for each grade and both types and using the predicted stock of child human capital with variance (in expectation) at the beginning of the grade as starting point. Columns (9)–(12) show the cumulative total effect as well as the cumulative *ex ante* effect of variance on H_{t+3} and on s_t , respectively (tables D.10 and D.11 show similar calculations using various alternative beginning-of-the-grade human capital stocks H_t for the low and for the high type, respectively.)

We find, using this approach, that reducing risk by 100% would increase child human capital accumulated over the child’s entire schooling lifetime by 15% for both (low and high) types. Incorporating the information on the naive households we find that, for the low types, approximately 54% of that total effect can be attributed to the *ex ante* effect, yielding an *ex ante* elasticity of child human capital with respect to variance of -0.08. The corresponding figures are 53% and -0.08 for the high types. To gauge the validity of our approach, we also calculate the final stock of child human capital without variance and with variance (in expectation). Using these final stocks to compute elasticities, rather than the linear approximations, yields an elasticity of child human capital with respect to variance of -0.14 for both types. Reassuringly, the linear approximations are not far off. The elasticity of school time with respect to variance is -0.18 for both types and fairly constant throughout the child’s lifetime. About 72-73% of this effect can be attributed to an *ex ante* response.

A few points regarding the decomposition are worth highlighting. *First*, the cumulative *ex ante* effect of variance on the stock of child human capital and on study time declines slightly over time, while the cumulative total effect on H_{t+3} increases over time. This result is not surprising, as the *ex post* effects increase over time as new shocks are realized. The *ex ante* effect in contrast can be observed from the very first grade. By targeting a lower level of accumulated human capital from the start they minimize the loss they might experience in the future should they experience an adverse shock. Effectively, parents only start what they can finish. While there is some additional *ex ante* under-investment in later grades, this gap is modest by comparison with that in the initial grade. For example, for the low type this gap is 11% on average over the three seasons in grade one but only 2% in grade seven, conditional on beginning-of-the-grade human capital.

Second, the fact that $\delta > 1$ leads to a considerable decline in study times (both with and without variance) with increasing grades. While we do observe some decline in study time with age in our data, it seems that extrapolating the δ estimated within the school year to the whole school-age period to learn about the allocation of study time across school years might be problematic —

and is indeed one reason that our estimation focused on an early grade. Two pieces of evidence, however, show that our estimates of the effect of risk on schooling are not driven by this feature of the data. For one, the estimate of $1 - \phi - \rho$ if we constrain $\delta \leq 1$ is 1.42, thus relatively close to the 1.95 obtained from the main estimates. Second, when we perform simulations of the life-time effect of risk on the original set of structural parameters but setting $\delta = 1$, the schooling time decline disappears but the effects of risk are virtually unchanged (see table D.12).

3.2 Estimating the Effect of Risk on Schooling Investments

With these theoretical insights in hand, we are in a position to better interpret more direct estimates of the effect of income risk on schooling in our data. Our concept of risk exploits three key facts from rural India. *First*, agriculture is the dominant source of income for the majority of households over the time period of our study, either through own agricultural production or through casual agricultural employment. *Second*, agriculture is inherently risky. A large body of literature has documented that agricultural yields, wages and employment levels in rural India are strongly influenced by rainfall conditions (see e.g. Jayachandran, 2006; Shah and Steinberg, 2017; Kaur, 2019). High rainfall leads to good harvests, high demand for labor, and high wages. In contrast, low rainfall levels lead to poor harvests, and low demand for agricultural labor.²⁴ *Third*, the use of irrigation in agricultural production expanded rapidly in rural India over the time-period covered by our data, reducing the exposure to weather shocks in some places but not in others (Duflo and Pande, 2007; Mondal et al., 2015; Ambika et al., 2016).²⁵

3.2.1 Rainfall Shocks in Rural India

Before discussing the construction of our risk variable, we first show that rainfall shocks matter economically in our context. In particular, we test if rainfall shocks affect agricultural income per capita (in logs), consumption expenditures per capita (in logs), and — given the focus of this paper — the time allocation of children per day to home production (own agricultural production and household chores) and to studying (in school or at home).²⁶ Similarly to previous studies that look at the role of rainfall shocks in determining economic outcomes in India, we focus on variation in relative rainfall (see e.g. Jayachandran, 2006; Shah and Steinberg, 2017; Kaur, 2019). Given the structure of our data, we construct historical rainfall distributions for each village and estimate if deviations from village average rainfall affect the outcomes of interest. Exploiting the exogenous

²⁴More recent work also highlights the role of temperature on agricultural output in India: directly, because it affects the pace of crop growth, and indirectly, because temperature influences the rate at which water evaporates from the ground, and thus soil moisture levels (Mondal et al., 2015; Colmer, 2021b).

²⁵Irrigation water in India mainly stems from two sources: dams (surface water irrigation) and bore-wells (groundwater irrigation).

²⁶Agricultural income is the total value of agricultural production corrected for (paid-out) operating expenses (but not for the value of family labor or the rental value of owned machinery).

nature of rainfall shocks, we estimate

$$y_{ijst} = \beta_0 + \beta_1' \theta_{jst} + \beta_2 p_{jst} + \sum_{n=1}^3 \beta_{2+n} \theta_{jst-n} + \gamma_j + \psi_{st} + \epsilon_{ijst}, \quad (29)$$

where y_{ijst} is any of the outcome variables discussed above, observed in year i , village j , state s and year t . θ_{jst} is a vector of dummies that indicate in which quintile of the historical rainfall distribution (of each village) the current realization of rainfall is located. We also control for the share of village agricultural area that is irrigated p_{jst} , for rainfall (in log mms) in the last 3 years $\sum_{n=1}^3 \theta_{jst-n}$, for village fixed effects γ_j , and for state-by-round of interview fixed effects ψ_{st} .²⁷ We report standard errors that are corrected to allow for clustering at the level of the grid cell by round of interview, as well as Conley (1999) standard errors that allow for arbitrary spatial auto-correlation (300km cut-off, Bartlett weights).²⁸

We present the fitted values obtained from estimating eq. (29) in figure 8. In line with previous work, our findings suggest that low rainfall, measured as rainfall outcome that falls below the 20th percentile of the historical rainfall distribution in a particular village, severely reduces incomes from agricultural production and consumption expenditures, while rainfall above the 80th percentile (60th percentile) of the historical rainfall distribution seems to increase agricultural output (consumption), see panels A and B. We also find that low rainfall is associated with more time spent in home production (*i.e.* own agricultural production or household chores), and less time spent studying (in school or at home). Conversely, high rainfall (*i.e.* rainfall above the 80th percentile of the rainfall distribution) reduces the time spend in home production, and increases the time spent studying (see panels C and D of figure 8). These effects are economically meaningful: Moving from the lowest rainfall quintile to the highest rainfall quintile increases the time spent in school on a typical day by a little over 40 minutes on average. These results differ from previous work: Shah and Steinberg (2017) find that positive rainfall shocks are associated with lower school attendance among children in rural India. Given that the rainfall-school effect reflects the sum of opposite-signed income and substitution effects it is perhaps not surprising that the overall effect will vary given different sampling schemes, time periods, and measures of educational investment. In any case, what matters for our purpose is that rainfall affects the choice of study time, not the sign of this effect.

3.2.2 Consumption Variability as Measure of Risk

After having established that rainfall shocks matter economically in our sample, we turn to developing a measure of risk. In order to do this, we need to move away from a relative measure of rainfall to an absolute one. In other words, we need an estimate of the absolute effect of rainfall (conditional on village characteristics) on the outcomes of interest. Given the growing importance

²⁷When the dependent variable is child time use, we additionally control for age-by-gender and season-by-round of interview fixed effects.

²⁸We use the spatial-HAC standard error correction implemented by Hsiang (2010) and Colella et al. (2019). The Conley standard errors are largely identical to standard errors clustered at the grid cell by round, *c.f.* table D.13.

of irrigation in our context, we allow the absolute effect of rainfall to vary with the availability of irrigation. We hypothesize that a higher the share of irrigated land in a village will mediate the effect of rainfall shocks on household-level outcomes. This could happen because households are more likely to have irrigation on their own land in villages with a high share of area under irrigation or because (wages and employment in) casual agricultural employment will be less dependent on good rainfall realizations (as only a small share of agricultural production in the village would be solely rain-fed).

We estimate

$$y_{ijst} = \beta_0 + \beta_1\theta_{jst} + \beta_2p_{jst} + \beta_3\theta_{jst}p_{jst} + \sum_{n=1}^3\beta_{3+n}\theta_{jst-n} + \gamma_j + \psi_{st} + \epsilon_{ijst}, \quad (30)$$

where θ_{jst} now is the total precipitation in the current agricultural year (in log mms).²⁹ All remaining controls are defined as in eq. (29). y_{ijst} is the outcome of interest. As above, we estimate the effect of rainfall on agricultural income per capita (in logs), and consumption expenditures per capita (in logs). We also estimate the effect of rainfall on maternal time allocation to household chores and to the labor market, and on the time children allocate to own agricultural production, household chores, studying, and to leisure. In the time-use specifications, we disaggregate the sample by gender (child-level), and control for age and for season-by-round (of interview) fixed effects. Standard errors are again corrected to allow for arbitrary spatial auto-correlation.

Identification in this exercise relies on the exogenous nature of rainfall, and — because we are interested in the coefficients on irrigation and on its interaction with rainfall — on the assumption that irrigation at the village level is uncorrelated with the error term. The village fixed effect removes any time invariant heterogeneity, but cannot control for unobserved time-varying shocks that might affect the outcomes of interest and irrigation availability simultaneously. The identifying assumption is violated — for instance — if unobserved shocks or underlying trends affect the availability of irrigation at the village level as well as the covariance between consumption outcomes and precipitation in this village. Under which conditions would these violations generate a negative bias on the coefficient on the interaction term? If villages that are wealthier initially experience a faster expansion in irrigation, and also have better means of self-insuring against shocks, then this could bias our estimates. Similarly, unobserved shocks that improve self-insurance, and irrigation availability, would generate a similar negative bias. In both cases this would not necessarily mean that our risk variable does not capture risk, since improved self-insurance still generates a consumption portfolio that is less exposed to shocks. More worrisome are violations of the identifying assumption that might have direct or indirect effects on investments in child human capital through other mechanisms than risk. An example could be that places that have experienced adverse climatic conditions had to diversify out of agriculture. The few farmers that remain are profitable only because they irrigate their land. These places would probably also display a lower correlation between rainfall and consumption, but not because agricultural production became less risky.

²⁹We show in tables D.14 and D.15 that specifications that use rainfall in levels or in squares do not fit the data better than those that use the log of rainfall.

And places with a more diversified economic structure probably also have higher returns to human capital, which would incentivize parents to invest more in the human capital of their children. We return to this point in the robustness checks.

Estimates of eq. (30) are reported in tables 5 and 6. In table 5, we present results with and without controlling for village fixed effects to explore the robustness of our estimates. Controlling for village fixed effects seems to attenuate the coefficients somewhat, but the interpretation remains largely identical. Analogously to figure 8, we first explore the relationship between rainfall, irrigation and agricultural income (cols. (1) and (2)). Our results suggest that higher rainfall increases agricultural income; but that rainfall affects these outcomes to a smaller extent in villages with high irrigation levels. Columns (3) and (4) show the relationship between rainfall, irrigation and consumption expenditures. The effects are somewhat smaller, but more precisely estimated, and suggest a similar relationship: higher rainfall increases consumption, but less so in villages with a higher share of agricultural area that is irrigated. Consistent with the notion of an *ex post* labor supply response by the parent, log rainfall is negatively (positively) associated with the time mothers allocate to the labor market (household chores). Again these associations are weakened by the availability of irrigation.³⁰ Note, that we can only observe the time-allocation of the mother in our data, not of the father. We can expect her response to affect children’s time allocation as long as the mother’s response to shocks is not perfectly undone by the father’s response (*i.e.* the mother increases her working hours but the father reduces working hours by exactly the same amount, and *vice versa*).

The effects of rainfall on child time use are reported in table 6. For brevity, we only report estimates that control for village fixed effects.³¹ Again, rainfall is negatively associated with children’s time spent in own agricultural production and on household chores, and positively associated with time spent in school or studying at home. As with mothers, the availability of irrigation at the village level seems to weaken these associations. These findings are consistent with a model in which parental and child time are substitutes in home production.³² While responses to shocks seem to dominate in the farm-sector for boys, they are more pronounced in household chores for girls (at least in absolute terms). The schooling effects, then, are roughly similar for boys and girls (if anything slightly more pronounced for girls). Finally, we explore if rainfall shocks affect the leisure time of children. We find weak evidence that boys’ leisure increases with rainfall, but no evidence of an association between leisure and rainfall for girls. This suggests that boys can reduce (increase) leisure time to some extent when faced with negative (positive) shocks, while girls cannot.³³

With the estimates of column (3) of table 5, we then simulate the variability of consumption for

³⁰The coefficients are statistically significant only for time allocated to the labor market, not for household chores.

³¹The corresponding results without village fixed effects are largely identical, as reported in table D.16.

³²We cannot observe paternal time allocation in our data. Given that the effect of maternal time allocated to household chores is not statistically significant, while we still see strong responses of children, it seems likely that fathers also reduce the time they spend on chores and on-farm.

³³Note that leisure time includes sleep time.

each household in our sample at a particular point in time.³⁴ *First*, using the annual rainfall data from 1979 to 2019, we calculate the probabilities of all rainfall outcomes per grid cell.³⁵ *Second*, we predict consumption expenditures per capita (in logs) for each rainfall outcome observed in our data given the availability of irrigation in the village in a particular round of interview. *Third*, we construct the probability distribution of consumption outcomes for each household in a particular village and round of interview by multiplying the probability of each of these rainfall outcomes with the predicted log consumption at that rainfall level (given irrigation). *Fourth*, we calculate the interquartile range as well as the standard deviation of a household’s predicted log consumption per capita as measures of risk.

Figure 9 illustrates our calculations. Panel A depicts the historical rainfall distribution of all REDS villages, grouped by quartiles of the share of agricultural area that was irrigated in 1982. We can see small differences in the distribution of rainfall, but no clear link between the distribution of rainfall and the availability of irrigation. Panels B to D, then, plot the distribution of predicted consumption per capita (in logs) in the years 1981-82, 1998-99 and 2007-08, respectively. Again the villages are grouped into quartiles of irrigation availability in 1981-82. As expected, the villages with the highest levels of irrigation in 1981-82 show the least variability in predicted consumption in that same year. In 1998/99 this general patterns seems to remain, although the initial differences in terms of variance are somewhat attenuated. In 2007/08 then, the picture looks substantially different with villages of the lower 3 quartiles (of initial irrigation) displaying roughly similar variance in predicted consumption, while the villages in the highest quartile of initial irrigation seem to have developed in an entirely different direction, displaying substantially lower variability in consumption.

3.2.3 The *ex ante* Effect of Variance on Schooling

In a multi-period model with savings, the reduced form decision rule for schooling investments can be described by

$$s_t = s_t(A_t, H_t, \theta_t, y_t, \Omega_{t+1}), \quad (31)$$

where the Ω_{t+1} term captures how variance in future incomes affects current school time given the stock of assets and human capital at the start of the period as well as current income.³⁶ The formulation in eq. (31) illustrates the challenges in estimating the *ex ante* effect of risk in any single period: If the income process is stationary, future variability Ω_{t+1} will be the same as variability in the past Ω_t , and the stock of human capital at any point in time H_t will reflect the cumulative

³⁴This approach is inspired by Dercon and Christiaensen (2011). We also use the estimates in column (4) to calculate the risk variable, and repeat the main regression as robustness check. The results are largely identical, as reported in table D.17.

³⁵The probabilities are obtained by dividing the sample rainfall distribution in 0.025 intervals of annual log rainfall. We then calculate the historical probability of village-level rainfall to fall in each of these intervals.

³⁶The first derivative of eq. (4) with respect to s_t yields the first order condition: $-u'(y_t + \theta_t + A_t - a_t - s_t) + \beta \mathbb{E}_{t+1} \left[\frac{\partial V_{t+1}(H_{t+1}, A_{t+1})}{\partial H_{t+1}} \frac{\partial H_{t+1}}{\partial s_t} | \Omega_{t+1} \right] = 0$. A comparable first order condition can be obtained for saving.

effects of past shocks (*ex post* effects) as well as past adjustments in study time to future income variability (lagged *ex ante* effects).

For notational simplicity we focus on the values of t corresponding to the start of each crop/school-year, that is, the set \mathbf{B} of t such that $\text{mod}(t, 3) = 1$. If weather is i.i.d. across years but dependent within year we can think of y and Ω as summarizing the mean and covariance of seasonal incomes within a year. Because beginning-of-the-grade stock of human capital can be described by $H_t(\Omega_t, \theta_t, y_t, A_{t-1}, H_{t-1})$ we can then rewrite the decision rule for each $t \in \mathbf{B}$ to

$$s_t = s_t(A_t, \theta_t \dots \theta_1, y, \Omega), \quad (32)$$

which is obtained by recursive substitution. This formulation highlights that the *ex ante* effect can be isolated by controlling for past shocks rather than H_t .

In order to quantify the *ex ante* effect of variance on child schooling, we estimate a linear approximation to the decision rule described by in eq. (32). In particular, we estimate

$$s_{ijst} = \beta_0 + \beta_1 R_{jst} + \beta_2' A_{ijst} + \beta_3 y_{ijst} + \sum_{n=0}^3 \beta_{4+n} \theta_{jst-n} + X_{ijst} + \gamma_j + \psi_{st} + \xi_t + \epsilon_{ijst}, \quad (33)$$

where the dependent variable s_{ijst} is the study time of child i , living in village j , and state s in year t .³⁷ We focus on extensive margin and intensive margin responses by using school attendance (*i.e.* a dummy that equals one if study time is non-zero) and hours spent studying (in school and at home) on a typical day of each season as dependent variables. We are particularly interested in estimating β_1 , *i.e.* the effect of village-level risk R_{jst} on schooling.

Our specification controls for household income y_{ijst} and wealth A_{ijst} . We use household income per capita (in logs) during the reference period for y_{ijst} , and the value of all assets owned at the beginning of the reference period (again in per capita terms and in logs), the area owned (per capita, in log acres), and household size as measures of A_{ijst} . We use annual precipitation (in log mms) as measure of the shock θ_{jst} , and include the current realization, the realizations in the three years prior to the survey, as well as the standard deviation of annual rainfall over these three years as controls. We also interact the lagged rainfall realizations with a dummy that takes the value one if the child was of school age (*i.e.* 6 and above) in that particular year in order to allow for past rainfall realizations to matter more for older children, *i.e.* we add $\sum_{n=1}^3 [\theta_{jst-n} \times 1(\text{SchoolAge}_{ijst-n})]$ to eq. (33). As noted above, controlling for these shocks allows us to isolate the *ex ante* response to risk.

We also control for X_{ijst} , a vector of age-by-gender fixed effects, that captures for example any differences in the implicit price of schooling by age and gender (see appendix A.1). γ_j is a vector of village fixed effects, ψ_{st} a vector of state-by-round (of interview) fixed effects, and ξ_t a vector of season-by-round (of interview) fixed effects. This specification allows us to compare children of the same age and gender from the same village, who are surveyed in different rounds. State-by-round

³⁷Note the slight abuse of notation t here. We observe three seasons of study time for each child, and also control for season-by-round of interview fixed effects ξ_t . Yet, risk, rainfall and state-by-round (of interview) fixed effects are measures that remain constant within the year, and do not vary per child.

of interview fixed effects control for aggregate shocks at the state level, and season-by-round effects control for temporal variation in time allocation. The identifying assumption is that — conditional on the control variables and the fixed effects — no other (unobserved) characteristics that determine hours in school are correlated with our regressor of interest, *i.e.* risk. This assumption is quite strong and we discuss several concerns about identification (and how we address these) in the following section.

The term ϵ_{ijst} reflects unobserved time-specific shocks to study time. We cluster our standard errors at the level of the village by round of interview in all specifications (the level of variation of our main regressor of interest). However, because our measure of consumption risk is not directly observed but predicted for each village and round, these standard errors might not be correct. We therefore additionally report p-values obtained from percentile-t cluster bootstrap (null-imposed, 999 replications) in which we re-sample at the village level. The risk variable is re-calculated for every bootstrap sample, and the distribution of t-statistics is then obtained by regressing the outcome of interest on a newly predicted risk variable in every bootstrap sample.³⁸ Note that this is a highly conservative approach, as entire villages, rather than individual observations, are being re-sampled to calculate risk, and to estimate the effect of risk.

OLS estimates of extensive margin responses are reported in table 7. We add control variables step-wise to gauge the sensitivity of our estimates to the choice of specification. We have seasonal data for 22,767 children in the age group 6 to 15. We use two variables as proxies for risk: the interquartile range (panel A), and the standard deviation of predicted log consumption (panel B). Column (1) controls for village and for season-by-round of interview fixed effects, column (2) adds rainfall controls to account for the effect of current and past shocks. In column (3), we additionally interact each lag with a dummy that is unity if the child was of school age in that particular year. This allows for the effects of shocks to differ by age. Column (4), then, adds wealth controls, and column (5) adds state-by-round (of interview) fixed effects.

We find that risk considerably reduces study time at the extensive margin. The point estimates are largely insensitive to the inclusion of any of the controls. According to the most conservative estimates (column 5), a one standard deviation increase in the interquartile range of predicted log consumption (0.030) reduces the probability that children attend school by 2.7 percentage points. Likewise, an increase in the standard deviation of predicted log consumption by one standard deviation (0.023) reduces the probability that children are in school by 2.3 percentage points. The effect is statistically significant throughout: the percentile-t bootstrap p-values range from 0.028 to 0.081. These point estimates suggest that the elasticity of school attendance with respect to risk ranges between -0.047 and -0.036, which is somewhat smaller than the elasticity that was predicted from the structural estimates.

We explore the effect of risk on the study time (in school or at home) on a typical day in table D.18. To isolate extensive margin responses from intensive margin responses, we estimate eq. (33)

³⁸The percentile-t cluster bootstrap has a better asymptotic refinement than the simple pairs cluster bootstrap, and seems more appropriate in a sample with a limited number of clusters (*i.e.* 242). For more details see Cameron and Miller (2015).

in OLS for the full sample of children, and for the sub-sample of children who have positive hours in school. Our results suggest that changes in school time are driven by extensive margin responses, and not by intensive margin responses. The estimated elasticity of study time with respect to risk ranges between -0.050 and -0.037 when considering the full sample of children. These findings imply that children are less likely to attend school due to risk, while there seems to be little scope for adjusting the time in school at the intensive margin. This finding could be suggestive of an (unmodeled) fixed cost of school attendance, such as travel or material costs, that causes households to withdraw a child from school entirely when faced with higher risk. It could also be a sign for limited flexibility on the side of the teacher, that forces children to withdraw from school entirely rather than reducing hours.

We split the sample by gender and by age group to explore some heterogeneity in effect sizes in table D.19. We find some differences in effect size by gender, with effects being more pronounced for girls than for boys (and being statistically significant for girls only). This is in line with our earlier finding that the effect of rainfall on school time is slightly stronger for girls than for boys, which suggests that girls are exposed to higher risk of dropping out. It could also be interpreted as evidence for differential parental preferences towards boys' schooling.³⁹ We find hardly any differences in effect size by age group. If anything the effect of risk might be slightly more pronounced in older age groups, but again the difference in effect size is very small. We also interact the risk variable with age dummies to explore more nuanced changes by age in figure C.2. Again, the results suggest that the effect of risk is more pronounced in older age groups, however, we find the strongest effects at age 10, thus around the transition to middle school.

3.2.4 Accumulated Schooling

While we have demonstrated that analysis of study time is the best way to assess *ex ante* effects of risk, it remains to be shown that these differences in investment translate into differences in human capital acquisition or at least in some sort of credential. Study time may not effectively raise completed schooling if promotion from grade to grade is not influenced by performance or if reduced study time simply leads to delays in the time of completion. Of course, our structural estimates of the production function estimates indicate how study time translates into future human capital. But those estimates were based on input choice rather than observed outcomes and thus, among other things, may reflect parental views of how human capital acquired rather than the actual production process.

However, given the length of the REDS panel data set, it is possible to track individuals into adulthood and thus to evaluate at least one measure of completed human capital. We focus on adult men observed in 2007-08 round of REDS who were between 10 and 21 during one of the previous survey rounds. In particular, risk observed in village in 1981-82 is used as explanatory

³⁹An earlier draft of this paper used only the 2007-08 round of data and found effects of risk for girls but not for boys. The fact that we now find risk effects for boys and girls (when the data are pooled) might be due to gains in precision (the sample size increased almost three-fold), or due to the fact that households were poorer in 1981-82 and 1998-99, which made it more difficult to protect boys from the adverse effects of risk.

variable for men aged 35-46, and risk observed in village in 1998-99 is used for men aged 18-29.

Table 8 presents estimates of the effects of school-age consumption variability on primary schooling. Note that as these estimates focus on completed schooling they combine *ex ante* and *ex post* effects. The first column uses village fixed effects and the second incorporates household fixed effects and is thus restricted to panel households. The village fixed effects estimates suggest that a one standard deviation increase in the interquartile range of predicted log consumption leads to a 4.1 percent decline in the probability of completing primary school. The household fixed effects estimates suggest a 7.7 percent decline. It is instructive that these are the same order of magnitude of the declines in school attendance and study hours in percentage terms.

3.2.5 Robustness Checks

The identification strategy of his paper largely exploits differential changes over time in the availability of irrigation across villages as approximation for changes in risk.⁴⁰ There are two major reasons why this strategy might be inappropriate. *First*, village level irrigation (or any other variable that correlates with irrigation) might have an effect on schooling decisions that does not work through the risk portfolio of a household. For example, one would expect wealthier villages to exhibit more irrigation and, consequently, lower consumption variability, while also investing more in education. Similarly, differences in village characteristics that co-evolve with the availability of irrigation might affect the supply side of education. *Second*, any endogeneity concerns about the expansion of irrigation will necessarily also concern our risk variable.

We conduct several robustness checks to address these concerns. The first set of robustness checks essentially seeks to rule out the possibility that we are not capturing risk but differences in wealth, or in other household or village-level characteristics. We begin by exploring if household characteristics are systematically associated with our measure of risk. In particular, we focus on household income, consumption expenditures, asset ownership (as measure of household wealth), savings accumulated during the reference period, whether the household is involved in agricultural production, area owned, and maternal labor supply. As reported in table 9, we find some evidence that households exposed to more risk may be slightly poorer. While the point estimates on income, wealth and land ownership are not statistically significant, we do find significant negative effects of risk on consumption expenditures per capita. This finding could be consistent with households saving a higher share of their income if exposed to more risk, but we do not find this to be the case, at least not for households that report positive savings over the reference period. It seems more likely that these households are indeed somewhat poorer, and that consumption expenditures are just more precisely measured than incomes.⁴¹ Finally, households exposed to more risk are no

⁴⁰Remember that the construction of the risk variable from the availability of irrigation and the historical rainfall distribution adds precision to our estimates (we specifically allow an increase in irrigation to have a differential effect on schooling outcomes conditional on the amount of variance in each village's rainfall distribution), but does not introduce any additional source of variation (since we use the same rainfall distribution for all three rounds of interviews).

⁴¹Note that we control for these measures in order to rule out the possibility that differences in wealth are driving our results.

more or less likely to be involved in agricultural production. And we do not find any evidence that mothers in our sample are working more or less hours in the labor market due to risk.

We then explore whether school infrastructure, and village characteristics more generally, are systematically co-evolving with our measure of risk. We focus on educational infrastructure because we have to worry that any schooling response of children is driven by supply side constraints rather than demand side factors. We find no evidence that risk is associated with a lower number of primary, middle, lower secondary or higher secondary schools in a village. We also test if our results are affected by controlling for the presence of primary, middle or secondary schools, and find no change in the magnitude or statistical significance of the effect of risk (*c.f.* table D.20). We take this as evidence that our results are not driven by the differential supply of education infrastructure. We then address the concern raised in Section 3.2.2, namely that villages with a higher share of irrigated land might be villages that have diversified out of agriculture (leaving rain-fed land largely uncultivated), and thus are less exposed to rainfall shocks. These would also be places that invest more in education, if the returns to education are higher in the non-agricultural sector than in the agricultural sector (as is typically assumed). In table D.21, we show that the economic structure in places with more or less risk is not systematically different: neither the share of households that are involved in agriculture, nor the share of income that is generated in agriculture are affected by our risk variable. We also explore if these villages might differ in the presence of factories, or in the number of villagers employed in factories, and find no systematic differences. Finally, we explore if villages with higher risk experience differential access to cooperatives or banks, and again find no evidence that this could be explaining our results.

One caveat with the exercise above is that all these variables might be measured with considerable error, such that we are biased against finding any significant effects. Since we are using a regressor that is calculated for the entire village we are somewhat limited in the amount of tests that we can do. We can, however, explore if any variation within the village (in terms of risk exposure) is pointing in a similar direction, after accounting for village-by-round of interview fixed effects. In order to do this, we first restrict the sample to households that were observed and could be matched in at least two rounds. We then augment equation (30) by an interaction term of rainfall with the log of cash-equivalent savings (*i.e.* savings that can relatively easily be sold in the case of a shock, such as jewelery or cattle). We these estimates, we predict consumption risk at the household level, and re-estimate table 7 with the new risk variable. After controlling for household fixed effects, for household-specific time-trends, and for village-by-round fixed effects, we still find a large negative effect of risk on school attendance, that is statistically significant at the 0.1% level.⁴²

⁴²Results are reported in tables D.22, D.23 and D.24. While the effect of risk on education is clearly negative, and the effect is robust to controlling for the standard controls, as well as for initial conditions multiplied by time fixed effects, we cannot rule out that this household-level risk variable also captures other differences between households. Note that the same risk variable seems to be negatively associated with land ownership, wealth, consumption expenditures and with the probability of being engaged in agricultural production. Yet, it is reassuring to see that we find similar associations within villages as we find across villages. We also exploit the panel data to control for household fixed effects in the original specification, *i.e.* with village-level risk, and find that our results are robust to using this subset of the data and to controlling for household fixed effects (see table D.25).

Taken together, these results are consistent with the notion that we are indeed capturing *ex ante* effects of variance, rather than some other underlying household or village characteristics that correlate with our risk variable.

The second set of robustness checks seeks to address potential sources of endogeneity in the expansion of irrigation. Endogeneity could arise from differences in initial levels of risk that might put villages on very different paths of agricultural production, wealth accumulation, and investments in education. To address the concern that underlying trends are driving the contemporaneous effect of risk on schooling, we conduct robustness checks in which we control for (non-linear) time trends that vary with baseline levels of village average irrigation, education of the household head, income, wealth, and land ownership, and our results are robust in four out of five specifications (*c.f.* table D.26). We also slice our data to perform placebo tests. In particular, we regress current schooling (in rounds 1981-82 and 1998-99) on future risk in table D.27, and find that these are unrelated.

Endogeneity could also arise from unobserved shocks that are not captured by rainfall nor by state-by-round (of interview) fixed effects, and that simultaneously affect irrigation availability (and therewith risk) at the village level and investments in education at the child level. Lagged negative shocks, for example, could slow down investments in irrigation, and — through wealth effects — also negatively affect investments in education. While we control for contemporaneous and lagged rainfall in all our specifications, we might not be capturing the full effect of these shocks. We therefore also explore if our results are robust to flexibly controlling for lagged (3 years) rainfall and temperature. As we show in table D.28, this is indeed the case.⁴³ It remains to be said, however, that we cannot fully rule out the possibility that other unobserved shocks simultaneously affect irrigation and schooling outcomes in ways that could bias our results.

3.3 Alternative Explanations

The results presented so far are consistent with the existence of dynamic complementarities in the human capital production function. However, other mechanisms could also result in a negative association between risk and human capital investments, leading us to erroneously attribute this effect to dynamic complementarities. We analyze each of these explanations in detail to show that they are unlikely to produce the results we are finding.

The first alternative explanation to the negative effect of risk on child schooling could be a precautionary savings motive. Study time might be lower in high-risk villages, because parents are less willing to spend their limited resources on education if concerned about next season's income and rather save more, or because parents are working more in order to accumulate savings, and therefore spend less time on home production, which has to be taken on by the children of the family. *Ceteris paribus*, this would increase the time children spend on household chores or agricultural production, and reduce study time.⁴⁴ However, the available empirical evidence does

⁴³We cannot use more than three lags since the weather data we are using only goes back to 1979. We also show that neither lagged, nor current or future rainfall is systematically related to our measure of consumption risk (*c.f.* table D.29).

⁴⁴That households tend to save more in the presence of uninsured risk is by now well established. However, it is

not suggest that this is what is driving our results. As discussed previously, we do not find that households save more year-to-year nor that women work more in areas with higher risk.⁴⁵ While year-to-year saving does not necessarily capture the mechanism we outlined above, the simulations with savings showed clearly that the negative effect of risk would vary substantially over the year (in particular *ex ante* and *ex post* to shocks) in a precautionary savings world, while it is mostly constant over time if driven by dynamic complementarities. Empirically, we find the effect of risk to be surprisingly constant over the three seasons of the year, with the most pronounced negative effects *de facto* arising in season one (October/November), thus *ex post* to the main shock (*i.e.* monsoon), see table D.30. This result is consistent with the dynamic complementarity mechanism, yet less so with a savings motive.

And finally, liquidity (or savings) concerns should generally be less of concern in contexts like India, where schooling is generally for free. Obviously, one-time expenditures for school material or uniforms could be just as prohibitively expensive, but this seems not to be driving our results. In table D.31, we allow the effect of risk on school attendance to vary by quartiles of consumption, income, and wealth, and contrast the results with the average expenditures on education (per child), and average income (per capita) in these subsamples. This exercise reveals that the effect of risk is generally most pronounced in the poorest quartile. Yet, we also find effects of similar magnitudes in the second and third quartiles of the consumption, income or wealth distributions, and in these groups the share of income that is spent on education is substantially lower than in the poorest quartile.

The second alternative explanation is that risk could negatively affect returns to human capital, either through the economic structure or due to a reduced propensity to adopt technological innovations, causing parents to invest less in the education of their children. We suggest two strategies to address this concern. First, we test directly whether the wage return to education varies with risk. Since we only observe wage data in the 2007-08 round of data, we restrict this exercise to an analysis of cross-sectional associations. We regress log daily wages (averaged over all observed activities for the individual) on risk interacted with the level of education, and on a number of individual-level characteristics (results are reported in table D.32). We do not find any evidence that risk in fact reduces the returns to education either for men or for women.⁴⁶ To the contrary, if anything evidence suggests that returns to education are positively associated with risk, which suggests that diversification out of agriculture is higher in more risk prone areas or that risk prone areas rely on production technologies that require higher levels of human capital. Second, for girls

unclear to which extent this affects labor supply decisions. Typically, the literature has looked at foregone current consumption for the sake of higher (and more secure) future consumption (see e.g. Zeldes, 1989; Deaton, 1991). On the other side, one might think that households would tend to work harder in order to accumulate higher savings. A positive labor supply response to risk in the context of India was found by Rose (2001).

⁴⁵Table 9 shows the effect of risk on savings and on maternal time allocation. Unfortunately, the same data is not available for fathers, because the time-use questions are included in the demographic questionnaire (that is administered to women exclusively).

⁴⁶Of course, this approach does not account for other possible consequences of risk such as the returns to entrepreneurship or the returns to migration. But it is at least consistent with the idea that our effect of interest is not driven by the return to schooling.

at least, who are likely to leave the village following marriage, we can examine if the negative effect of risk is driven by the level of risk in the village of birth or by the level of risk in the relevant marriage market. The concern would be that risk is likely correlated across space, such that we would be confounding the level of risk in the the village of origin with the level of risk in the marriage market. We explore this idea by adapting a strategy from Foster and Rosenzweig (2001). In particular, that paper uses a radius of 67km to define the marriage market (the 90th percentile of the marriage distance). It then distinguishes the returns to human capital for boys (that in the village) and the returns to human capital for girls (among sample villages in the broader marriage market). In this paper, we control for the average risk among sample villages within 67km of each village of interest. We exclude the village of interest from this mean and drop those villages in which there are no other sample villages in the marriage market. As reported in table D.33, the coefficient on marriage-market risk is close to zero and its inclusion does not substantially affect the point estimate on the variable of interest. Thus, as posited in our model, the risk estimates appear to reflect the consequences of risk during childhood rather than adulthood.

A third alternative explanation could be differences in fertility — or (gender specific) child survival rates — that are due to risk. If households exposed to higher risk have more children, they might consequently invest less in each of them. Likewise, a higher gender-imbalance might increase (or reduce) investments in girls (and boys) due to selection effects or due to changes in marriage-market returns. Yet, as reported in table D.34, we find no evidence that these mechanisms are empirically meaningful in our data.

4 Simulating the Effect of the NREGS on Consumption Variability and Schooling Investments

Given the magnitude of the effects of risk observed above, it seems worthwhile to explore potential policy tools to mediate these. Obviously, any policy that helps farmers insure against agricultural production risk, could be a viable option. But as Mobarak and Rosenzweig (2013) pointed out, providing insurance to farmers might actually increase overall risk in village economies as farmers become more risk taking in their production decisions. This would then be particularly harmful for the poorest households with no own land and no access to agricultural insurance. An alternative could be an workfare program, such as the Indian National Rural Employment Guarantee Scheme, that was introduced in rural India in 2006.⁴⁷

A workfare program like the NREGS can affect schooling decisions via two mechanisms: first, the program could provide employment at higher wages than casual agricultural wages, which might affect total labor supply and the amount of labor supplied to the private sector (with potential general equilibrium effects on wages, as documented by Imbert and Papp (2015)). Second, such a program could reduce consumption risk in rural areas, by providing employment at a fixed wage

⁴⁷The National Rural Employment Guarantee Act is India’s flagship anti-poverty program; it entitles every household in rural India to a maximum of 100 days of employment per year at state minimum wages.

independently of rainfall shocks. Assuming income effects of rainfall shocks dominate in the time allocation of adults, the fact that wages do not fall during periods of need is likely to reduce the magnitude of the labor supply response of households relative to the case in which there are negative general equilibrium effects on wages. This effect should reduce the extent of the response among children *ex post* to shocks. This lower response, in turn, increases the return to first-period school time.

To understand the effectiveness of a workfare program in mediating the negative effects on income risk on school attendance, we explore the follow-up data to the REDS collected between 2014 and 2016, and simulate the effect of the NREGS.⁴⁸ This simulation is expected to shed light on the importance of the second mechanism and is intended to complement existing empirical evidence on the short-term impact of the NREGS on school enrollment, which estimates that the NREGS decreased enrollment by 1-3.5 percentage points (Shah and Steinberg, 2021).⁴⁹

Using the data from 1982 to 2016, we estimate the extent to which the presence of the NREGS, or more specifically, the amount of employment generated by the NREGS per year, mediates the effect of rainfall on household consumption. Formally, we re-estimate eq. (30), but now add employment per capita (in person-days) generated in a given village by the NREGS and its interaction with rainfall to the estimation. We find that a one unit increase in employment per capita, reduces the effect of rainfall on consumption by 4.2 percentage points (see table D.36).

We use these results to predict the standard deviation of predicted log consumption at different levels of NREGS employment per capita. At the mean of irrigation and the minimum of NREGS employment (*i.e.* zero employment), an increase in NREGS employment by 2.24 days per capita would reduce the standard deviation of log consumption by 34%.⁵⁰ A reduction of risk by this magnitude would increase the probability of attending school by 1.0 percentage points, according to estimates presented in table 7, column (4).

Our simulation thus suggests that an employment guarantee such as the NREGS could have positive effects on child school attendance by reducing consumption risk. However, as documented by previous literature the NREGS had substantial wage effects, which seem to explain the negative effect on school enrollment documented by Shah and Steinberg (2021) and Li and Sekhri (2020). If the wage effect outweighs the risk reducing effect, or in other words, if the risk reducing effect is not high enough to offset the negative effect of increasing wages, then an employment guarantee such

⁴⁸Summary statistics of the SEPRI data are shown in table D.35. More details about these data are available in appendix B.3.

⁴⁹We expect that the risk mitigating effect of the NREGS is internalized by households only after a certain period of time. Therefore this effect cannot be captured by impact evaluations that explore the sequenced phase-in of the NREGS and which can thus only provide information about the immediate effects of the program. This simulation is intended to provide insights about potential long-run effects associated with public works programs such as the NREGS.

⁵⁰We use 2.24 days of employment to make our results comparable to previous work. Imbert and Papp (2015) and Shah and Steinberg (2021) use NSS data collected in 2004-05 (pre) and 2007-08 (post). In 2007-08, the NREGS had been implemented in 330 (out of 625) rural districts of India, covering a population of 642 Million (projections for March 2007). To get the population estimates, we apply the India-wide projected population growth rate to the Census 2001 population count of these districts. That year, the total amount of employment-days created within the NREGS was 1,435.9 Million (Ministry of Rural Development, 2012).

as the NREGS decreases school attendance not only in the first years of program implementation, but permanently.

5 Conclusions

This paper analyzes the joint effect of risk and dynamic complementarity on investments in human capital, with a focus on the study time of children in rural India. The model developed in this paper shows that parents respond to income variance by reducing schooling investments *ex ante* if there is dynamic complementarity in the demand for child human capital, which is a function of the curvature in the parental demand for child human capital and of the perceived dynamic complementarity in the production of human capital.

Using indirect inference, we then estimate the structural parameters of the model exploiting — in the auxiliary model — a surprisingly robust relationship in study times across seasons observed in the data. Our results suggest that parents perceive a strong degree of dynamic complementarity in child schooling, and exhibit substantial curvature in the demand for child human capital. The estimated elasticity of intertemporal substitution in study times is 0.29 which is considerably smaller than what has been found using other approaches. One possible reason is that because we focus primarily on inputs, our estimates may reflect differences between parental perceptions about the production function and the ‘true’ production function. While our results on accumulated human capital are of the same order as are our study time results, suggesting this may not be an issue, it is worth noting that, if present, such a gap would be important to understand. Any policy intervention needs to account for the choices that parents make as well as the actual accumulation process in order to generate desired effects. We would also note that our approach offers a way of identifying parameters of interest even in contexts in which skills are measured with considerable error.

We simulate the lifetime effect of variance on child schooling and decompose the overall effect into *ex ante* and *ex post* components. We find that risk is expected to substantially reduce schooling in this context and that a large share of this effect can be explained by an *ex ante* response. Interestingly, because parents underinvest from the start the extent of *ex post* variation explains only about half of the total effect.

Of course, the results of this simulation are sensitive to the assumption of i.i.d. shocks across years. The results would likely be quite different if there were a substantial probability of a permanent reduction, say, in the variance of income realizations at some later grade. However, we believe the current model of the income process is a reasonable characterization of the income processes faced by rural households in India over the period covered by our data. We also recognize that there are other permanent shocks such as newly revealed information about the ability of a child or about the long-term returns to human capital that could have additional — and potentially more pronounced — effects.

We test the model’s predictions by exploiting village-level variation over time in the use of

irrigation in agriculture, and find strong evidence for a negative effect of risk on the probability that children attend school.

These findings contribute to a better understanding of the consequences of dynamic complementarity in the production of human capital. Much of the literature to date on dynamic complementarity and schooling has emphasized the importance of early child interventions for eventual schooling outcomes. Our results suggest that — in settings characterized by risk — such interventions may not automatically transfer into higher accumulated capital in the longer run unless this exposure to risk is addressed directly. Even though the immediate human-capital returns to subsequent investment will be higher, the longer term expected returns may still be low due to concerns about the ability to follow through on those investments.

While direct schooling subsidies to at-risk children in later years may be the most direct means of addressing this issue, other forms of insurance for households can play an important role. We simulate the effect of the National Rural Employment Guarantee Scheme on the child schooling. We estimate the extent to which NREGS reduces variability in consumption and use this reduction to simulate the program's effect on child study time. We find that the program would increase school attendance by 1 percentage point, if it were holding the level of wages fixed. Of course, if this result also raises child wages then the salutary effect of risk reduction may be offset by the rising opportunity cost of schooling.

While our study focuses in particular on the problem of risk in low income rural areas, we believe our results may have some relevance in other settings. Students, in any context, may underinvest in studying at the start of the semester if they think there is a chance that they will not be able to maintain that level of investment throughout the semester and thus do well on the exam. Similarly, students engaged in college preparation may be importantly affected by the perceived risk that they will not be able to fully finance the years of study needed to complete a degree. These questions await further research.

References

- Agostinelli, F. and M. Wiswall (2020). Estimating the Technology of Children’s Skill Formation. Working Paper 22442, National Bureau of Economic Research.
- Ambika, A. K., B. Wardlow, and V. Mishra (2016). Remotely sensed high resolution irrigated area mapping in India for 2000 to 2015. *Scientific data* 3, 160118.
- Andrews, I., M. Gentzkow, and J. M. Shapiro (2017, 06). Measuring the Sensitivity of Parameter Estimates to Estimation Moments. *The Quarterly Journal of Economics* 132(4), 1553–92.
- Atkin, D. (2016). Endogenous Skill Acquisition and Export Manufacturing in Mexico. *American Economic Review* 106(8), 2046–85.
- Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2020). Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia. *American Economic Review* 110(1), 48–85.
- Attanasio, O., C. Meghir, and E. Nix (2020). Human Capital Development and Parental Investment in India. *The Review of Economic Studies* 87(6), 2511–2541.
- Beegle, K., R. H. Dehejia, and R. Gatti (2006). Child labor and agricultural shocks. *Journal of Development Economics* 81(1), 80–96.
- Bjorkman-Nyqvist, M. (2013). Income shocks and gender gaps in education: Evidence from Uganda. *Journal of Development Economics* 105, 237 – 253.
- Cameron, C. A. and D. L. Miller (2015). A practitioner’s guide to cluster-robust inference. *Journal of Human Resources* 50(2), 317–372.
- Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019). *Inference with arbitrary clustering*. IZA Discussion Paper.
- Colmer, J. (2021a). Rainfall variability, child labor, and human capital accumulation in rural ethiopia. *American Journal of Agricultural Economics* 103(3), 858–877.
- Colmer, J. (2021b, October). Temperature, labor reallocation, and industrial production: Evidence from india. *American Economic Journal: Applied Economics* 13(4), 101–24.
- Conley, T. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics* 92(1), 1–45.
- Copernicus Climate Change Service (2017). *ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate*. Copernicus Climate Change Service Climate Data Store (CDS).
- Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review* 97(2), 31–47.
- Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78(3), 883–931.
- Das, J., A. Singh, and A. Yi Chang (2022). Test scores and educational opportunities: Panel evidence from five low- and middle-income countries. *Journal of Public Economics* 206, 104570.

- Deaton, A. (1991). Saving and liquidity constraints. *Econometrica* 59(5), 1221–1248.
- Dercon, S. and L. Christiaensen (2011). Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics* 96(2), 159–173.
- Directorate of Economics and Statistics (2019). *Agricultural Statistics at a Glance 2018*. Ministry of Agriculture & Farmers Welfare. Government of India.
- Duflo, E. and R. Pande (2007). Dams. *The Quarterly Journal of Economics* 122(2), 601–646.
- Duryea, S., D. Lam, and D. Levison (2007). Effects of economic shocks on children’s employment and schooling in Brazil. *Journal of Development Economics* 84(1), 188–214.
- Fitzsimons, E. (2007). The Effects of Risk on Education in Indonesia. *Economic Development and Cultural Change* 56(1), 1–25.
- Foster, A. and M. Rosenzweig (2001). Missing women, the marriage market, and economic growth. Manuscript, Brown University.
- Foster, A. D. and M. R. Rosenzweig (1996). Technical change and human-capital returns and investments: Evidence from the green revolution. *American Economic Review* 86(4), 931–53.
- Glewwe, P. and H. G. Jacoby (2004). Economic growth and the demand for education: is there a wealth effect? *Journal of Development Economics* 74(1), 33–51.
- Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect Inference. *Journal of Applied Econometrics* 8, S85–S118.
- Gubert, F. and A.-S. Robilliard (2008). Risk and Schooling Decisions in Rural Madagascar: A Panel Data-Analysis. *Journal of African Economies* 17(2), 207–238.
- Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences* 107(35), 15367–15372.
- Ilahi, N. (2000). *The intra-household allocation of time and tasks: what have we learnt from the empirical literature?* World Bank, Development Research Group/Poverty Reduction and Economic Management Network.
- Imbert, C. and J. Papp (2015). Labor Market Effects of Social Programs: Evidence from India’s Employment Guarantee. *American Economic Journal: Applied Economics* 7(2), 233–63.
- Jacoby, H. G. and E. Skoufias (1997). Risk, financial markets, and human capital in a developing country. *Review of Economic Studies* 64(3), 311–335.
- Jayachandran, S. (2006). Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy* 114(3), 538–575.
- Jensen, R. (2000). Agricultural volatility and investments in children. *American Economic Review* 90(2), 399–404.
- Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics* 125(2), 515–548.

- Jensen, R. (2012). Do Labor Market Opportunities Affect Young Women’s Work and Family Decisions? Experimental Evidence from India. *Quarterly Journal of Economics* 127(2), 753–792.
- Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review* 109(10), 3585–3616.
- Kazianga, H. (2012). Income Risk and Household Schooling Decisions in Burkina Faso. *World Development* 40(8), 1647 – 1662.
- Kochar, A. (1999). Smoothing consumption by smoothing income: hours-of-work responses to idiosyncratic agricultural shocks in rural India. *Review of Economics and Statistics* 81(1), 50–61.
- Li, T. and S. Sekhri (2020). The Spillovers of Employment Guarantee Programs on Child Labor and Education. *The World Bank Economic Review* 34(1), 164–178.
- Ministry of Rural Development (2012). *MGNREGA Sameeksha: An Anthology of Research Studies on the Mahatma Gandhi National Rural Employment Guarantee Act, 2005. 2006-2012*. New Delhi: Orient BlackSwan.
- Mobarak, A. M. and M. R. Rosenzweig (2013). Informal risk sharing, index insurance, and risk taking in developing countries. *American Economic Review* 103(3), 375–80.
- Mondal, P., M. Jain, R. S. DeFries, G. L. Galford, and C. Small (2015). Sensitivity of crop cover to climate variability: Insights from two Indian agro-ecoregions. *Journal of Environmental Management* 148, 21 – 30. Land Cover/Land Use Change (LC/LUC) and Environmental Impacts in South Asia.
- Oster, E. and B. M. Steinberg (2013). Do it service centers promote school enrollment? evidence from india. *Journal of Development Economics* 104(Supplement C), 123–35.
- Rose, E. (2001). Ex ante and ex post labor supply response to risk in a low-income area. *Journal of Development Economics* 64(2), 371–388.
- Saha, A. and J. Stroud (1994). A household model of on-farm storage under price risk. *American Journal of Agricultural Economics* 76(3), 522–534.
- Shah, M. and B. M. Steinberg (2017). Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *Journal of Political Economy* 125(2), 527–561.
- Shah, M. and B. M. Steinberg (2021). Workfare and Human Capital Investment: Evidence from India. *Journal of Human Resources* 56(2), 380–405.
- Skoufias, E. (1993). Labor market opportunities and intrafamily time allocation in rural households in South Asia. *Journal of Development Economics* 40(2), 277–310.
- Skoufias, E. and S. W. Parker (2006). Job loss and family adjustments in work and schooling during the Mexican peso crisis. *Journal of Population Economics* 19(1), 163–181.
- von Hippel, P. T. and C. Hamrock (2019). Do test score gaps grow before, during, or between the school years? measurement artifacts and what we can know in spite of them. *Sociological Science* 6, 43–80.

Waldman, K. B., S. Giroux, J. P. Blekking, K. Baylis, and T. P. Evans (2020, February). Small-holder food storage dynamics and resilience. *Food Security: The Science, Sociology and Economics of Food Production and Access to Food* 12(1), 7–20.

Zeldes, S. P. (1989). Optimal consumption with stochastic income: Deviations from certainty equivalence. *Quarterly Journal of Economics* 104(2), 275–298.

Figures

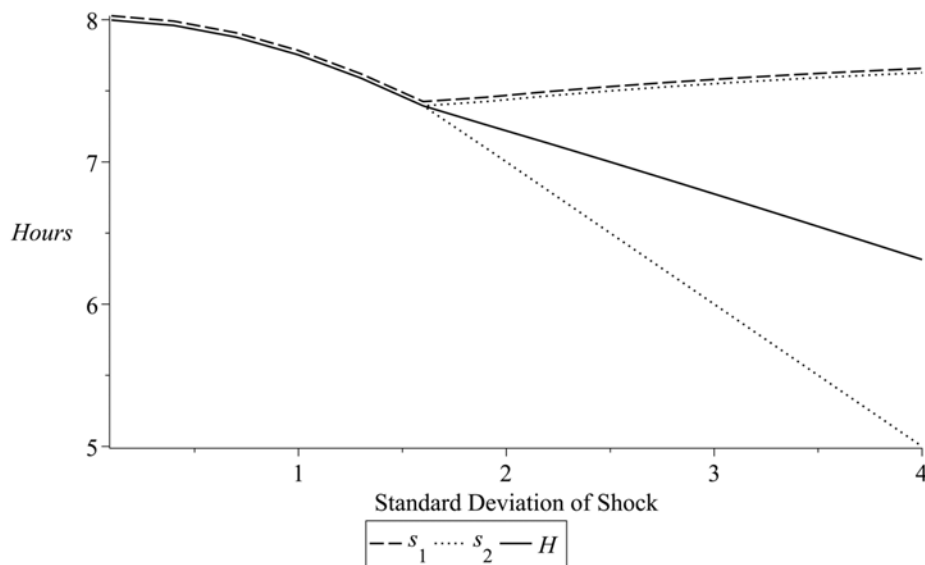


Figure 1: Risk, Study Time and Expected Human Capital

Notes: This figure presents numerical solutions to the optimal choice of s_1 , s_2 by state and $\mathbb{E}[H_3]$ by the standard deviation of the shock ($\sigma_\theta = r$), for a two-period model as described by eq. (5) with perfect complements and linear $g(H)$. Assumed parameter values are $y_1 = 10$, $y_2 = 10$, and $\zeta = 1$.

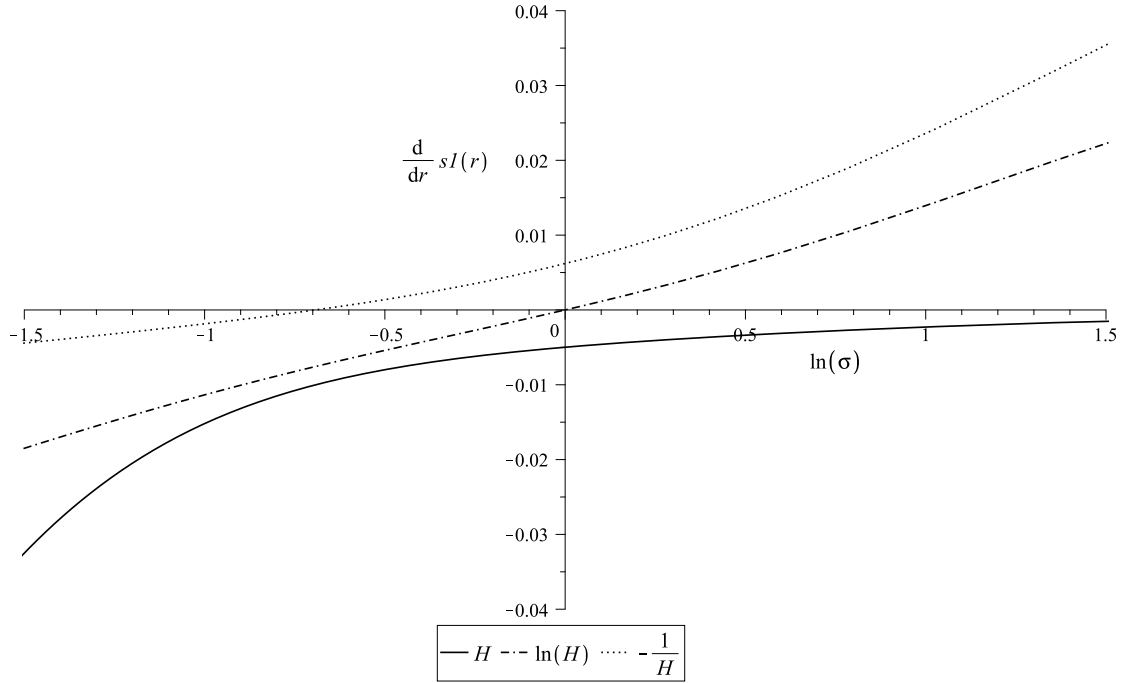


Figure 2: Numeric Approximation of the Effect of Risk on First Period Schooling

Notes: This figure presents numerical solutions to the derivative of s_1 wrt to the standard deviation of the shock ($\sigma_\theta = r$), *i.e.* the *ex ante* effect of an increase in variance, for a two-period model as described by eq. (5) and for various values of $\sigma = 1/(1 - \rho)$ and three specifications of $g(H)$: $g(H) = H$, $g(H) = \ln(H)$ and $g(H) = -1/H$. Assumed parameter values are $y_1 = 10$, $y_2 = 10$, $\nu = 1/2$, $r = 1$ and $\zeta = 1$.

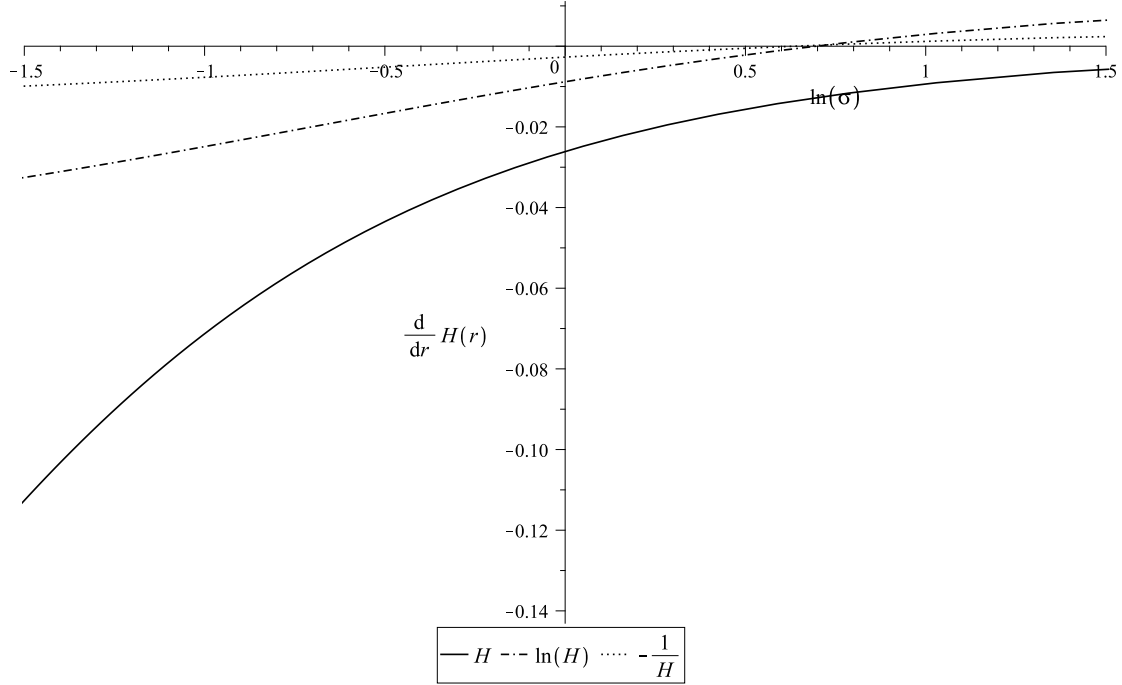


Figure 3: Numeric Approximation of the Effect of Risk on the Expected Value of Human Capital

Notes: This figure presents numerical solutions to the derivative of $\mathbb{E}(H_3)$ wrt to the standard deviation of the shock ($\sigma_\theta = r$), *i.e.* the combined *ex ante* and *ex post* effect of an increase in variance, for a two-period model as described by eq. (5) and for various values of $\sigma = 1/(1 - \rho)$ and three specifications of $g(H)$: $g(H) = H$, $g(H) = \ln(H)$ and $g(H) = -1/H$. Assumed parameter values are $y_1 = 10$, $y_2 = 10$, $\nu = 1/2$, $r = 1$ and $\zeta = 1$.

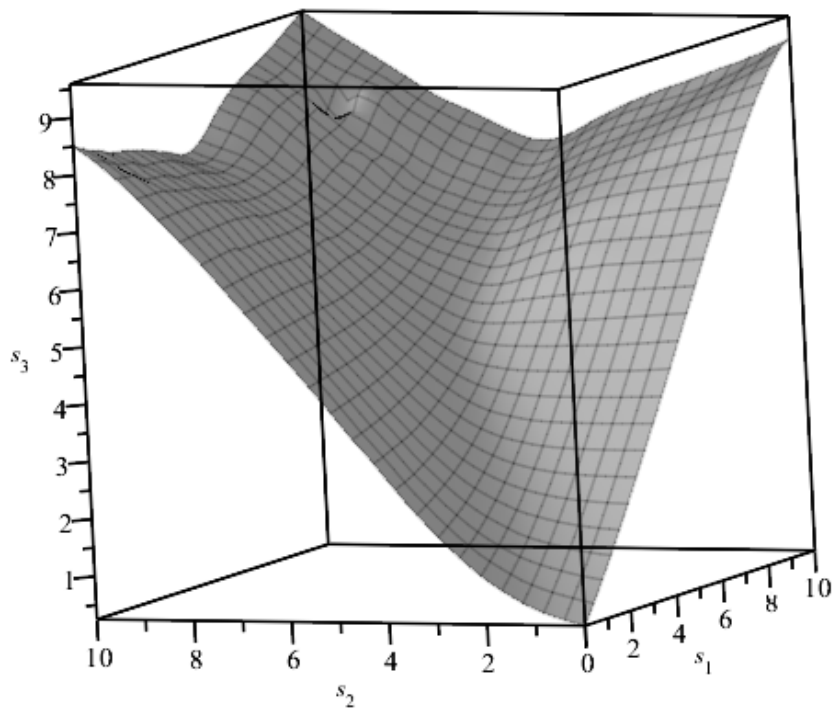


Figure 4: Lowess-smoothed Study Time by Season
Notes: This figure plots the fitted values of a lowess-regression (bandwidth = 0.3, degree = 1) of s_3 on s_1 and s_2 .

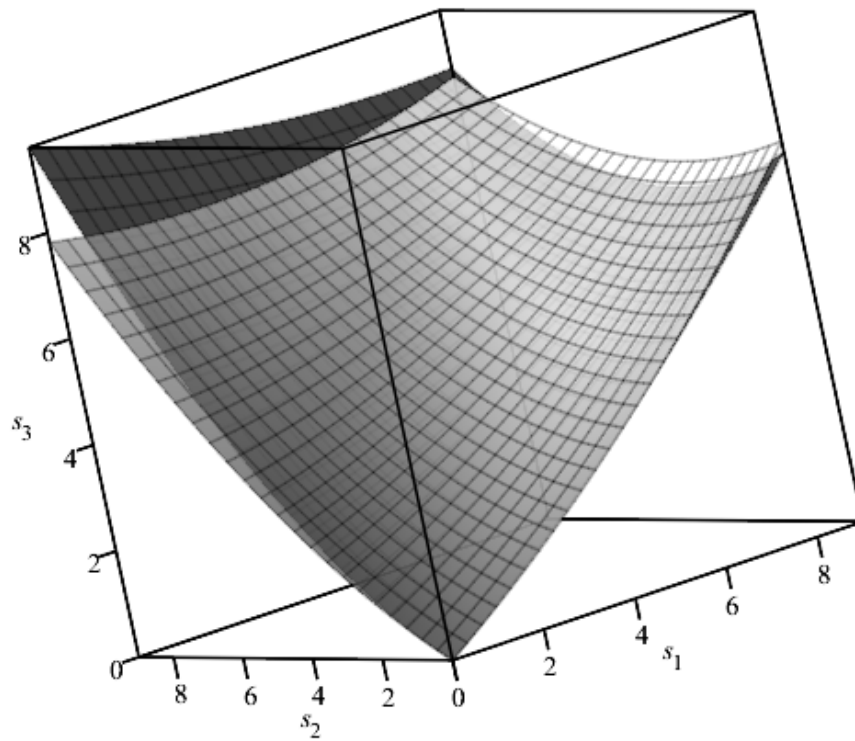


Figure 5: Study Time by Season: Model and Data Estimates

Notes: This figure plots the estimated quadratic relationship between s_3 and s_1 and s_2 from the simulated data (black) and actual data (grey). The parameter estimates ($\alpha_1 - \alpha_5$) are given by table 2.

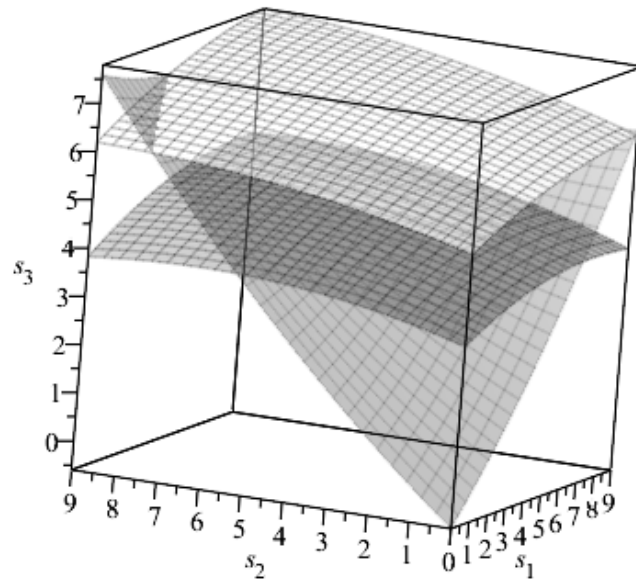
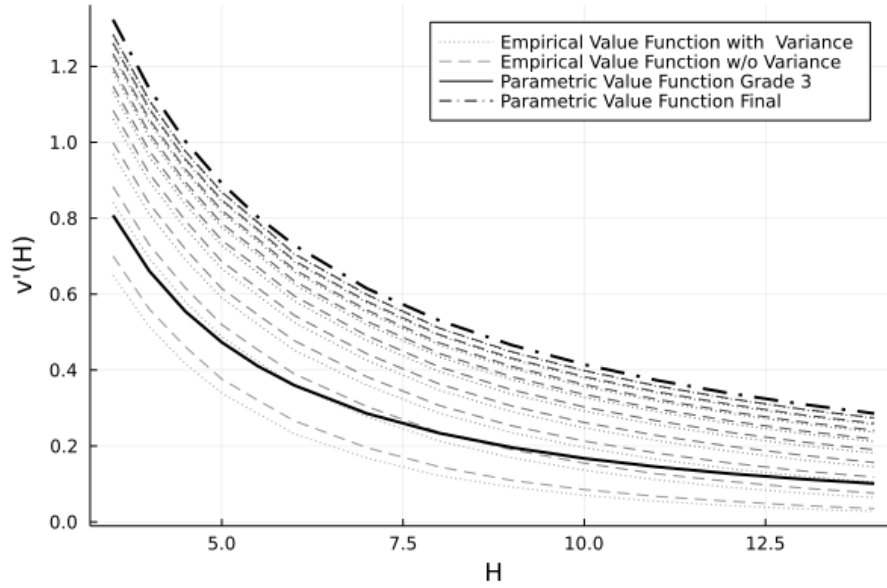


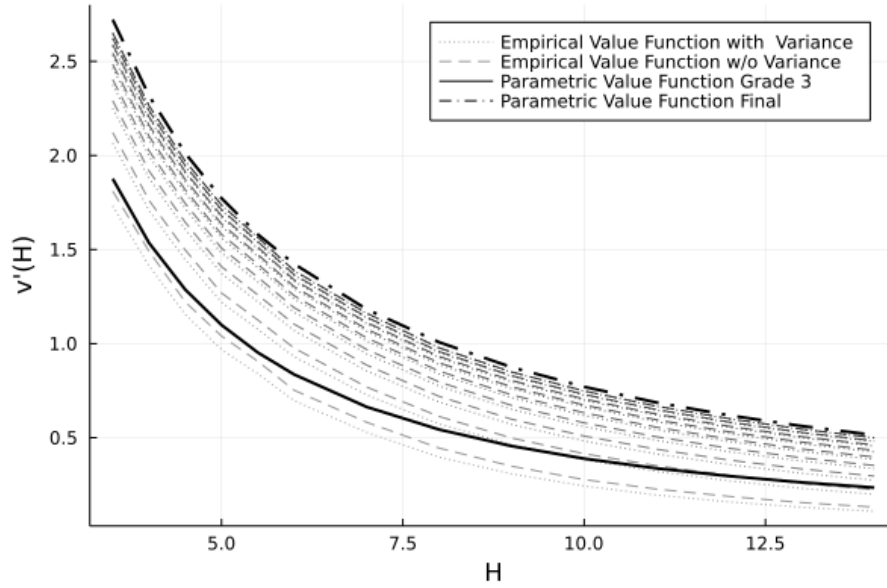
Figure 6: Study Time by Season and by Parental Type

Notes: This figure plots the estimated quadratic relationship between s_3 and s_1 and s_2 for the high-type parent (white), for the low-type parent (dark grey), and for both types combined (grey).

Figure 7: First Derivatives of Parametric and of Empirical Value Functions at Various Grades



(a) Low type



(b) High type

Notes: This graph shows the first derivatives w.r.t. H of the estimated (grade 3) parametric value function ($\zeta^i g(H_7)$) and of the adjusted grade 9 parametric value function ($\zeta^i g(H_{28})$), as well as of the derived empirical value functions from grade 2 (light grey) through grade 9 (dark grey). Dotted lines are no risk cases, dash-dot lines are the cases with risk. Panel (a) shows the derivatives for the low type ($\phi = 1.105$) and panel (b) for the high type ($\phi = 1.202$).

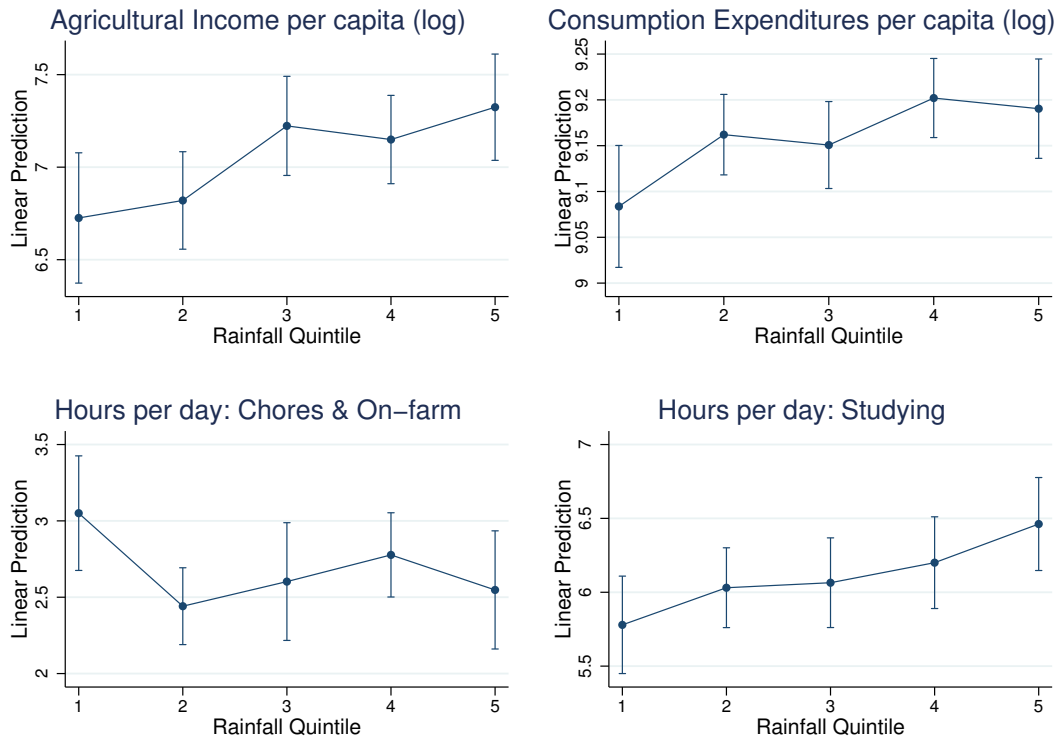


Figure 8: Effect of Rainfall Shocks

Notes: The figures plot linear predictions and 95% confidence intervals from regressions of each outcome variable on dummies for each quintile of the rainfall distribution. Each quintile dummy equals 1 if total annual rainfall in the current agricultural year fell within the given quintile of the village's usual rainfall distribution and equals 0 otherwise. Each regression controls for state-by-round of interview fixed effects, for the share of area that is irrigated within a village, and for lagged rainfall (in log mms) in the past 3 years (2 years in Panel A). Panels C and D additionally control for age-by-gender and season-by-round of interview fixed effects. Standard errors are corrected to allow for clustering at the level of the grid cell by round of interview.

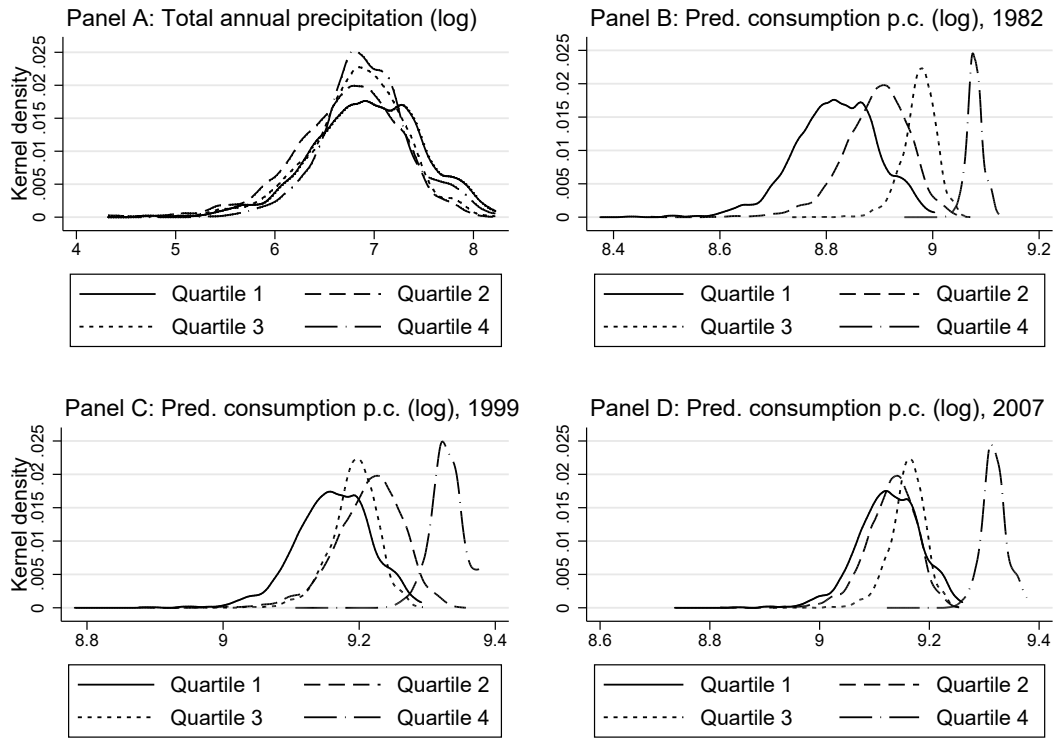


Figure 9: Distribution of Rainfall and Predicted Consumption by Quartiles of Baseline Irrigation
 Notes: This figure plots the kernel densities of total annual rainfall (log), and predicted consumption expenditures per capita (log) in REDS villages in the rounds 1981-82, 1998-99 and 2007-08. In all four graphs, the REDS villages are split into quartiles of irrigation share in 1981-82.

Tables

Table 1: Study Time across Seasons

| | (1) | (2) | (3) | (4) | (5) |
|--|----------------------|----------------------|----------------------|----------------------|----------------------|
| Study time, season 1 | 0.573*** (0.040) | 0.571*** (0.039) | 0.571*** (0.039) | 0.576*** (0.039) | 0.568*** (0.040) |
| Study time, season 1 (square) | 0.024*** (0.004) | 0.024*** (0.004) | 0.024*** (0.004) | 0.024*** (0.004) | 0.025*** (0.004) |
| Study time, season 2 | 0.453*** (0.041) | 0.471*** (0.040) | 0.471*** (0.040) | 0.473*** (0.040) | 0.474*** (0.041) |
| Study time, season 2 (square) | 0.045*** (0.004) | 0.044*** (0.004) | 0.044*** (0.004) | 0.044*** (0.004) | 0.044*** (0.004) |
| Study time, season 1 \times season 2 | -0.080*** (0.005) | -0.081*** (0.005) | -0.081*** (0.005) | -0.081*** (0.005) | -0.082*** (0.005) |
| Rainfall controls | N | Y | Y | Y | Y |
| Income, wealth, area, hh size, risk | N | N | Y | Y | Y |
| State-by-round FE | N | N | N | Y | Y |
| Village-by-round FE | N | N | N | N | Y |
| Observations | 17186 | 17186 | 17186 | 17186 | 17186 |
| Adjusted R^2 | 0.796 | 0.806 | 0.806 | 0.808 | 0.816 |

Notes: The dependent variable is study time in season 3. Each regression controls for village, round of interview, and age-by-gender fixed effects. Rainfall controls are current and lagged rainfall in past 3 years (in log mms). Income, wealth and area are measured in per capita terms (and are in logs). Risk is the interquartile range of predicted log consumption. Sample is restricted to 1998-99 and 2007-08 rounds of interview. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Match of the Auxiliary Model

| Auxiliary model parameter | Symbol | Target | Simulated | Difference |
|---------------------------------------|----------------|--------|-----------|------------|
| Effect of s_1 on s_3 | α_1 | 0.568 | 0.596 | 0.028 |
| Effect of s_2 on s_3 | α_2 | 0.474 | 0.483 | 0.009 |
| Effect of $(s_1)^2$ on s_3 | α_3 | 0.025 | 0.020 | -0.006 |
| Effect of $(s_2)^2$ on s_3 | α_4 | 0.044 | 0.066 | 0.022 |
| Effect of $(s_1 \times s_2)$ on s_3 | α_5 | -0.082 | -0.100 | -0.018 |
| Effect of s_1 on s_2 | γ_1 | 1.003 | 1.237 | 0.234 |
| Effect of $(s_1)^2$ on s_2 | γ_2 | -0.014 | -0.028 | -0.014 |
| Mean of s_1 | μ_{s_1} | 6.860 | 7.065 | 0.205 |
| Standard deviation of s_1 | σ_{s_1} | 2.884 | 3.179 | 0.294 |
| Mean of s_2 | μ_{s_2} | 6.900 | 6.972 | 0.071 |
| Standard deviation of s_2 | σ_{s_2} | 2.879 | 3.095 | 0.215 |
| Mean of s_3 | μ_{s_3} | 6.608 | 6.810 | 0.202 |
| Standard deviation of s_3 | σ_{s_3} | 2.905 | 3.371 | 0.465 |
| Correlation $(y_2 - s_2, y_1 - s_1)$ | $r_{1,2}$ | 0.260 | 0.307 | 0.047 |
| Correlation $(y_3 - s_3, y_2 - s_2)$ | $r_{2,3}$ | 0.286 | 0.459 | 0.172 |
| Correlation $(y_3 - s_3, y_1 - s_1)$ | $r_{1,3}$ | -0.358 | -0.217 | 0.141 |

Notes: Target parameter estimates α_1 to α_5 are given by table 1, col. (5), and γ_1 and γ_2 by table D.7, col (5).

Table 3: Estimated Structural Parameters and Sensitivity Matrix

| Symbol | ζ^L | $\Delta\zeta$ | ϕ | σ | δ | y^L | Δy | ω_1 | ω_2 | ω_3 |
|--------------------------------|-----------|---------------|--------|----------|----------|--------|------------|------------|------------|------------|
| Panel A: Structural Parameters | | | | | | | | | | |
| Estimate | 5.284 | 6.992 | 1.499 | 0.290 | 1.357 | 14.020 | 26.088 | 0.072 | 0.087 | -0.023 |
| | 0.831 | 1.408 | 0.074 | 0.007 | 0.015 | 0.198 | 0.826 | 0.008 | 0.005 | 0.006 |
| Panel B: Sensitivity Matrix | | | | | | | | | | |
| α_1 | -0.101 | -0.129 | -0.025 | -0.006 | 0.002 | 0.015 | 0.004 | -0.002 | -0.010 | -0.021 |
| α_2 | -0.148 | -0.183 | -0.035 | -0.001 | -0.006 | 0.024 | 0.009 | -0.031 | 0.009 | -0.081 |
| α_3 | -0.171 | -0.206 | -0.045 | -0.017 | 0.000 | 0.019 | -0.003 | -0.002 | -0.015 | -0.024 |
| α_4 | -0.144 | -0.158 | -0.032 | -0.025 | -0.001 | 0.021 | 0.015 | -0.054 | 0.006 | -0.131 |
| α_5 | -0.163 | -0.174 | -0.038 | -0.031 | 0.001 | 0.022 | 0.012 | -0.041 | -0.015 | -0.085 |
| γ_1 | -0.090 | -0.093 | -0.039 | -0.020 | -0.006 | -0.015 | -0.043 | 0.028 | 0.035 | -0.071 |
| γ_2 | -0.113 | -0.123 | -0.044 | -0.016 | -0.006 | -0.010 | -0.041 | 0.026 | 0.032 | -0.078 |
| μ_{s_1} | -0.138 | -0.200 | -0.045 | 0.016 | -0.005 | 0.011 | -0.033 | 0.026 | 0.022 | -0.007 |
| σ_{s_1} | 0.083 | 0.097 | 0.013 | 0.000 | 0.000 | -0.024 | -0.014 | 0.020 | 0.025 | 0.009 |
| μ_{s_2} | 0.327 | 0.455 | 0.107 | -0.050 | 0.024 | -0.025 | 0.083 | -0.105 | -0.068 | -0.074 |
| σ_{s_2} | 0.153 | 0.228 | 0.070 | -0.052 | 0.019 | 0.014 | 0.090 | -0.085 | -0.125 | -0.056 |
| μ_{s_3} | -0.192 | -0.252 | -0.064 | 0.033 | -0.019 | 0.010 | -0.050 | 0.081 | 0.043 | 0.074 |
| σ_{s_3} | -0.238 | -0.310 | -0.083 | 0.053 | -0.015 | 0.005 | -0.068 | 0.064 | 0.091 | 0.036 |
| $r_{1,2}$ | 0.002 | 0.001 | 0.000 | 0.001 | -0.001 | -0.001 | -0.004 | 0.070 | 0.010 | 0.007 |
| $r_{2,3}$ | 0.029 | 0.035 | 0.007 | -0.005 | 0.002 | -0.004 | 0.000 | 0.005 | 0.056 | 0.012 |
| $r_{1,3}$ | 0.049 | 0.064 | 0.016 | -0.005 | 0.003 | -0.001 | 0.011 | -0.093 | -0.012 | -0.257 |

Notes: Parameters of the structural model $\Psi = [\zeta^L, \Delta\zeta, \phi, \sigma, \delta, y^L, \Delta y, \omega_1, \omega_2, \omega_3]$ are estimated by indirect inference. $\Delta\zeta = \zeta^H - \zeta^L$, and $\Delta y = y^H - y^L$. Standard errors are in parentheses. The parameters of the auxiliary model are $\hat{B} = [\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \gamma_1, \gamma_2, \mu_{s_1}, \sigma_{s_1}, \mu_{s_2}, \sigma_{s_2}, \mu_{s_3}, \sigma_{s_3}, r_{1,2}, r_{2,3}, r_{1,3}]$. The Sensitivity Matrix is expressed in semi-elasticities multiplied with the standard deviation of the moment of \hat{B} so that the elements of Λ can be interpreted as the relative change in the parameter of Ψ that can be attributed to a one standard deviation change in the parameter of \hat{B} .

Table 4: Predicted Human Capital, Study Time and Elasticities with respect to Variance

| Grade | t | H_t | Predicted levels | | | | | Variance Elasticities | | | |
|--------------------------------|-----|--------|------------------|----------------------------|---------------|----------------------------|------------|-----------------------|--------|-----------|--------|
| | | | without variance | | with variance | | | total | | ex ante | |
| | | | H_{t+3} | $\overline{s_t - s_{t+2}}$ | H_{t+3} | $\overline{s_t - s_{t+2}}$ | $var(s_t)$ | H_{t+3} | s_t | H_{t+3} | s_t |
| Panel A: Low-return household | | | | | | | | | | | |
| 1 | 1 | n.a. | 7.845 | 7.776 | 6.929 | 7.017 | 1.276 | -0.132 | -0.175 | -0.094 | -0.175 |
| 2 | 4 | 6.929 | 6.580 | 5.866 | 6.339 | 5.510 | 0.589 | -0.144 | -0.183 | -0.094 | -0.167 |
| 3 | 7 | 6.339 | 6.032 | 4.498 | 5.933 | 4.303 | 0.301 | -0.147 | -0.180 | -0.090 | -0.154 |
| 4 | 10 | 5.933 | 5.696 | 3.466 | 5.646 | 3.346 | 0.162 | -0.150 | -0.186 | -0.088 | -0.150 |
| 5 | 13 | 5.646 | 5.471 | 2.662 | 5.442 | 2.586 | 0.090 | -0.151 | -0.190 | -0.086 | -0.146 |
| 6 | 16 | 5.442 | 5.314 | 2.049 | 5.296 | 1.998 | 0.051 | -0.152 | -0.183 | -0.084 | -0.138 |
| 7 | 19 | 5.296 | 5.203 | 1.577 | 5.191 | 1.541 | 0.029 | -0.152 | -0.183 | -0.083 | -0.136 |
| 8 | 22 | 5.191 | 5.122 | 1.212 | 5.114 | 1.187 | 0.017 | -0.152 | -0.185 | -0.082 | -0.135 |
| 9 | 25 | 5.114 | 5.064 | 0.931 | 5.058 | 0.913 | 0.010 | -0.152 | -0.184 | -0.082 | -0.133 |
| Total | | | 5.775 | | 5.058 | | | -0.142 | | | |
| Panel B: High-return household | | | | | | | | | | | |
| 1 | 1 | n.a. | 11.603 | 11.524 | 10.232 | 10.405 | 2.617 | -0.134 | -0.150 | -0.092 | -0.150 |
| 2 | 4 | 10.232 | 9.854 | 9.005 | 9.488 | 8.463 | 1.624 | -0.147 | -0.185 | -0.093 | -0.169 |
| 3 | 7 | 9.488 | 9.094 | 6.994 | 8.944 | 6.688 | 0.839 | -0.150 | -0.182 | -0.089 | -0.156 |
| 4 | 10 | 8.944 | 8.630 | 5.432 | 8.556 | 5.246 | 0.442 | -0.151 | -0.180 | -0.086 | -0.146 |
| 5 | 13 | 8.556 | 8.320 | 4.209 | 8.278 | 4.088 | 0.242 | -0.152 | -0.179 | -0.084 | -0.140 |
| 6 | 16 | 8.278 | 8.104 | 3.253 | 8.078 | 3.171 | 0.135 | -0.152 | -0.179 | -0.083 | -0.136 |
| 7 | 19 | 8.078 | 7.950 | 2.510 | 7.933 | 2.452 | 0.076 | -0.152 | -0.177 | -0.081 | -0.132 |
| 8 | 22 | 7.933 | 7.839 | 1.934 | 7.827 | 1.893 | 0.044 | -0.152 | -0.176 | -0.081 | -0.129 |
| 9 | 25 | 7.827 | 7.758 | 1.488 | 7.750 | 1.459 | 0.025 | -0.152 | -0.175 | -0.080 | -0.127 |
| Total | | | 8.834 | | 7.750 | | | -0.140 | | | |

Notes: Levels of s_t and H_{t+3} are obtained from numerically solving the dynamic model for three seasons of each grade as described by eq. (12) and parameter vector Ψ shown in table 3, and calculating the weighted average over 27 possible combinations of y^H , y^M and y^L for each type.

Table 5: Effect of Rainfall on Economic Outcomes

| | Household | | | Maternal time | | | | |
|---|---------------------|--------------------|----------------------|---------------------|----------------------|-------------------|-------------------------|-------------------|
| | Agr. income (1) | (2) | Cons. expend. (3) | (4) | Labor market (5) | Market (6) | Household chores (7) | (8) |
| Precipitation, current year (log mms) | 0.960** (0.456) | 0.826** (0.413) | 0.229*** (0.076) | 0.138** (0.061) | -1.914*** (0.563) | -0.841 (0.520) | 0.972* (0.504) | 0.593 (0.459) |
| Irrigated area (share of village agr. area) | 6.580** (2.948) | 3.583 (2.466) | 1.608*** (0.397) | 0.907*** (0.339) | -1.754 (2.535) | -4.313 (2.662) | 0.621 (2.636) | 2.200 (2.210) |
| Precipitation \times Irrigated area | -0.900** (0.425) | -0.497 (0.358) | -0.210*** (0.057) | -0.116** (0.047) | 0.256 (0.366) | 0.640* (0.382) | -0.101 (0.384) | -0.274 (0.317) |
| State-by-round FE | Y | Y | Y | Y | Y | Y | Y | Y |
| Village FE | N | Y | N | Y | N | Y | N | Y |
| Dependent variable mean | ln(9455) | ln(9455) | ln(11274) | ln(11274) | 3.04 | 3.04 | 9.04 | 9.04 |
| Observations | 15849 | 15849 | 20445 | 20445 | 33486 | 33486 | 33486 | 33486 |
| Adjusted R^2 | 0.062 | 0.144 | 0.266 | 0.334 | 0.225 | 0.306 | 0.245 | 0.310 |

Notes: Agricultural income and consumption expenditures are in log INR per capita. Female time allocation is in hours per day for three seasons. Each regression controls for state-by-round of interview fixed effects, and for lagged rainfall (in logs) in the last 3 years (2 years in cols. (1) and (2)). Columns (5) to (8) also control for season-by-round, and age fixed effects. Columns (2), (4), (6) and (8) additionally control for village fixed effects. Conley (1999) standard errors in parentheses (cut-off 300km, Bartlett weights). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Effect of of Rainfall on Child Time Use

| Hours per day: | On-farm | | Chores | | Studying | | Leisure | |
|---|----------------------|----------------------|---------------------|----------------------|---------------------|----------------------|---------------------|-------------------|
| | Boys (1) | Girls (2) | Boys (3) | Girls (4) | Boys (5) | Girls (6) | Boys (7) | Girls (8) |
| Precipitation, current year (log mms) | -0.822** (0.344) | -0.446* (0.253) | -0.165 (0.573) | -0.532 (0.643) | 1.156*** (0.381) | 1.516*** (0.456) | -0.093 (0.494) | -0.522 (0.507) |
| Irrigated area (share of village agr. area) | -4.670*** (1.702) | -3.363*** (1.296) | -5.256** (2.326) | -8.686*** (2.777) | 5.867** (2.534) | 13.088*** (2.998) | 5.025** (2.325) | 1.079 (2.640) |
| Precipitation \times Irrigated area | 0.664*** (0.240) | 0.476*** (0.182) | 0.694** (0.334) | 1.198*** (0.398) | -0.766** (0.359) | -1.814*** (0.428) | -0.737** (0.333) | -0.159 (0.372) |
| Dep. Var. Mean | 0.35 | 0.24 | 1.92 | 2.94 | 6.44 | 5.71 | 11.27 | 11.11 |
| Observations | 37005 | 31296 | 37005 | 31296 | 37005 | 31296 | 37005 | 31296 |
| Adjusted R^2 | 0.180 | 0.156 | 0.310 | 0.402 | 0.218 | 0.288 | 0.498 | 0.485 |

Notes: Time allocation is reported in hours per day over three seasons. Each regression controls for village, state-by-round of interview, season-by-round, and age fixed effects, and for lagged rainfall (in logs) in the last 3 years. Conley (1999) standard errors in parentheses (cut-off 300km, Bartlett weights).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Effect of Risk on School Attendance

| | (1) | (2) | (3) | (4) | (5) |
|----------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|-------------------------------|
| Panel A: | | | | | |
| IQR of predicted log consumption | -0.998** (0.297) [0.039] | -1.025** (0.299) [0.037] | -1.019** (0.299) [0.036] | -0.969** (0.278) [0.028] | -0.889* (0.302) [0.057] |
| Treatment Mean | 0.035 | 0.035 | 0.035 | 0.035 | 0.035 |
| Treatment Std. Dev. | 0.030 | 0.030 | 0.030 | 0.030 | 0.030 |
| Adjusted R^2 | 0.207 | 0.208 | 0.209 | 0.228 | 0.240 |
| Panel B: | | | | | |
| SD of predicted log consumption | -1.154** (0.379) [0.050] | -1.166** (0.388) [0.050] | -1.158* (0.388) [0.055] | -1.096** (0.358) [0.045] | -1.006* (0.383) [0.081] |
| Treatment Mean | 0.027 | 0.027 | 0.027 | 0.027 | 0.027 |
| Treatment Std. Dev. | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 |
| Adjusted R^2 | 0.207 | 0.208 | 0.208 | 0.228 | 0.239 |
| Rainfall controls | N | Y | Y | Y | Y |
| Rainfall by schoolage | N | N | Y | Y | Y |
| Income, wealth, area, hh size | N | N | N | Y | Y |
| State-by-round FE | N | N | N | N | Y |
| Dep. Var. mean | 0.762 | 0.762 | 0.762 | 0.762 | 0.762 |
| Observations | 68301 | 68301 | 68301 | 68301 | 68301 |

Notes: Each regression controls for village, season-by-round of interview, and age-by-gender fixed effects. Rainfall controls are current and lagged rainfall in past 3 years (in log mms), and the standard deviation of rainfall over the last three years. Rainfall by schoolage additionally interacts each rainfall lag with a dummy that equals 1 if the child was 6 years or older in that year. Income, wealth and area are measured in per capita terms (and are in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table 8: Effect of risk on primary completion (male adults)

| | (1) | (2) |
|----------------------------------|-------------------------------|-------------------------------|
| Panel A: | | |
| IQR of predicted log consumption | -1.104* (0.496) [0.077] | -2.043* (0.958) [0.089] |
| Treatment Mean | 0.036 | 0.036 |
| Treatment Std. Dev. | 0.030 | 0.030 |
| R^2 | 0.234 | 0.711 |
| Panel B: | | |
| SD of predicted log consumption | -1.254 (0.621) [0.108] | -2.280 (1.173) [0.113] |
| Treatment Mean | 0.028 | 0.028 |
| Treatment Std. Dev. | 0.023 | 0.023 |
| R^2 | 0.234 | 0.710 |
| Dep. var. mean | 0.800 | 0.800 |
| Fixed Effects | Village | Household |
| Observations | 7488 | 3396 |

Sample: Adult men observed in 2007-08 round of REDS.

Notes: Risk observed in village in 1981-82 is used as explanatory variable for men aged 35-46, and risk observed in village in 1998-99 for men aged 18-29, hence we merge the level of risk at ages between 10 and 21. Each regression controls for age and village fixed effects, and column (2) additionally controls for household fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table 9: Effect of Risk on Economic Outcomes

| | Income | Wealth | Land | Cons. exp | Curr. savings | Agr. prod. (=1) | Mat. labor supply |
|----------------------------------|------------------------------|------------------------------|-----------------------------|--------------------------------|------------------------------|-----------------------------|------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| PANEL A: | | | | | | | |
| IQR of predicted log consumption | -2.018 (1.138) [0.299] | -1.853 (0.841) [0.141] | 1.411 (1.438) [0.463] | -1.513** (0.429) [0.024] | -3.244 (1.673) [0.207] | 0.337 (0.286) [0.406] | -3.427 (2.703) [0.333] |
| Treatment Mean | 0.035 | 0.035 | 0.035 | 0.035 | 0.035 | 0.035 | 0.035 |
| Treatment Std. Dev. | 0.030 | 0.030 | 0.030 | 0.030 | 0.030 | 0.030 | 0.030 |
| Adjusted R^2 | 0.147 | 0.287 | 0.209 | 0.347 | 0.277 | 0.171 | 0.305 |
| PANEL B: | | | | | | | |
| SD of predicted log consumption | -2.328 (1.356) [0.292] | -2.326 (1.089) [0.167] | 2.313 (1.763) [0.339] | -1.985** (0.563) [0.030] | -3.634 (2.162) [0.261] | 0.624 (0.391) [0.290] | -3.168 (3.508) [0.507] |
| Treatment Mean | 0.027 | 0.027 | 0.027 | 0.027 | 0.027 | 0.027 | 0.027 |
| Treatment Std. Dev. | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 |
| Adjusted R^2 | 0.147 | 0.287 | 0.209 | 0.347 | 0.277 | 0.171 | 0.305 |
| Dependent variable mean | ln(16688) | ln(31314) | ln(0.778) | ln(11274) | ln(7090) | 0.789 | 3.044 |
| Observations | 20445 | 20445 | 20445 | 20445 | 11496 | 20445 | 33486 |

Notes: Income, land owned, wealth (value of all assets owned), consumptions expenditures and savings (if > 0) are expressed in per capita terms and logs. Maternal labor supply is in hours per day for three seasons. Each regression controls for village, state-by-round of interview fixed effects, and for current and lagged rainfall in past 3 years (1 years in cols. (1) and (6), in log mms), and for the standard deviation of rainfall over the last three years. Col. (6) additionally controls for age and for season-by-round of interview fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile Bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

A Mathematical Appendix

A.1 Incorporating child home production

A model that fits our data better links risk and schooling via the time children allocate to home production. While the share of time that children in our sample allocate to the labor market is declining strongly over time, we find that boys and girls allocate time to household chores and to own agricultural production throughout. Importantly, this time-allocation seems to be responsive to rainfall shocks (with children spending more time on these activities when faced with a negative income shock).

To generate this pattern in the absence of a child labor market, we incorporate a home produced good in the model, as well an *ex post* labor supply response to income shocks on the part of the parent. The idea is that parents supply more labor to the labor market in the face of an adverse shock to cope with the shock, as found by Kochar (1999); Rose (2001); Jayachandran (2006) and — at least for mothers — also in our data.⁵¹ As long as there is substitution between the parent’s and the child’s time in home production, the opportunity cost of time of the child must rise during bad periods even though the child is unlikely to be active in the labor market (Skoufias, 1993; Ilahi, 2000).

The single-period utility function u over the market-produced good c_t^m and the home-produced good c_t^h is assumed to be log Cobb-Douglas with share parameter α :

$$U_t = u(c_t^m, c_t^h) = \alpha \ln c_t^m + (1 - \alpha) \ln c_t^h. \quad (\text{A.1})$$

Market goods are purchased using cash income that is earned by the parent with wage w_t and a time allocation of $T^a - h_t^a$, where T^a is her time endowment and h_t^a is the time spent by the parent in home production,

$$c_t^m = I_t + \theta_t + w_t(T^a - h_t^a). \quad (\text{A.2})$$

I_t is the non-stochastic part of income (for example non-labor income), and θ_t is a shock realized at time t . Consumption of the home-produced good is linear in the allocation of time by the parent h_t^a and the child h_t^c ,

$$c_t^h = h_t^a + \eta h_t^c, \quad (\text{A.3})$$

where η captures the relative productivity of the child in home production. School time, then, is the difference between the child’s time endowment T^c and the amount of time spent in home production:

$$h_t^c = T^c - s_t \quad (\text{A.4})$$

Parental time at home in each period is:

$$h_t^a = (1 - \alpha)T^a - \alpha\eta h_t^c + \frac{(1 - \alpha)}{w_t}(I_t + \theta_t). \quad (\text{A.5})$$

The result is decreasing in the child’s home time and the wage, and increasing in the income endowment and shock. With a more favorable shock or a lower wage, the parent spends less time

⁵¹As described in the empirical section, we do not observe paternal time use.

in the labor market and more time at home. If we then define

$$y_t = T^a + \eta T^c + \frac{(I_t + \theta_t)}{w_t}, \quad (\text{A.6})$$

we can rewrite eq. (A.1) to get a period utility that corresponds to that used in our initial model plus a constant, in which the price of schooling now varies with η

$$U_t = (1 - \alpha) \ln(1 - \alpha) + \alpha \ln(w_t) + \alpha \ln(\alpha) + \ln(y_t - \eta s_t) \quad (\text{A.7})$$

In the theoretical section, we normalize η to 1 without loss of generality. In the empirical analysis, we control for a vector of age-by-gender fixed effects to capture any differences in the price of schooling that can be attributed to η . Note that w_t appears as additive component in eq. (A.7) but not in the the main model. As wages fluctuate with rainfall, a rainfall shock (*i.e.* θ is negative) would reduce wages and labor supply in eq. (A.5), which could counteract the positive direct effect of the shock on parental labor supply. Empirically, this effect seems to be minimal, however, as parents do respond to rainfall shocks by increasing labor supply (such that the income effect of the shock must be dominating the wage effect). We therefore believe that treating w_t as a constant is a reasonable approximation in this context.

A.2 Incorporating savings

If we extend the simple two-period model to allow for savings or borrowings in period one against period two, the objective function becomes

$$v_1 = \max_{s_1, s_2} \ln(y_1 - s_1 - a_1) + \mathbb{E}_1[\ln(y_2 + \theta_2 + a_1 - s_2) + \zeta g(H_3)], \quad (\text{A.8})$$

where a_1 is the amount saved or borrowed in period one, and we abstract from the interest rate (*i.e.* we assume it equals 1). Note that this formulation does not allow the household to borrow beyond period 2. We again solve this expression numerically for various ρ and three functional forms of $g(H)$. As before, we use parameter values of $y_1 = y_2 = 10, r = 1, \nu = 1/2, \zeta = 1$.

We now find a negative effect of risk on first-period schooling over all values of ϕ and ρ as depicted in Figure C.3. This can be explained by a additional income-smoothing mechanism, that encourages households to save in the first period, in order to smooth consumption in the second period. Yet, this effect reverses in the second period, where the effect of risk on the expected value of second period schooling is positive except for very small ρ (see Figure C.4). The net effect on total human capital, finally, is exactly identical to the case without savings (see Figure C.5).

This exercise highlights the importance of distinguishing a negative effect of risk on first-period schooling that is due to income smoothing from a negative effect of risk that is due to dynamic complementarity. Yet, it also identifies a potential strategy to distinguish both mechanisms: While the effect of risk on schooling is positive in expectation *ex post* to shocks in a situation where the *ex ante* negative effect is driven by income smoothing concerns, it continues to be negative throughout for very concave preferences over human capital and no savings (see figure C.6). For log-utility and for linear preferences, the predicted effect of risk on the expected value of s_2 turns positive for moderate complementarity in production, but remains close to zero throughout.

B Data Appendix

B.1 REDS Data

Our primary data set is the Rural Economic and Demographic Survey (REDS). The REDS is the follow-up survey of the Additional Rural Incomes Survey (ARIS), which was carried out between 1969 and 1971. The ARIS sample was designed to represent the rural population of India across 17 major states, and covers 4,527 households in 259 villages. The REDS expanded the thematic coverage of the ARIS, and in addition to revisiting the original ARIS households, increased the sample size by randomly sampling additional households from the same villages. The REDS data were collected in 1981-82, 1998-99, and 2007-08.⁵² We use data from all three REDS rounds, but restrict the sample to the 242 villages from which data were collected consistently.⁵³ Because the sample size increased substantially over time, we re-weight all observations to give equal weight to each of the three survey rounds.

The REDS is particularly suited for the analysis proposed here, as it collects detailed information on time use for all women in the household and their children (linked through a mother identifier). This information refers the hours per day allocated to an extensive list of activities on a typical day, and is collected for three seasons of the year, which are also marked by very different levels of agricultural activity. The reference months are: October/November (season one), February (season two), and April/May (season three).⁵⁴ Importantly, this data explicitly differentiates between time in school, time allocated to household chores, to farm work, to the labor market and to leisure.

We restrict the sample to households with complete information on village irrigation, and on household consumption and income. This sample consists of 4,704 households in 1981-82, 7,148 households in 1998-999 and 8,593 households in 2007-08. Table D.1 reports summary statistics for these households. In order to use the available information efficiently, we use this sample to estimate the effect of shocks on household level outcomes, and to calculate the risk variable.

During the time period we study, agriculture is the dominant source of income for the majority of households in rural India, either through own agricultural production or through casual agricultural employment. In the REDS sample, the share of households involved in agricultural production decreased from 86% in 1982 to 71% in 2007-08. Yet, even in 2007-08, total household income still consists to 57% of income from casual employment in agriculture and from own agricultural production (see Table D.1).⁵⁵

Over this time period, the share of the net area cultivated that is under irrigation has increased from 51% to 60% in our sample (see Table D.1). This is higher than the national average, but follows similar trends: According to the Directorate of Economics and Statistics, the share of the net area sown that is under irrigation at the national level increased from 28% in 1980-81 to 45% in 2007-08 (Directorate of Economics and Statistics, 2019).

⁵²A reduced version of the questionnaire was also administered between 2014 and 2016. This survey did not collect information on time allocation, but can be used to simulate the effects of NREGS in Section 5.

⁵³Due to ongoing conflict, no data were collected in Jammu & Kashmir in the 1998-99 and 2007-08 round of interviews (11 villages), and in Assam in the 1981-81 and 2007-08 round of interviews (8 villages). The 242 villages for which data were collected in all three rounds are distributed across 15 states (17 states since the foundation of Chhattisgarh and Jharkhand in 2000).

⁵⁴As such, season one corresponds to the harvest period of rice (and planting period of wheat). Season two is the slack season, and season three the planting period of rice (and harvest period of wheat). In 2007-08, the questionnaire did not make reference to precise months.

⁵⁵The agricultural year in India typically lasts from July to June, and the main harvest is strongly influenced by the onset of the monsoon and precipitation levels during the monsoon. The monsoon onset varies across places in India, but is usually between May and July, and its end is mostly around September or October.

In our main analysis, then, we focus on all children aged 6 to 15 with complete information on time allocation across all three seasons. We focus on the age group 6 to 15 because most children still live at home at this age, and because the questionnaire only collects this information for women in the household and her own (co-resident) children (not her sons or daughters-in-law) in 1981-82 and in 1998-99. By restricting ourselves to younger ages, we seek to minimize any selection bias that could arise if children with low education were more likely to leave the house at younger ages (*e.g.* in order to get married). Because not all households have children in this age group, the child sample comprises 5,581 children from 2,388 households in 1981-82, 8,424 children from 3,818 households in 1998-99, and 8,762 children from 4,472 households in 2007-08. The unit of observation in the main estimation is the child by agricultural season and round of interview, resulting in a final sample size of 68,301 observations. Child-by-season level summary statistics are reported in Table D.2. Note that it is not possible to track individuals across time in the REDS data. While we can follow a subset of households over time (and control for household fixed effects in part of the analysis), we cannot control for individual fixed effects.

Achieving universal education has been the declared goal of Indian governments since independence. The Right to Education Act of 2002 declares free and compulsory education a fundamental right of children aged 6 to 14. Since then, substantial improvements have been made in the enrollment rates of boys and girls and in closing the gender gap in primary school enrollment, which are also reflected in our data. In the REDS sample, the share of boys aged 6 to 15 that are attending school increased from 68% in 1981-81 to 93% in 2007-08. Likewise, the share of girls that are attending school increased from 53% to 91% in the same age group.⁵⁶

The time use data presented in Table D.2 shows that boys and girls spend most of the day in school but that other activities are important, too. Boys in our sample spend on average 5.5 (7.1) hours per day in school or studying in 1981-82 (2007-08). They also spend an average of 2.9 (1.0) hours per day on household chores, and 1.6 (0.1) hours per day working on-farm or on the labor market. Girls spend slightly less time in school (4.2 hours per day in 1981-82 and 6.9 hours in 2007-08), and more time on household chores (4.8 hours per day in 1981-82 and 1.3 hours in 2007-08). As found in previous work, agricultural work and wage work are a little less relevant for girls: they spend about 1.1 (0.1) hours per day these activities in 1981-82 (2007-08). In terms of leisure, finally, we do not observe substantial differences between boys and girls, but an increase in time allocated to leisure over time.⁵⁷ In figure C.7, we graph the average time children spend on different activities on a typical day by age and gender.

B.2 Weather Data

To obtain precise rainfall data, we merge our data with ERA5 (reanalysis) monthly averaged precipitation data (Copernicus Climate Change Service, 2017). ERA5 is the fifth generation reanalysis for the global climate and weather by the European Center for Medium-Range Weather Forecasts (ECMWF). The data is available since 1979, and in grid cells of 0.25x0.25 degrees (approximately 31x31 km). We calculate total annual rainfall in Millimeters, and merge the data point that is closest to the village centroid.

⁵⁶Note that these figures are likely smaller than actual enrollment rates, because this variable captures whether a child is actually attending school in a given time period. Also, the reference period of season three (April/May) could overlap with summer vacations in the southern states of India.

⁵⁷Note, that the questionnaire changed in how it elicited leisure over survey rounds. Leisure includes sleep time in all cases, but in the rounds 1981-82 and 1998-99, the questionnaire only accounts for hours between 04:00 and 24:00, whereas the 2007-08 questionnaire accounts for 24 hours. Thus by construction, leisure should be around 4 hours higher in 2007-08.

Some villages fall in the same cell, resulting in 195 unique data points per round of interview. We merge the precipitation data by the agricultural year to explain agricultural production and income, *i.e.* we use May 1998 to April 1999 rainfall to explain agricultural income, area cultivated and yields of the agricultural year 1998-99, and so forth. Note, that the consumption module and the time use module in the REDS do not refer to the last agricultural year, but to the last 12 months prior to the interview (except in 1998-99, where the consumption module also refers to the AY 1998-99). We therefore merge precipitation data of the current year, *i.e.* May 1982 to April 1983, for interviews conducted in the fall of 1982, and so forth. Apart from generating the best fit, this seems sensible because the bulk of household consumption consists of food products, and food consumption is typically extrapolated from last month's consumption to the entire year.

B.3 SEPRI Data

The Socio-Economic Profiles of Rural Households in India (SEPRI) data of 2014 and 2016 is a follow-up survey to the REDS. It that was collected in 13 states of India: Andhra Pradesh, Bihar, Chhattisgarh, Gujarat, Haryana, Jharkhand, Madhya Pradesh, Maharashtra, Orissa, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal.

It samples the entire population of the REDS survey villages, but applies a questionnaire that is considerably shorter than the REDS. We drop Gujarat from the analysis because information about the NREGS was not collected for all villages in that state. The rainfall data covers the agricultural year 2012/13 for interviews conducted in 2014, and the agricultural year 2014/15 for interviews conducted in 2016. We also use a one year lag of employment generation within the NREGS (captured in the village questionnaire) to minimize concerns about reverse causality.

C Supplementary Figures

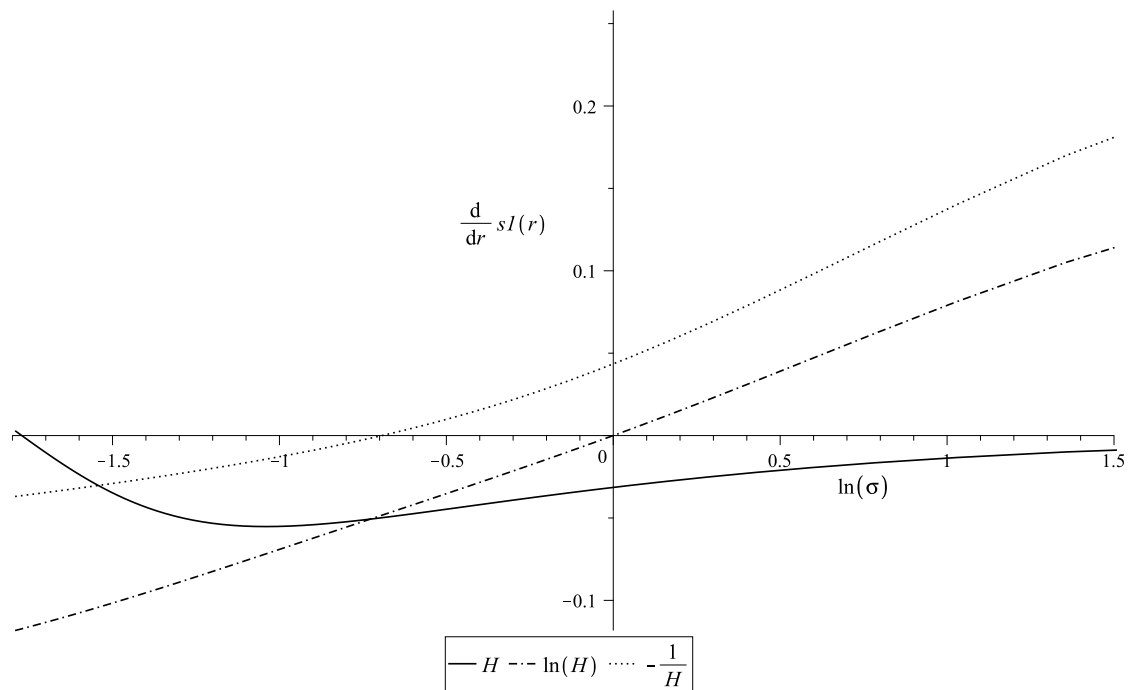


Figure C.1: Numeric Approximations of the Effect of Risk on First Period Schooling with High Risk

Notes: This figure presents numerical solutions to the derivative of s_1 wrt to the standard deviation of the shock ($\sigma_\theta = r$), *i.e.* the *ex ante* effect of risk, for a two-period model as described by eq. (5) and for various values of $\sigma = 1/(1 - \rho)$ and three specifications of $g(H)$: $g(H) = H$, $g(H) = \ln(H)$ and $g(H) = -1/H$. Assumed parameter values are $y_1 = 10$, $y_2 = 10$, $\nu = 1/2$, $r = 5$ and $\zeta = 1$.

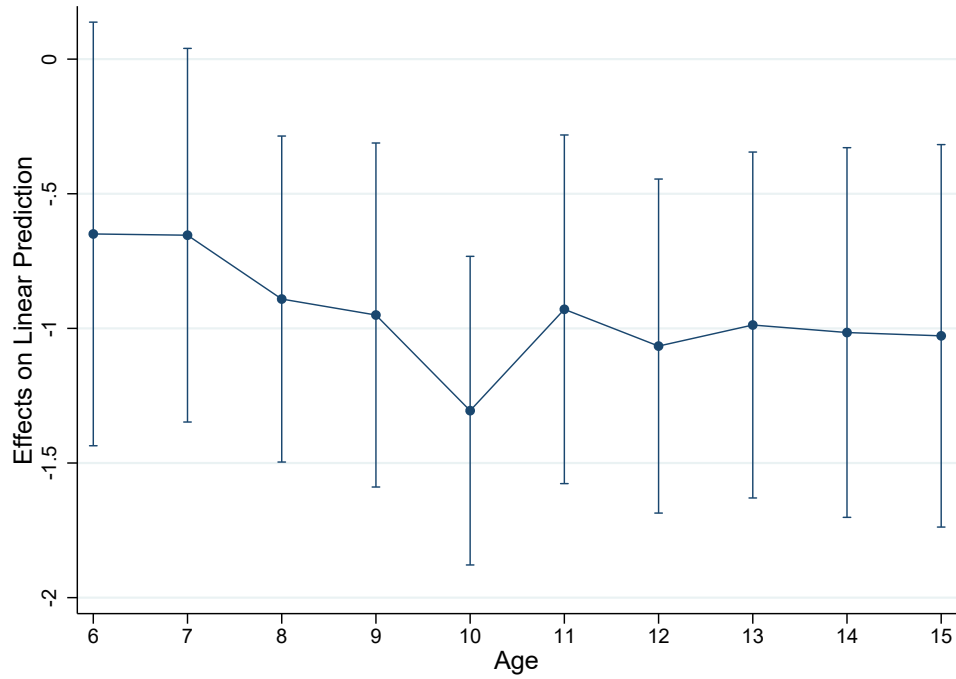


Figure C.2: Effect of Risk on School Attendance by Age

Notes: This figure plots the age-specific coefficients of risk and 90% confidence intervals from a regression of school attendance on risk interacted with dummies for each age. The regression controls for village, season-by-round of interview and gender fixed effects, for current and for lagged rainfall in the past 3 years (in log mms), and the standard deviation of rainfall over the last 3 years. The regression also controls for income, wealth and area (all measured in per capita terms, in logs) and for household size. Standard errors are corrected to allow for clustering at the level of the grid cell by round of interview.

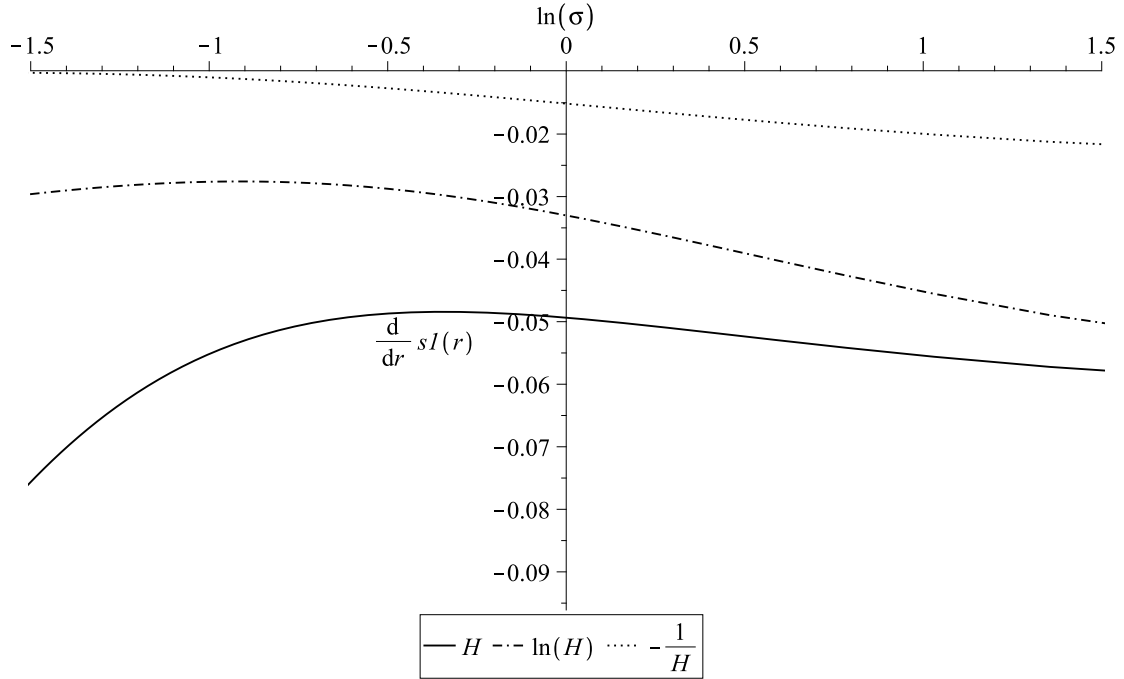


Figure C.3: Numeric Approximations of the Effect of Risk on First Period Schooling with Savings

Notes: This figure presents numerical solutions to the derivative of s_1 wrt to the standard deviation of the shock ($\sigma_\theta = r$), *i.e.* the *ex ante* effect of risk, for a two-period model that allows for precautionary savings as described by eq. (A.8) and for various values of $\sigma = 1/(1 - \rho)$ and three specifications of $g(H)$: $g(H) = H$, $g(H) = \ln(H)$ and $g(H) = -1/H$. Assumed parameter values are $y_1 = 10$, $y_2 = 10$, $\nu = 1/2$, $r = 1$ and $\zeta = 1$.

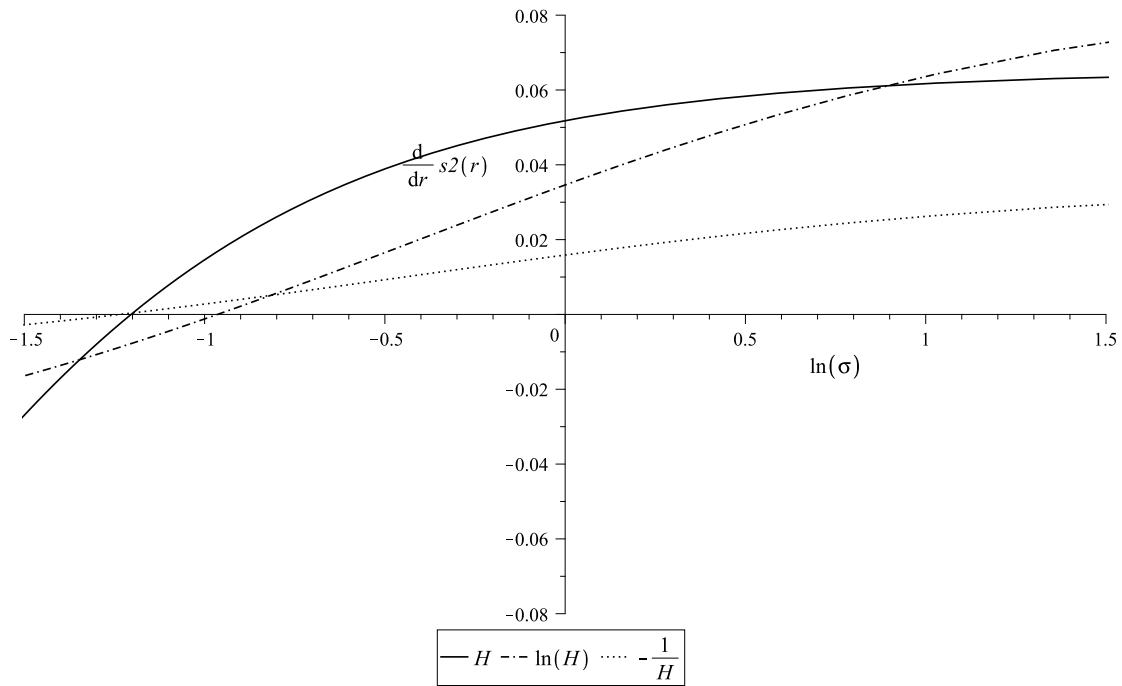


Figure C.4: Numeric Approximations of the Effect of Risk on Expected Second Period Schooling with Savings

Notes: This figure presents numerical solutions to the derivative of $\mathbb{E}[s_2]$ wrt to the standard deviation of the shock ($\sigma_\theta = r$), *i.e.* the *ex post* effect of risk, for a two-period model with precautionary savings as described by eq. (A.8) and for various values of $\sigma = 1/(1 - \rho)$ and three specifications of $g(H)$: $g(H) = H$, $g(H) = \ln(H)$ and $g(H) = -1/H$. Assumed parameter values are $y_1 = 10$, $y_2 = 10$, $\nu = 1/2$, $r = 1$ and $\zeta = 1$.

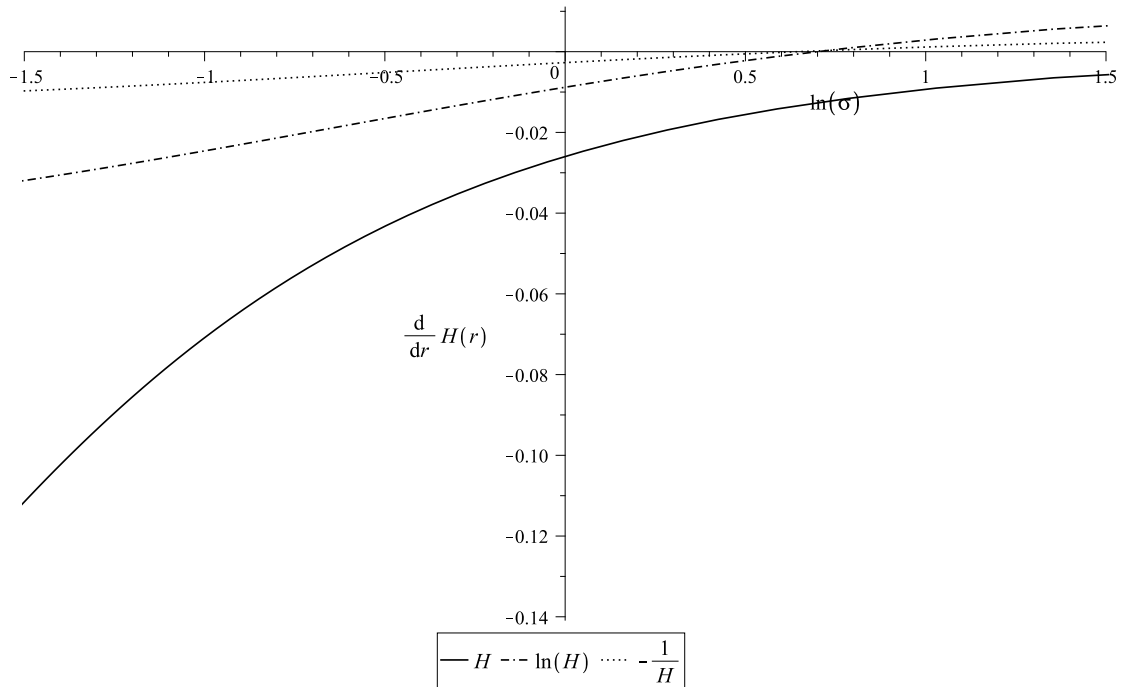


Figure C.5: Numeric Approximations of the Effect of Risk on Final Human Capital with Savings
Notes: This figure presents numerical solutions to the derivative of $\mathbb{E}[H_3]$ wrt to the standard deviation of the shock ($\sigma_\theta = r$), *i.e.* the combined *ex ante* and *ex post* effect of risk, for a two-period model with precautionary savings as described by eq. (A.8) and for various values of $\sigma = 1/(1 - \rho)$ and three specifications of $g(H)$: $g(H) = H$, $g(H) = \ln(H)$ and $g(H) = -1/H$. Assumed parameter values are $y_1 = 10$, $y_2 = 10$, $\nu = 1/2$, $r = 1$ and $\zeta = 1$.

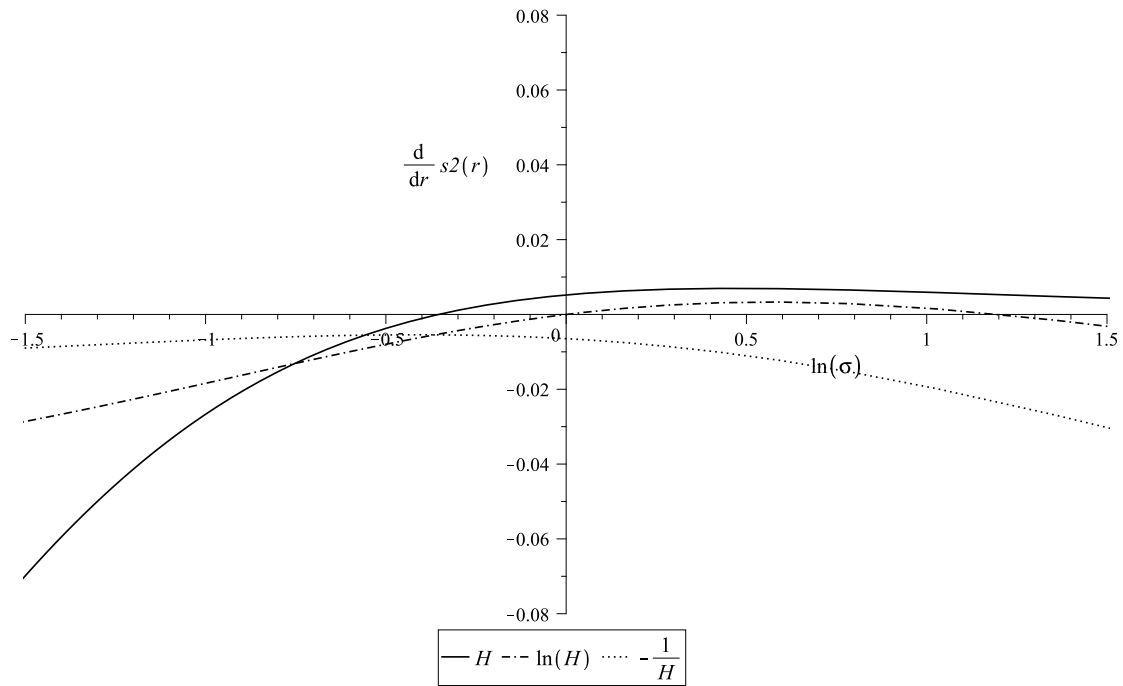


Figure C.6: Numeric Approximations of the Effect of Risk on Expected Second Period Schooling
— No Savings

Notes: This figure presents numerical solutions to the derivative of $\mathbb{E}[s_2]$ wrt to the standard deviation of the shock ($\sigma_\theta = r$), *i.e.* the *ex post* effect of risk, for a two-period model as described by eq. (5) and for various values of $\sigma = 1/(1 - \rho)$ and three specifications of $g(H)$: $g(H) = H$, $g(H) = \ln(H)$ and $g(H) = -1/H$. Assumed parameter values are $y_1 = 10$, $y_2 = 10$, $\nu = 1/2$, $r = 1$ and $\zeta = 1$.

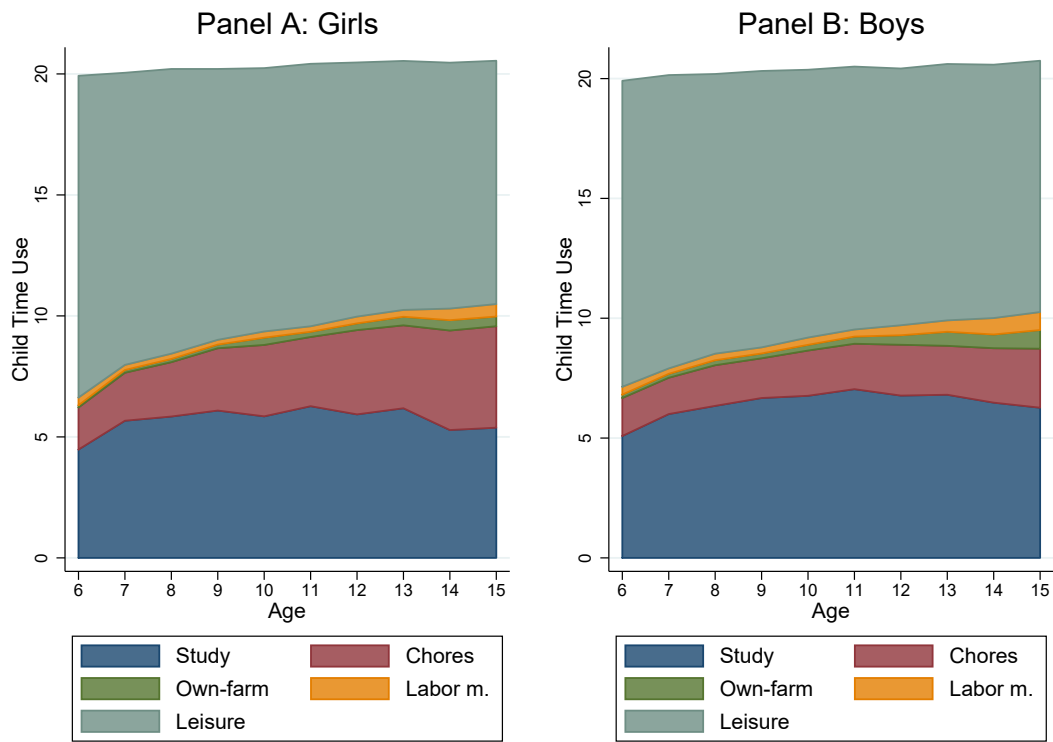


Figure C.7: Time-Use of Children by Gender and Age

Notes: This figure shows the average reported time use of children in hours per day by age and gender.

D Supplementary Tables

Table D.1: Household Characteristics

| | 1981-82 | | 1998-99 | | 2007-08 | |
|--|----------|----------|----------|----------|----------|----------|
| | Mean | SD | Mean | SD | Mean | SD |
| Household size | 6.86 | 3.56 | 6.06 | 3.40 | 5.17 | 2.60 |
| Hh owns any agr. land | 0.78 | 0.42 | 0.75 | 0.43 | 0.63 | 0.48 |
| Area owned (acres) | 6.37 | 9.36 | 4.05 | 7.42 | 2.73 | 5.44 |
| Area owned per capita | 1.01 | 1.67 | 0.74 | 1.36 | 0.59 | 1.32 |
| Household cultivated any land | 0.86 | 0.35 | 0.80 | 0.40 | 0.71 | 0.45 |
| Share of income from agriculture | 0.44 | 0.39 | 0.43 | 0.42 | 0.33 | 0.37 |
| Income from wages | 14709.3 | 25610.0 | 32335.0 | 51654.5 | 27994.7 | 44424.4 |
| Income from agr. wage work | 4371.2 | 9473.7 | 6424.2 | 14302.8 | 6501.0 | 14236.9 |
| Annual income, agriculture and livestock | 40506.3 | 57921.5 | 43282.0 | 85627.1 | 56836.5 | 104822.9 |
| Income from self-employment | 7107.5 | 33741.6 | 7073.5 | 35060.2 | 17735.2 | 86250.2 |
| Non-labor income | 2737.5 | 21902.7 | 3546.0 | 19269.8 | 9185.3 | 36442.8 |
| Annual income, total | 65060.6 | 74125.4 | 86236.5 | 103161.0 | 111751.7 | 142369.1 |
| Annual income per capita | 10043.8 | 11327.7 | 16026.8 | 21766.6 | 23993.0 | 33258.6 |
| Total consumption expenditures | 56239.7 | 39150.9 | 66026.0 | 54179.0 | 59163.8 | 42267.5 |
| Consumption expenditures per capita | 9092.6 | 6320.7 | 11983.7 | 8538.1 | 12746.4 | 9682.2 |
| Household wealth | 165323.5 | 340774.7 | 175813.7 | 377987.9 | 157351.8 | 239054.1 |
| Household wealth per capita | 26977.3 | 57602.9 | 32451.1 | 68614.3 | 34514.2 | 55612.3 |
| Irrigated area (fraction of village agr. area) | 0.51 | 0.39 | 0.58 | 0.36 | 0.60 | 0.34 |
| Precipitation, current year | 864.5 | 406.8 | 1087.1 | 506.8 | 1212.1 | 516.3 |
| Precipitation, one year lag | 1012.0 | 502.0 | 1182.3 | 521.7 | 1148.1 | 660.8 |
| Precipitation, two years lag | 1080.7 | 560.5 | 1113.9 | 500.7 | 1111.8 | 508.5 |
| Precipitation, three years lag | 867.7 | 482.0 | 1074.9 | 553.6 | 1053.0 | 494.5 |
| Rainfall Quintile | 1.96 | 1.15 | 3.32 | 1.21 | 3.80 | 1.27 |
| IQR of predicted log consumption | 0.039 | 0.032 | 0.033 | 0.028 | 0.032 | 0.027 |
| SD of predicted log consumption | 0.030 | 0.025 | 0.026 | 0.022 | 0.025 | 0.022 |
| Observations | 4704 | | 7148 | | 8593 | |

Notes: All nominal values are reported in constant 2010 INR. Prices are deflated using the All India Consumer Price Index for Agricultural Labourers in the month of July.

Table D.2: Individual Characteristics

| | 1981-82 | | 1999-99 | | 2007-08 | |
|---|---------|------|---------|------|---------|------|
| | Mean | SD | Mean | SD | Mean | SD |
| PANEL A: Boys | | | | | | |
| Age | 10.4 | 2.79 | 10.2 | 2.81 | 10.9 | 2.76 |
| Attending school | 0.68 | 0.47 | 0.81 | 0.39 | 0.93 | 0.25 |
| Hours per day: wage work | 0.78 | 2.26 | 0.27 | 1.15 | 0.10 | 0.89 |
| Hours per day: on-farm agriculture | 0.85 | 2.56 | 0.13 | 0.83 | 0.037 | 0.42 |
| Hours per day: household chores | 2.91 | 3.56 | 1.69 | 2.37 | 1.04 | 1.56 |
| Hours per day: studying | 5.45 | 4.22 | 6.88 | 3.70 | 7.08 | 2.62 |
| Hours per day: leisure | 8.87 | 3.70 | 11.0 | 3.57 | 14.3 | 1.40 |
| Observations | 9048 | | 13440 | | 14517 | |
| PANEL B: Girls | | | | | | |
| Age | 10.2 | 2.78 | 10.2 | 2.79 | 10.8 | 2.77 |
| Attending school | 0.53 | 0.50 | 0.74 | 0.44 | 0.91 | 0.29 |
| Hours per day: wage work | 0.53 | 1.82 | 0.26 | 1.10 | 0.058 | 0.65 |
| Hours per day: on-farm agriculture | 0.57 | 2.02 | 0.11 | 0.73 | 0.011 | 0.21 |
| Hours per day: household chores | 4.75 | 4.39 | 2.48 | 2.95 | 1.33 | 1.84 |
| Hours per day: studying | 4.19 | 4.31 | 6.23 | 4.01 | 6.89 | 2.81 |
| Hours per day: leisure | 8.76 | 3.85 | 10.9 | 3.82 | 14.2 | 1.48 |
| Observations | 7695 | | 11832 | | 11769 | |
| PANEL C: Mothers | | | | | | |
| Age of mother | 34.9 | 6.90 | 34.3 | 6.65 | 36.0 | 7.27 |
| Number of children born alive to mother | 4.88 | 2.21 | 3.53 | 1.67 | 3.35 | 1.67 |
| Hours per day: wage work | 1.54 | 2.98 | 1.54 | 2.81 | 1.29 | 2.85 |
| Hours per day: on-farm agriculture | 2.39 | 3.55 | 1.43 | 2.32 | 0.88 | 1.96 |
| Hours per day: household chores | 9.38 | 3.78 | 9.71 | 2.93 | 7.92 | 2.83 |
| Observations | 7806 | | 12504 | | 13176 | |

Notes: The unit of observation is the individual by season.

Table D.3: Study Time Across Seasons (All Interview Rounds)

| | (1) | (2) | (3) | (4) | (5) |
|--|----------------------|----------------------|----------------------|----------------------|----------------------|
| Study time, season 1 | 0.663*** (0.044) | 0.667*** (0.045) | 0.666*** (0.046) | 0.679*** (0.048) | 0.689*** (0.048) |
| Study time, season 1 (square) | 0.019*** (0.004) | 0.018*** (0.004) | 0.018*** (0.004) | 0.019*** (0.004) | 0.018*** (0.004) |
| Study time, season 2 | 0.361*** (0.049) | 0.359*** (0.050) | 0.357*** (0.050) | 0.344*** (0.052) | 0.337*** (0.054) |
| Study time, season 2 (square) | 0.049*** (0.005) | 0.049*** (0.005) | 0.049*** (0.005) | 0.050*** (0.005) | 0.051*** (0.005) |
| Study time, season 1 \times season 2 | -0.079*** (0.005) | -0.079*** (0.005) | -0.079*** (0.005) | -0.081*** (0.005) | -0.082*** (0.005) |
| Rainfall controls | N | Y | Y | Y | Y |
| Income, wealth, area, hh size, risk | N | N | Y | Y | Y |
| State-by-round FE | N | N | N | Y | Y |
| Village-by-round FE | N | N | N | N | Y |
| Observations | 22767 | 22767 | 22767 | 22767 | 22767 |
| Adjusted R^2 | 0.811 | 0.818 | 0.819 | 0.823 | 0.834 |

Notes: The dependent variable is study time in season 3. Each regression controls for village, round of interview, and age-by-gender fixed effects. Rainfall controls are current and lagged rainfall in past 3 years (in log mms). Income, wealth and area are measured in per capita terms (and are in logs). Risk is the interquartile range of predicted log consumption. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.4: Study Time Across Seasons (Panel Households Only)

| | (1) | (2) | (3) | (4) | (5) |
|--|----------------------|----------------------|----------------------|----------------------|----------------------|
| Study time, season 1 | 0.603*** (0.048) | 0.604*** (0.048) | 0.604*** (0.048) | 0.557*** (0.065) | 0.557*** (0.066) |
| Study time, season 1 (square) | 0.019*** (0.005) | 0.019*** (0.005) | 0.019*** (0.005) | 0.018*** (0.007) | 0.016** (0.007) |
| Study time, season 2 | 0.438*** (0.049) | 0.440*** (0.050) | 0.440*** (0.050) | 0.535*** (0.067) | 0.548*** (0.069) |
| Study time, season 2 (square) | 0.041*** (0.005) | 0.041*** (0.005) | 0.041*** (0.005) | 0.030*** (0.007) | 0.027*** (0.007) |
| Study time, season 1 \times season 2 | -0.075*** (0.006) | -0.075*** (0.006) | -0.075*** (0.006) | -0.068*** (0.009) | -0.066*** (0.009) |
| Rainfall controls | N | Y | Y | Y | Y |
| Income, wealth, area, hh size, risk | N | N | Y | Y | Y |
| Household-specific time trend | N | N | N | Y | Y |
| Village-by-round FE | N | N | N | N | Y |
| Observations | 18404 | 18404 | 18404 | 18404 | 18403 |
| Adjusted R^2 | 0.860 | 0.860 | 0.860 | 0.895 | 0.893 |

Notes: The dependent variable is study time in season 3. Each regression controls for household, state-by-round of interview, and age-by-gender fixed effects. Rainfall controls are current and lagged rainfall in past 3 years (in log mms). Income, wealth and area are measured in per capita terms (and are in logs). Risk is the interquartile range of predicted log consumption. Sample is restricted to households that are observed at least twice across the REDS rounds 1981-82, 1998-99, 2007-08. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.5: Study Time Across Seasons (Northern States only)

| | (1) | (2) | (3) | (4) | (5) |
|--|----------------------|----------------------|----------------------|----------------------|----------------------|
| Study time, season 1 | 0.721*** (0.052) | 0.721*** (0.054) | 0.721*** (0.054) | 0.744*** (0.057) | 0.757*** (0.057) |
| Study time, season 1 (square) 0.014*** | 0.014*** (0.005) | 0.014*** (0.005) | 0.014*** (0.005) | 0.013*** (0.005) | 0.014*** (0.005) |
| Study time, season 2 | 0.305*** (0.062) | 0.309*** (0.063) | 0.306*** (0.063) | 0.275*** (0.066) | 0.262*** (0.068) |
| Study time, season 2 (square) | 0.054*** (0.006) | 0.053*** (0.006) | 0.053*** (0.006) | 0.057*** (0.006) | 0.058*** (0.006) |
| Study time, season 1 \times season 2 | -0.081*** (0.006) | -0.081*** (0.006) | -0.081*** (0.006) | -0.084*** (0.006) | -0.084*** (0.006) |
| Rainfall controls | N | Y | Y | Y | Y |
| Income, wealth, area, hh size, risk | N | N | Y | Y | Y |
| State-by-round FE | N | N | N | Y | Y |
| Village-by-round FE | N | N | N | N | Y |
| Observations | 13982 | 13982 | 13982 | 13982 | 13982 |
| Adjusted R^2 | 0.809 | 0.810 | 0.810 | 0.815 | 0.825 |

Notes: The dependent variable is study time in season 3. Each regression controls for village, round of interview, and age-by-gender fixed effects. Rainfall controls are current and lagged rainfall in past 3 years (in log mms). Income, wealth and area are measured in per capita terms (and are in logs). Risk is the interquartile range of predicted log consumption. Sample is restricted to northern states of India, *i.e.* we drop Karnataka, Kerala, Tamil Nadu, Andhra Pradesh, and Maharashtra. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.6: Study Time Across Seasons by Gender and Agegroup

| | Boys (1) | Girls (2) | Age \leq 10 (3) | Age $>$ 10 (4) |
|--|----------------------|----------------------|----------------------|----------------------|
| Study time, season 1 | 0.606*** (0.052) | 0.503*** (0.054) | 0.576*** (0.049) | 0.578*** (0.055) |
| Study time, season 1 (square) | 0.025*** (0.005) | 0.028*** (0.005) | 0.025*** (0.005) | 0.025*** (0.006) |
| Study time, season 2 | 0.430*** (0.051) | 0.535*** (0.056) | 0.509*** (0.052) | 0.437*** (0.056) |
| Study time, season 2 (square) | 0.051*** (0.005) | 0.036*** (0.006) | 0.042*** (0.005) | 0.046*** (0.006) |
| Study time, season 1 \times season 2 | -0.088*** (0.006) | -0.076*** (0.007) | -0.083*** (0.006) | -0.081*** (0.007) |
| Observations | 9319 | 7866 | 8849 | 8335 |
| Adjusted R^2 | 0.792 | 0.840 | 0.816 | 0.815 |

Notes: The dependent variable is study time in season 3. Each regression controls for village-by-round of interview and age-by-gender fixed effects, as well as for current and lagged rainfall in past 3 years (in log mms), for income, wealth and area (all measured in per capita terms and are in logs), and for village risk (the interquartile range of predicted log consumption). Sample is restricted to REDS 1998-99 and 2007-08. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.7: Study Time Across Seasons 1 & 2

| | (1) | (2) | (3) | (4) | (5) |
|-------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Study time, season 1 | 0.986*** (0.026) | 0.994*** (0.026) | 0.990*** (0.026) | 0.994*** (0.026) | 1.003*** (0.028) |
| Study time, season 1 (square) | -0.011*** (0.002) | -0.013*** (0.002) | -0.012*** (0.002) | -0.013*** (0.002) | -0.014*** (0.002) |
| Rainfall controls | N | Y | Y | Y | Y |
| Income, wealth, area, hh size, risk | N | N | Y | Y | Y |
| State-by-round FE | N | N | N | Y | Y |
| Village-by-round FE | N | N | N | N | Y |
| Observations | 17186 | 17186 | 17186 | 17186 | 17186 |
| Adjusted R^2 | 0.791 | 0.794 | 0.795 | 0.795 | 0.798 |

Notes: The dependent variable is study time in season 2. Each regression controls for village, round of interview, and age-by-gender fixed effects. Rainfall controls are current and lagged rainfall in past 3 years (in log mms). Income, wealth and area are measured in per capita terms (and are in logs). Risk is the interquartile range of predicted log consumption. Sample is restricted to REDS 1998-99 and 2007-08. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.8: Effect of Rainfall on Income by Season

| | (1) | (2) | (3) |
|--|----------------------|----------------------|----------------------|
| Lagged Monthly Rainfall, January (log mms) | -0.122** (0.054) | -0.149** (0.060) | -0.053 (0.082) |
| Lagged Monthly Rainfall, February (log mms) | -0.019 (0.056) | 0.040 (0.061) | 0.240*** (0.085) |
| Lagged Monthly Rainfall, March (log mms) | 0.064 (0.060) | 0.110* (0.066) | 0.059 (0.091) |
| Lagged Monthly Rainfall, April (log mms) | 0.104 (0.070) | -0.026 (0.077) | -0.251** (0.106) |
| Lagged Monthly Rainfall, May (log mms) | -0.082 (0.081) | 0.173* (0.088) | 0.190 (0.122) |
| Lagged Monthly Rainfall, June (log mms) | -0.028 (0.107) | 0.064 (0.117) | -0.140 (0.161) |
| Lagged Monthly Rainfall, July (log mms) | 0.223* (0.115) | 0.086 (0.126) | -0.005 (0.174) |
| Lagged Monthly Rainfall, August (log mms) | -0.087 (0.110) | -0.015 (0.120) | -0.063 (0.165) |
| Lagged Monthly Rainfall, September (log mms) | 0.142 (0.126) | 0.252* (0.138) | -0.409** (0.191) |
| Lagged Monthly Rainfall, October (log mms) | -0.181*** (0.050) | -0.177*** (0.055) | 0.273*** (0.076) |
| Lagged Monthly Rainfall, November (log mms) | -0.040 (0.070) | 0.097 (0.077) | -0.012 (0.106) |
| Lagged Monthly Rainfall, December (log mms) | 0.122* (0.063) | 0.000 (0.069) | -0.383*** (0.095) |
| Monthly Rainfall, January (log mms) | 0.124* (0.072) | 0.019 (0.078) | -0.326*** (0.108) |
| Monthly Rainfall, February (log mms) | -0.080* (0.043) | 0.030 (0.047) | 0.079 (0.064) |
| Monthly Rainfall, March (log mms) | 0.015 (0.059) | 0.022 (0.065) | 0.239*** (0.089) |
| Monthly Rainfall, April (log mms) | -0.092 (0.076) | -0.101 (0.083) | 0.237** (0.114) |
| Observations | 480 | 480 | 480 |
| Dep. Var. mean | 10.066 | 9.685 | 8.217 |
| Adjusted R^2 | 0.487 | 0.569 | 0.373 |

Notes: The dependent variable is seasonal income from agricultural production and wage labor. Unit of observation is the village-by-round of interview. Each regression controls for village and round of interview fixed effects. Sample is restricted to REDS 1998-99 and 2007-08. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.9: Estimated Structural Parameters — Robustness

| Symbol | ζ^L | $\Delta\zeta$ | ϕ | σ | δ | y^L | Δy | ω_1 | ω_2 | ω_3 |
|---|-----------|---------------|--------|----------|----------|--------|------------|------------|------------|------------|
| Panel A: Main estimates | | | | | | | | | | |
| Estimate | 5.284 | 6.992 | 1.499 | 0.290 | 1.357 | 14.020 | 26.088 | 0.072 | 0.087 | -0.023 |
| Std error | 0.831 | 1.408 | 0.074 | 0.007 | 0.015 | 0.198 | 0.826 | 0.008 | 0.005 | 0.006 |
| Panel B: Constraining $\delta \leq 1$ | | | | | | | | | | |
| Estimate | 12.049 | 28.975 | 2.738 | 0.143 | 0.974 | 18.721 | 129.828 | 0.012 | 0.011 | -0.007 |
| Std error | 0.864 | 2.947 | 0.043 | 0.002 | 0.015 | 0.197 | 1.138 | 0.001 | 0.001 | 0.001 |
| Panel C: Initial probabilities of high, medium, low incomes = [1/5, 3/5, 1/5] | | | | | | | | | | |
| Estimate | 2.718 | 2.573 | 1.640 | 0.242 | 1.127 | 19.572 | 73.621 | 0.066 | 0.044 | -0.016 |
| Std error | 0.133 | 0.149 | 0.022 | 0.006 | 0.006 | 0.109 | 0.666 | 0.005 | 0.004 | 0.005 |
| Panel D: Weighting matrix is inverse variance-covariance matrix | | | | | | | | | | |
| Estimate | 4.899 | 6.963 | 1.542 | 0.289 | 1.376 | 13.418 | 29.212 | 0.062 | 0.079 | -0.028 |
| Std error | 0.810 | 1.661 | 0.062 | 0.013 | 0.026 | 0.488 | 0.861 | 0.006 | 0.005 | 0.005 |
| Panel E: Weighting matrix is identity matrix | | | | | | | | | | |
| Estimate | 8.028 | 11.479 | 1.672 | 0.239 | 1.439 | 11.711 | 28.988 | 0.070 | 0.061 | -0.030 |
| Std error | 0.916 | 1.781 | 0.060 | 0.008 | 0.018 | 0.242 | 1.283 | 0.008 | 0.005 | 0.006 |

Notes: Parameters of the structural model $\Psi = [\zeta^L, \Delta\zeta, \phi, \sigma, \delta, y^L, \Delta y, \omega_1, \omega_2, \omega_3]$ are estimated by indirect inference. $\Delta\zeta = \zeta^H - \zeta^L$, and $\Delta y = y^H - y^L$.

Table D.10: Predicted Human Capital, Study Time and Variance Elasticities — Low-return Hh

| Grade | H_t | Predicted levels | | | | | Variance Elasticities | | | |
|-------|-------|------------------|-----------------------|---------------|-----------------------|------------|-----------------------|--------|-----------|--------|
| | | without variance | | with variance | | | total | | ex ante | |
| | | H_{t+3} | $\bar{s}_t - s_{t+2}$ | H_{t+3} | $\bar{s}_t - s_{t+2}$ | $var(s_t)$ | H_{t+3} | s_t | H_{t+3} | s_t |
| 1 | 4 | 7.845 | 7.776 | 6.929 | 7.017 | 1.276 | -0.132 | -0.175 | -0.094 | -0.175 |
| 1 | 6 | 7.845 | 7.776 | 6.929 | 7.017 | 1.276 | -0.132 | -0.175 | -0.094 | -0.175 |
| 1 | 8 | 7.845 | 7.776 | 6.929 | 7.017 | 1.276 | -0.132 | -0.175 | -0.094 | -0.175 |
| 1 | 10 | 7.845 | 7.776 | 6.929 | 7.017 | 1.276 | -0.132 | -0.175 | -0.094 | -0.175 |
| 1 | 12 | 7.845 | 7.776 | 6.929 | 7.017 | 1.276 | -0.132 | -0.175 | -0.094 | -0.175 |
| 1 | 14 | 7.845 | 7.776 | 6.929 | 7.017 | 1.276 | -0.132 | -0.175 | -0.094 | -0.175 |
| 2 | 4 | 4.161 | 4.648 | 4.090 | 4.442 | 0.325 | -0.137 | -0.175 | -0.100 | -0.155 |
| 2 | 6 | 5.875 | 5.547 | 5.699 | 5.239 | 0.495 | -0.141 | -0.180 | -0.096 | -0.163 |
| 2 | 8 | 7.323 | 6.174 | 7.002 | 5.769 | 0.672 | -0.137 | -0.177 | -0.084 | -0.164 |
| 2 | 10 | 8.489 | 6.582 | 8.010 | 6.104 | 0.793 | -0.148 | -0.183 | -0.082 | -0.171 |
| 2 | 12 | 9.421 | 6.875 | 8.774 | 6.326 | 0.887 | -0.155 | -0.183 | -0.079 | -0.174 |
| 2 | 14 | 10.082 | 7.013 | 9.340 | 6.469 | 0.957 | -0.142 | -0.171 | -0.064 | -0.164 |
| 3 | 4 | 3.946 | 3.530 | 3.912 | 3.413 | 0.165 | -0.139 | -0.173 | -0.099 | -0.151 |
| 3 | 6 | 5.740 | 4.380 | 5.653 | 4.192 | 0.281 | -0.147 | -0.181 | -0.094 | -0.156 |
| 3 | 8 | 7.396 | 5.019 | 7.231 | 4.773 | 0.395 | -0.142 | -0.176 | -0.078 | -0.151 |
| 3 | 10 | 8.885 | 5.480 | 8.627 | 5.193 | 0.488 | -0.153 | -0.183 | -0.074 | -0.154 |
| 3 | 12 | 10.238 | 5.873 | 9.853 | 5.525 | 0.590 | -0.157 | -0.181 | -0.066 | -0.154 |
| 3 | 14 | 11.441 | 6.188 | 10.908 | 5.777 | 0.676 | -0.153 | -0.191 | -0.053 | -0.160 |
| 4 | 4 | 3.911 | 2.759 | 3.890 | 2.680 | 0.095 | -0.139 | -0.172 | -0.097 | -0.147 |
| 4 | 6 | 5.757 | 3.488 | 5.706 | 3.367 | 0.164 | -0.150 | -0.178 | -0.092 | -0.148 |
| 4 | 8 | 7.522 | 4.068 | 7.426 | 3.910 | 0.236 | -0.145 | -0.183 | -0.075 | -0.145 |
| 4 | 10 | 9.190 | 4.525 | 9.033 | 4.332 | 0.306 | -0.153 | -0.179 | -0.067 | -0.140 |
| 4 | 12 | 10.777 | 4.934 | 10.540 | 4.698 | 0.377 | -0.163 | -0.195 | -0.060 | -0.146 |
| 4 | 14 | 12.267 | 5.276 | 11.936 | 5.003 | 0.451 | -0.156 | -0.182 | -0.046 | -0.139 |
| 5 | 4 | 3.916 | 2.160 | 3.902 | 2.105 | 0.056 | -0.140 | -0.172 | -0.096 | -0.145 |
| 5 | 6 | 5.802 | 2.767 | 5.769 | 2.686 | 0.097 | -0.151 | -0.180 | -0.090 | -0.143 |
| 5 | 8 | 7.635 | 3.266 | 7.576 | 3.161 | 0.141 | -0.147 | -0.182 | -0.072 | -0.137 |
| 5 | 10 | 9.414 | 3.689 | 9.314 | 3.555 | 0.191 | -0.155 | -0.184 | -0.064 | -0.134 |
| 5 | 12 | 11.134 | 4.056 | 10.990 | 3.902 | 0.235 | -0.163 | -0.186 | -0.056 | -0.130 |
| 5 | 14 | 12.801 | 4.390 | 12.597 | 4.208 | 0.284 | -0.159 | -0.193 | -0.043 | -0.129 |
| 6 | 4 | 3.930 | 1.688 | 3.921 | 1.649 | 0.033 | -0.140 | -0.173 | -0.094 | -0.142 |
| 6 | 6 | 5.845 | 2.180 | 5.824 | 2.125 | 0.057 | -0.152 | -0.182 | -0.088 | -0.141 |
| 6 | 8 | 7.726 | 2.596 | 7.687 | 2.525 | 0.084 | -0.148 | -0.179 | -0.070 | -0.130 |
| 6 | 10 | 9.572 | 2.963 | 9.510 | 2.874 | 0.114 | -0.155 | -0.185 | -0.061 | -0.127 |
| 6 | 12 | 11.383 | 3.288 | 11.290 | 3.182 | 0.146 | -0.164 | -0.190 | -0.054 | -0.124 |
| 6 | 14 | 13.154 | 3.579 | 13.027 | 3.459 | 0.176 | -0.160 | -0.186 | -0.041 | -0.117 |
| 7 | 4 | 3.945 | 1.314 | 3.938 | 1.286 | 0.019 | -0.140 | -0.173 | -0.093 | -0.141 |
| 7 | 6 | 5.882 | 1.708 | 5.867 | 1.669 | 0.034 | -0.152 | -0.179 | -0.086 | -0.136 |
| 7 | 8 | 7.795 | 2.046 | 7.768 | 1.997 | 0.050 | -0.148 | -0.178 | -0.068 | -0.126 |
| 7 | 10 | 9.686 | 2.352 | 9.645 | 2.291 | 0.068 | -0.156 | -0.184 | -0.060 | -0.122 |
| 7 | 12 | 11.554 | 2.628 | 11.493 | 2.556 | 0.087 | -0.164 | -0.190 | -0.052 | -0.118 |
| 7 | 14 | 13.397 | 2.879 | 13.313 | 2.796 | 0.108 | -0.160 | -0.188 | -0.039 | -0.111 |
| 8 | 4 | 3.958 | 1.020 | 3.953 | 1.000 | 0.011 | -0.141 | -0.173 | -0.093 | -0.140 |
| 8 | 6 | 5.910 | 1.331 | 5.900 | 1.303 | 0.020 | -0.152 | -0.179 | -0.085 | -0.134 |
| 8 | 8 | 7.846 | 1.603 | 7.828 | 1.567 | 0.030 | -0.148 | -0.178 | -0.067 | -0.123 |
| 8 | 10 | 9.768 | 1.851 | 9.740 | 1.809 | 0.041 | -0.156 | -0.184 | -0.059 | -0.118 |
| 8 | 12 | 11.673 | 2.078 | 11.633 | 2.028 | 0.052 | -0.164 | -0.190 | -0.051 | -0.114 |
| 8 | 14 | 13.563 | 2.288 | 13.508 | 2.231 | 0.065 | -0.160 | -0.187 | -0.038 | -0.106 |
| 9 | 4 | 3.967 | 0.790 | 3.964 | 0.774 | 0.007 | -0.141 | -0.173 | -0.092 | -0.138 |
| 9 | 6 | 5.932 | 1.034 | 5.924 | 1.013 | 0.012 | -0.152 | -0.178 | -0.085 | -0.132 |
| 9 | 8 | 7.885 | 1.249 | 7.872 | 1.224 | 0.018 | -0.148 | -0.178 | -0.067 | -0.121 |
| 9 | 10 | 9.827 | 1.448 | 9.808 | 1.418 | 0.024 | -0.156 | -0.184 | -0.058 | -0.116 |
| 9 | 12 | 11.759 | 1.631 | 11.731 | 1.596 | 0.031 | -0.164 | -0.190 | -0.050 | -0.111 |
| 9 | 14 | 13.680 | 1.802 | 13.642 | 1.762 | 0.038 | -0.160 | -0.187 | -0.038 | -0.103 |

Notes: Levels of s_t and H_{t+3} are obtained from numerically solving the dynamic model for three seasons of each grade as described by eq. (12) and parameter vector Ψ shown in table 3, and calculating the weighted average over 27 possible combinations of y^H , y^M and y^L for each type. Elasticities with respect to variance are obtained from numerically solving the structural model for a counterfactual income structure with no variance.

Table D.11: Predicted Human Capital, Study Time and Variance Elasticities — High-return Hh

| Grade | Predicted levels | | | | | | Variance Elasticities | | | |
|-------|------------------|-----------|----------------------------|---------------|----------------------------|------------|-----------------------|--------|-----------|--------|
| | without variance | | | with variance | | | total | | ex ante | |
| | H_t | H_{t+3} | $\overline{s_t - s_{t+2}}$ | H_{t+3} | $\overline{s_t - s_{t+2}}$ | $var(s_t)$ | H_{t+3} | s_t | H_{t+3} | s_t |
| 1 | 4 | 11.603 | 11.524 | 10.232 | 10.405 | 2.617 | -0.134 | -0.150 | -0.092 | -0.150 |
| 1 | 6 | 11.603 | 11.524 | 10.232 | 10.405 | 2.617 | -0.134 | -0.150 | -0.092 | -0.150 |
| 1 | 8 | 11.603 | 11.524 | 10.232 | 10.405 | 2.617 | -0.134 | -0.150 | -0.092 | -0.150 |
| 1 | 10 | 11.603 | 11.524 | 10.232 | 10.405 | 2.617 | -0.134 | -0.150 | -0.092 | -0.150 |
| 1 | 12 | 11.603 | 11.524 | 10.232 | 10.405 | 2.617 | -0.134 | -0.150 | -0.092 | -0.150 |
| 1 | 14 | 11.603 | 11.524 | 10.232 | 10.405 | 2.617 | -0.134 | -0.150 | -0.092 | -0.150 |
| 2 | 4 | 4.348 | 6.055 | 4.307 | 5.835 | 0.574 | -0.136 | -0.173 | -0.101 | -0.150 |
| 2 | 6 | 6.304 | 7.320 | 6.198 | 6.998 | 0.936 | -0.138 | -0.173 | -0.098 | -0.154 |
| 2 | 8 | 8.082 | 8.229 | 7.880 | 7.825 | 1.269 | -0.139 | -0.171 | -0.094 | -0.154 |
| 2 | 10 | 9.679 | 8.931 | 9.333 | 8.404 | 1.585 | -0.146 | -0.188 | -0.093 | -0.170 |
| 2 | 12 | 11.058 | 9.410 | 10.569 | 8.834 | 1.877 | -0.145 | -0.179 | -0.086 | -0.165 |
| 2 | 14 | 12.284 | 9.842 | 11.606 | 9.154 | 2.122 | -0.148 | -0.181 | -0.081 | -0.170 |
| 3 | 4 | 4.045 | 4.495 | 4.025 | 4.361 | 0.283 | -0.138 | -0.172 | -0.102 | -0.151 |
| 3 | 6 | 5.960 | 5.598 | 5.911 | 5.409 | 0.469 | -0.139 | -0.170 | -0.098 | -0.149 |
| 3 | 8 | 7.792 | 6.463 | 7.694 | 6.212 | 0.671 | -0.142 | -0.177 | -0.092 | -0.153 |
| 3 | 10 | 9.529 | 7.158 | 9.359 | 6.832 | 0.896 | -0.150 | -0.182 | -0.089 | -0.157 |
| 3 | 12 | 11.161 | 7.711 | 10.908 | 7.342 | 1.087 | -0.149 | -0.177 | -0.080 | -0.153 |
| 3 | 14 | 12.699 | 8.200 | 12.332 | 7.758 | 1.271 | -0.155 | -0.192 | -0.075 | -0.162 |
| 4 | 4 | 3.975 | 3.467 | 3.963 | 3.379 | 0.154 | -0.138 | -0.171 | -0.101 | -0.148 |
| 4 | 6 | 5.896 | 4.386 | 5.866 | 4.259 | 0.265 | -0.140 | -0.171 | -0.096 | -0.146 |
| 4 | 8 | 7.766 | 5.125 | 7.709 | 4.959 | 0.383 | -0.143 | -0.176 | -0.089 | -0.147 |
| 4 | 10 | 9.581 | 5.750 | 9.484 | 5.540 | 0.514 | -0.152 | -0.182 | -0.086 | -0.148 |
| 4 | 12 | 11.338 | 6.282 | 11.189 | 6.031 | 0.643 | -0.151 | -0.180 | -0.075 | -0.145 |
| 4 | 14 | 13.028 | 6.732 | 12.822 | 6.457 | 0.763 | -0.157 | -0.185 | -0.069 | -0.144 |
| 5 | 4 | 3.961 | 2.696 | 3.953 | 2.633 | 0.089 | -0.138 | -0.169 | -0.100 | -0.145 |
| 5 | 6 | 5.895 | 3.438 | 5.876 | 3.350 | 0.152 | -0.140 | -0.169 | -0.094 | -0.142 |
| 5 | 8 | 7.797 | 4.053 | 7.761 | 3.940 | 0.221 | -0.144 | -0.174 | -0.087 | -0.141 |
| 5 | 10 | 9.665 | 4.590 | 9.606 | 4.450 | 0.298 | -0.153 | -0.181 | -0.083 | -0.142 |
| 5 | 12 | 11.497 | 5.061 | 11.406 | 4.893 | 0.379 | -0.152 | -0.181 | -0.072 | -0.138 |
| 5 | 14 | 13.288 | 5.474 | 13.160 | 5.282 | 0.458 | -0.158 | -0.184 | -0.066 | -0.135 |
| 6 | 4 | 3.963 | 2.096 | 3.957 | 2.050 | 0.052 | -0.138 | -0.169 | -0.100 | -0.143 |
| 6 | 6 | 5.911 | 2.687 | 5.898 | 2.625 | 0.088 | -0.140 | -0.168 | -0.093 | -0.139 |
| 6 | 8 | 7.838 | 3.188 | 7.814 | 3.109 | 0.129 | -0.144 | -0.173 | -0.086 | -0.137 |
| 6 | 10 | 9.742 | 3.635 | 9.704 | 3.539 | 0.174 | -0.153 | -0.180 | -0.081 | -0.136 |
| 6 | 12 | 11.623 | 4.034 | 11.565 | 3.920 | 0.222 | -0.152 | -0.180 | -0.070 | -0.131 |
| 6 | 14 | 13.479 | 4.392 | 13.397 | 4.261 | 0.272 | -0.158 | -0.184 | -0.063 | -0.129 |
| 7 | 4 | 3.969 | 1.625 | 3.964 | 1.592 | 0.030 | -0.138 | -0.168 | -0.099 | -0.142 |
| 7 | 6 | 5.929 | 2.093 | 5.920 | 2.048 | 0.052 | -0.141 | -0.167 | -0.092 | -0.137 |
| 7 | 8 | 7.874 | 2.495 | 7.858 | 2.438 | 0.075 | -0.144 | -0.172 | -0.085 | -0.134 |
| 7 | 10 | 9.805 | 2.859 | 9.779 | 2.791 | 0.102 | -0.153 | -0.179 | -0.080 | -0.133 |
| 7 | 12 | 11.719 | 3.187 | 11.681 | 3.108 | 0.130 | -0.153 | -0.179 | -0.069 | -0.127 |
| 7 | 14 | 13.617 | 3.488 | 13.563 | 3.396 | 0.160 | -0.158 | -0.183 | -0.062 | -0.123 |
| 8 | 4 | 3.975 | 1.257 | 3.972 | 1.233 | 0.018 | -0.139 | -0.168 | -0.099 | -0.141 |
| 8 | 6 | 5.945 | 1.625 | 5.939 | 1.592 | 0.030 | -0.140 | -0.167 | -0.091 | -0.135 |
| 8 | 8 | 7.904 | 1.944 | 7.893 | 1.903 | 0.044 | -0.144 | -0.171 | -0.084 | -0.132 |
| 8 | 10 | 9.853 | 2.235 | 9.835 | 2.186 | 0.060 | -0.153 | -0.178 | -0.079 | -0.130 |
| 8 | 12 | 11.790 | 2.500 | 11.764 | 2.444 | 0.076 | -0.153 | -0.178 | -0.068 | -0.123 |
| 8 | 14 | 13.717 | 2.746 | 13.681 | 2.682 | 0.094 | -0.158 | -0.182 | -0.061 | -0.120 |
| 9 | 4 | 3.981 | 0.971 | 3.978 | 0.953 | 0.010 | -0.139 | -0.168 | -0.098 | -0.140 |
| 9 | 6 | 5.958 | 1.258 | 5.953 | 1.233 | 0.018 | -0.140 | -0.166 | -0.091 | -0.134 |
| 9 | 8 | 7.927 | 1.509 | 7.919 | 1.479 | 0.026 | -0.144 | -0.171 | -0.083 | -0.130 |
| 9 | 10 | 9.889 | 1.740 | 9.876 | 1.704 | 0.035 | -0.153 | -0.178 | -0.078 | -0.128 |
| 9 | 12 | 11.843 | 1.951 | 11.825 | 1.910 | 0.045 | -0.152 | -0.177 | -0.067 | -0.121 |
| 9 | 14 | 13.790 | 2.149 | 13.765 | 2.103 | 0.055 | -0.158 | -0.182 | -0.060 | -0.117 |

Notes: Levels of s_t and H_{t+3} are obtained from numerically solving the dynamic model for three seasons of each grade as described by eq. (12) and parameter vector Ψ shown in table 3, and calculating the weighted average over 27 possible combinations of y^H , y^M and y^L for each type. Elasticities with respect to variance are obtained from numerically solving the structural model for a counterfactual income structure with no variance.

Table D.12: Predicted Human Capital, Study Time and Variance Elasticities — $\delta = 1$

| Grade | t | H_t | Predicted levels | | | | | Variance Elasticities | | | |
|--------------------------------|-----|-------|------------------|----------------------------|---------------|----------------------------|------------|-----------------------|--------|-----------|--------|
| | | | without variance | | with variance | | | total | | ex ante | |
| | | | H_{t+3} | $\overline{s_t - s_{t+2}}$ | H_{t+3} | $\overline{s_t - s_{t+2}}$ | $var(s_t)$ | H_{t+3} | s_t | H_{t+3} | s_t |
| Panel A: Low-return household | | | | | | | | | | | |
| 1 | 1 | n.a. | 3.452 | 3.452 | 3.160 | 3.206 | 0.140 | -0.093 | -0.141 | -0.065 | -0.141 |
| 2 | 4 | 3.160 | 3.289 | 3.447 | 3.157 | 3.226 | 0.145 | -0.095 | -0.137 | -0.069 | -0.136 |
| 3 | 7 | 3.157 | 3.238 | 3.429 | 3.159 | 3.238 | 0.149 | -0.097 | -0.140 | -0.071 | -0.135 |
| 4 | 10 | 3.159 | 3.220 | 3.424 | 3.164 | 3.256 | 0.152 | -0.100 | -0.145 | -0.073 | -0.137 |
| 5 | 13 | 3.164 | 3.213 | 3.422 | 3.171 | 3.282 | 0.156 | -0.103 | -0.147 | -0.074 | -0.136 |
| 6 | 16 | 3.171 | 3.213 | 3.426 | 3.180 | 3.305 | 0.158 | -0.106 | -0.150 | -0.075 | -0.135 |
| 7 | 19 | 3.180 | 3.216 | 3.438 | 3.189 | 3.329 | 0.161 | -0.108 | -0.154 | -0.076 | -0.137 |
| 8 | 22 | 3.189 | 3.221 | 3.446 | 3.200 | 3.354 | 0.163 | -0.110 | -0.156 | -0.076 | -0.136 |
| 9 | 25 | 3.200 | 3.229 | 3.459 | 3.211 | 3.379 | 0.165 | -0.111 | -0.157 | -0.076 | -0.135 |
| Total | | | 3.556 | | 3.211 | | | -0.108 | | | |
| Panel B: High-return household | | | | | | | | | | | |
| 1 | 1 | n.a. | 5.791 | 5.791 | 5.187 | 5.277 | 0.486 | -0.116 | -0.168 | -0.079 | -0.168 |
| 2 | 4 | 5.187 | 5.460 | 5.773 | 5.192 | 5.337 | 0.501 | -0.128 | -0.168 | -0.089 | -0.163 |
| 3 | 7 | 5.192 | 5.363 | 5.741 | 5.203 | 5.380 | 0.497 | -0.135 | -0.169 | -0.094 | -0.159 |
| 4 | 10 | 5.203 | 5.326 | 5.727 | 5.217 | 5.418 | 0.498 | -0.140 | -0.174 | -0.097 | -0.159 |
| 5 | 13 | 5.217 | 5.314 | 5.724 | 5.233 | 5.459 | 0.502 | -0.143 | -0.178 | -0.098 | -0.159 |
| 6 | 16 | 5.233 | 5.313 | 5.729 | 5.251 | 5.502 | 0.507 | -0.145 | -0.182 | -0.099 | -0.159 |
| 7 | 19 | 5.251 | 5.319 | 5.741 | 5.269 | 5.545 | 0.514 | -0.147 | -0.184 | -0.099 | -0.158 |
| 8 | 22 | 5.269 | 5.329 | 5.758 | 5.289 | 5.589 | 0.521 | -0.148 | -0.186 | -0.099 | -0.157 |
| 9 | 25 | 5.289 | 5.343 | 5.778 | 5.309 | 5.633 | 0.528 | -0.149 | -0.187 | -0.098 | -0.156 |
| Total | | | 5.940 | | 5.309 | | | -0.119 | | | |

Notes: Levels of s_t and H_{t+3} are obtained from numerically solving the dynamic model for three seasons of each grade as described by eq. (12) and parameter vector Ψ shown in table 3, and calculating the weighted average over 27 possible combinations of y^H , y^M and y^L for each type. Elasticities with respect to variance are obtained from numerically solving the structural model for a counterfactual income structure with no variance.

Table D.13: Effect of Rainfall Shocks on Consumption and Time Use

| | Household | | Child time | |
|-------------------------|--------------------------------|------------------------------|------------------------------|------------------------------|
| | Agr.inc. (1) | Cons.exp. (2) | Home prod. (3) | Studying (4) |
| Rainfall Quintile=1 | -0.497** (0.229) [0.233] | -0.067 (0.045) [0.048] | 0.448 (0.283) [0.302] | -0.285 (0.242) [0.250] |
| Rainfall Quintile=2 | -0.403** (0.194) [0.203] | 0.011 (0.033) [0.033] | -0.161 (0.229) [0.220] | -0.034 (0.198) [0.203] |
| Rainfall Quintile=4 | -0.074 (0.181) [0.183] | 0.051 (0.031) [0.033] | 0.175 (0.248) [0.238] | 0.136 (0.217) [0.242] |
| Rainfall Quintile=5 | 0.101 (0.201) [0.232] | 0.040 (0.038) [0.040] | -0.055 (0.292) [0.298] | 0.397* (0.225) [0.229] |
| Dependent variable mean | ln(9455) | ln(11274) | 2.69 | 6.10 |
| Observations | 15849 | 20445 | 68301 | 68301 |
| Adjusted R^2 | 0.062 | 0.264 | 0.324 | 0.177 |

Notes: Agricultural income and consumption expenditures are in log INR per capita. Child time allocation is in hours per day for three seasons. Each regression controls for state-by-round of interview fixed effects, for irrigated area (share of village area), and for lagged rainfall (in logs) in the last 3 years (2 years in col. (1)). Columns (3) and (4) also control for season-by-round, and age-by-gender fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the grid cell by round of interview. Conley (1999) standard errors (300km cut-off, Bartlett weights) in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.14: Effect of Rainfall on Consumption Expenditures per Capita (log)

| | (1) | (2) | (3) | (4) | (5) |
|---|--------------------------------|--------------------------------|--------------------------------|---------------------------------|--------------------------------|
| Precipitation, current year (1,000 mms) | 0.105* (0.053) [0.056] | 0.115 (0.122) [0.132] | | | |
| Irrigated area (share of village agr. area) | 0.172*** (0.034) [0.035] | 0.172*** (0.034) [0.035] | 0.170*** (0.034) [0.035] | 1.608*** (0.394) [0.396] | 0.907*** (0.352) [0.338] |
| Precipitation, current year (square) | | -0.004 (0.033) [0.033] | | | |
| Precipitation, current year (log mms) | | | 0.119* (0.063) [0.071] | 0.229*** (0.071) [0.076] | 0.138** (0.061) [0.061] |
| Precipitation \times Irrigated area | | | | -0.210*** (0.057) [0.057] | -0.116** (0.050) [0.047] |
| Dependent variable mean | ln(11274) | ln(11274) | ln(11274) | ln(11274) | ln(11274) |
| Observations | 20445 | 20445 | 20445 | 20445 | 20445 |
| Adjusted R^2 | 0.277 | 0.277 | 0.277 | 0.280 | 0.347 |

Notes: Consumption expenditures are in log INR per capita. Each regression controls for state-by-round of interview fixed effects, and for lagged rainfall (in logs) in the last 3 years. Column (5) also controls for village fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the grid cell by round of interview. Conley (1999) standard errors in brackets (cut-off 300km, Bartlett weights). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, according to Conley standard errors.

Table D.15: Effect of Rainfall on Study Time

| | (1) | (2) | (3) | (4) | (5) |
|---|--------------------------------|--------------------------------|--------------------------------|--------------------------------|---------------------------------|
| Precipitation, current year (1,000 mms) | 0.616** (0.325) [0.299] | 1.764*** (0.690) [0.651] | | | |
| Irrigated area (share of village agr. area) | 0.854*** (0.203) [0.204] | 0.844*** (0.202) [0.204] | 0.850*** (0.201) [0.202] | 7.855*** (2.745) [2.792] | 9.427*** (2.277) [2.376] |
| Precipitation, current year (square) | | -0.424** (0.206) [0.191] | | | |
| Precipitation, current year (log mms) | | | 0.955*** (0.357) [0.337] | 1.524*** (0.433) [0.416] | 1.333*** (0.375) [0.380] |
| Precipitation \times Irrigated area | | | | -1.028** (0.393) [0.403] | -1.285*** (0.324) [0.336] |
| Dependent variable mean | 6.10 | 6.10 | 6.10 | 6.10 | 6.10 |
| Observations | 68301 | 68301 | 68301 | 68301 | 68301 |
| Adjusted R^2 | 0.176 | 0.177 | 0.177 | 0.179 | 0.245 |

Notes: Time spent studying is in hours per day for three seasons. Each regression controls for state-by-round of interview fixed effects, for age-by-gender fixed effects, for season-by-round fixed effects, and for lagged rainfall (in logs) in the last 3 years. Column (5) also controls for village fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the grid cell by round of interview. Conley (1999) standard errors in brackets (cut-off 300km, Bartlett weights). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, according to Conley standard errors.

Table D.16: Effect of Rainfall on Child Time Use – No village FE

| Hours per day: | On-farm | | Chores | | Studying | | Leisure | |
|---|-------------------|-------------------|--------------------|----------------------|---------------------|---------------------|-------------------|-------------------|
| | Boys (1) | Girls (2) | Boys (3) | Girls (4) | Boys (5) | Girls (6) | Boys (7) | Girls (8) |
| Precipitation, current year (log mms)s | -0.796 (0.506) | -0.478 (0.298) | -0.168 (0.520) | -0.678 (0.485) | 1.186*** (0.433) | 1.872*** (0.479) | 0.276 (0.603) | -0.203 (0.503) |
| Irrigated area (share of village agr. area) | -0.968 (1.892) | -0.543 (1.199) | -4.385* (2.625) | -7.303*** (2.776) | 6.109** (2.557) | 9.517*** (3.483) | 2.406 (2.465) | 0.975 (2.469) |
| Precipitation \times Irrigated area | 0.124 (0.266) | 0.072 (0.168) | 0.554 (0.380) | 0.972** (0.403) | -0.771** (0.368) | -1.274** (0.506) | -0.357 (0.358) | -0.128 (0.356) |
| Dep. Var. Mean | 0.35 | 0.24 | 1.92 | 2.94 | 6.44 | 5.71 | 11.27 | 11.11 |
| Observations | 37005 | 31296 | 37005 | 31296 | 37005 | 31296 | 37005 | 31296 |
| Adjusted R^2 | 0.112 | 0.096 | 0.252 | 0.356 | 0.150 | 0.209 | 0.467 | 0.456 |

Notes: Time allocation is reported in hours per day over three seasons. Each regression controls for state-by-round of interview, season-by-round, and age fixed effects, and for lagged rainfall (in logs) in the last 3 years. Conley (1999) standard errors in parentheses (cut-off 300km, Bartlett weights). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.17: Effect of Risk on School Attendance (with village FE in risk calculation)

| | (1) | (2) | (3) | (4) | (5) |
|----------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|
| Panel A: | | | | | |
| IQR of predicted log consumption | -2.087 (0.651) [0.143] | -2.420 (0.685) [0.139] | -2.404 (0.684) [0.141] | -2.361 (0.627) [0.125] | -1.807 (0.642) [0.187] |
| Treatment Mean | 0.025 | 0.025 | 0.025 | 0.025 | 0.025 |
| Treatment Std. Dev. | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 |
| Adjusted R^2 | 0.207 | 0.209 | 0.209 | 0.228 | 0.240 |
| Panel B: | | | | | |
| SD of predicted log consumption | -2.448 (0.835) [0.186] | -2.836 (0.901) [0.195] | -2.815 (0.900) [0.199] | -2.803 (0.818) [0.180] | -2.113 (0.816) [0.219] |
| Treatment Mean | 0.019 | 0.019 | 0.019 | 0.019 | 0.019 |
| Treatment Std. Dev. | 0.013 | 0.013 | 0.013 | 0.013 | 0.013 |
| Adjusted R^2 | 0.207 | 0.208 | 0.209 | 0.228 | 0.239 |
| Rainfall controls | N | Y | Y | Y | Y |
| Rainfall by schoolage | N | N | Y | Y | Y |
| Income, wealth, area, hh size | N | N | N | Y | Y |
| State-by-round FE | N | N | N | N | Y |
| Dep. var. mean | 0.762 | 0.762 | 0.762 | 0.762 | 0.762 |
| Observations | 68301 | 68301 | 68301 | 68301 | 68301 |

Notes: In this table, risk is calculated with the estimates of column (4), table 5. Each regression controls for village, season-by-round of interview, and age-by-gender fixed effects. Rainfall controls are current and lagged rainfall in past 3 years (in log mms), and the standard deviation of rainfall over the last three years. Rainfall by schoolage additionally interacts each rainfall lag with a dummy that equals 1 if the child was 6 years or older in that year. Income, wealth and area are measured in per capita terms (and are in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.18: Effect of Risk on Study Time

| | School children | | | | | | | | | |
|----------------------------------|--------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| IQR of predicted log consumption | -8.493** (2.799) [0.049] | -8.700** (2.803) [0.037] | -8.640** (2.804) [0.037] | -8.102** (2.772) [0.046] | -7.370** (2.381) [0.033] | -1.200 (2.744) [0.749] | -1.126 (2.651) [0.760] | -1.112 (2.653) [0.764] | -1.057 (2.662) [0.780] | -1.232 (2.118) [0.656] |
| Adjusted R^2 | 0.221 | 0.223 | 0.223 | 0.242 | 0.264 | 0.219 | 0.221 | 0.221 | 0.222 | 0.279 |
| SD of predicted log consumption | -9.750* (3.488) [0.066] | -9.826*** (3.511) [0.073] | -9.742*** (3.513) [0.072] | -9.067*** (3.484) [0.088] | -8.471*** (3.025) [0.052] | -1.284 (3.566) [0.804] | -1.171 (3.387) [0.824] | -1.151 (3.391) [0.829] | -1.048 (3.412) [0.837] | -1.416 (2.722) [0.716] |
| Adjusted R^2 | 0.221 | 0.223 | 0.223 | 0.242 | 0.264 | 0.219 | 0.221 | 0.221 | 0.222 | 0.279 |
| Rainfall controls | N | Y | Y | Y | Y | N | Y | Y | Y | Y |
| Rainfall by schoollage | N | N | Y | Y | Y | N | N | Y | Y | Y |
| Income, wealth, area, hh size | N | N | N | Y | Y | N | N | N | Y | Y |
| State-by-round FE | N | N | N | N | Y | N | N | N | N | Y |
| Dep. var. mean | 6.10 | 6.10 | 6.10 | 6.10 | 6.10 | 8.01 | 8.01 | 8.01 | 8.01 | 8.01 |
| Observations | 68301 | 68301 | 68301 | 68301 | 68301 | 54096 | 54096 | 54096 | 54096 | 54096 |

Notes: Each regression controls for village, season-by-round of interview, and age-by-gender fixed effects. Rainfall controls are current rainfall and lagged rainfall in the past 3 years (in log mm), and the standard deviation of rainfall over the last three years. Rainfall by schoollage additionally interacts each lag with a dummy that equals 1 of the child was 6 years or older in that year. Income, wealth and area are measured in per capita terms (and are in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.19: Effects of Risk on School Attendance by Gender and Agegroup

| | Boys | | | Girls | | | Age 6 -10 | | Age 11 -15 | |
|----------------------------------|------------------------------|------------------------------|--------------------------------|--------------------------------|------------------------------|------------------------------|--------------------------------|--------------------------------|------------|--|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | | |
| IQR of predicted log consumption | -0.768 (0.351) [0.138] | -0.700 (0.370) [0.205] | -1.283** (0.313) [0.018] | -1.160** (0.333) [0.039] | -0.927* (0.320) [0.06] | -0.797 (0.355) [0.133] | -1.025** (0.315) [0.029] | -1.007* (0.328) [0.052] | | |
| 0.141 Adjusted R^2 | 0.189 | 0.201 | 0.281 | 0.294 | 0.236 | 0.247 | 0.247 | 0.261 | | |
| SD of predicted log consumption | -0.875 (0.448) [0.184] | -0.792 (0.472) [0.243] | -1.422** (0.400) [0.034] | -1.285* (0.419) [0.053] | -0.981 (0.410) [0.108] | -0.812 (0.452) [0.230] | -1.235* (0.406) [0.057] | -1.244** (0.413) [0.050] | | |
| Adjusted R^2 | 0.188 | 0.201 | 0.281 | 0.294 | 0.235 | 0.247 | 0.247 | 0.261 | | |
| Rainfall controls | Y | Y | Y | Y | Y | Y | Y | Y | | |
| Rainfall by schoolage | Y | Y | Y | Y | Y | Y | Y | Y | | |
| Income, wealth, area, hh size | Y | Y | Y | Y | Y | Y | Y | Y | | |
| State-by-round FE | N | Y | N | Y | N | Y | N | Y | | |
| Dep. var. mean | 0.802 | 0.802 | 0.715 | 0.715 | 0.763 | 0.763 | 0.761 | 0.761 | | |
| Observations | 37005 | 37005 | 31296 | 31296 | 35625 | 35625 | 32676 | 32676 | | |

Notes: Each regression controls for village, season-by-round of interview, and age-by-gender fixed effects. Rainfall controls are current rainfall and lagged rainfall in the past 3 years (in log mms), and the standard deviation of rainfall over the last three years. Rainfall by schoolage additionally interacts each lag with a dummy that equals 1 of the child was 6 years or older in that year. Income, wealth and area are measured in per capita terms (and are in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.20: Effect of Risk on School Infrastructure

| TREATMENT VARIABLE: | IQR of exp. cons | | SD of exp. cons | |
|---|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | (1) | (2) | (3) | (4) |
| Panel A: Village-level outcomes | | | | |
| Primary school in village (=1) | -0.845 (0.969) [0.428] | 0.317 (1.009) [0.760] | -1.174 (1.131) [0.362] | 0.285 (1.206) [0.810] |
| Middle school in village (=1) | -1.210 (1.120) [0.321] | -1.341 (1.146) [0.307] | -1.273 (1.395) [0.381] | -1.345 (1.432) [0.395] |
| Lower secondary school in village (=1) | -1.091 (0.918) [0.281] | -0.891 (0.951) [0.380] | -1.504 (1.123) [0.234] | -1.257 (1.160) [0.307] |
| Higher secondary school in village (=1) | -0.036 (0.727) [0.952] | 0.169 (0.738) [0.814] | -0.204 (0.909) [0.829] | 0.106 (0.911) [0.906] |
| Village and round FE | Y | Y | Y | Y |
| State-by-round FE | N | Y | N | Y |
| Observations | 710 | 710 | 710 | 710 |
| Panel B: Child-level outcomes | | | | |
| School attendance (=1) | -1.084** (0.284) [0.024] | -1.028** (0.297) [0.038] | -1.255** (0.368) [0.033] | -1.192** (0.379) [0.044] |
| Village and round FE | Y | Y | Y | Y |
| Village & hh controls | Y | Y | Y | Y |
| School controls | Y | Y | Y | Y |
| State-by-round FE | N | Y | N | Y |
| Observations | 67485 | 67485 | 67485 | 67485 |

Notes: Each coefficient relates to an individual regression. The unit of observation is the village-by-round (Panel A), and the child (Panel B). Each regression controls for village and round of interview fixed effects. Cols. (2) and (4) add state-by-round of interview fixed effects. Child-level regressions additionally control for season-by-round of interview, and age-by-gender fixed effects, and for current rainfall and for lagged rainfall (in log mms) in the past 3 years. Each lag is additionally interacted with a dummy that equals 1 if the child was 6 years or older in that year. Child regressions also control for income, wealth and area (all measured in per capita terms, in logs), for household size, and for the presence of all four school types in the village (primary, middle, lower secondary and higher secondary). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile Bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.21: Effect of Risk on Village Characteristics

| TREATMENT VARIABLE: | IQR of pred. log cons | | SD of pred. log cons | |
|---|-------------------------------|--------------------------------|--------------------------------|---------------------------------|
| | (1) | (2) | (3) | (4) |
| Agricultural income (share of village income) | 0.138 (0.417) [0.756] | 0.313 (0.396) [0.468] | 0.318 (0.588) [0.634] | 0.579 (0.538) [0.357] |
| Share of households involved in agr. | -0.529 (0.540) [0.384] | -0.664 (0.452) [0.182] | -0.500 (0.722) [0.532] | -0.687 (0.608) [0.293] |
| Any factory in village | 0.082 (1.094) [0.927] | -0.407 (1.071) [0.690] | 0.047 (1.430) [0.973] | -0.681 (1.365) [0.619] |
| No. of villagers employed in factories | 86.030 (94.410) [0.394] | 107.830 (96.752) [0.317] | 83.880 (131.052) [0.579] | 107.467 (132.638) [0.479] |
| Any cooperative in village | -0.019 (1.127) [0.985] | 1.142 (1.060) [0.296] | -0.176 (1.438) [0.900] | 1.466 (1.335) [0.293] |
| Any bank in village | -0.508 (1.484) [0.739] | -0.675 (1.392) [0.626] | -1.121 (1.872) [0.630] | -1.142 (1.751) [0.584] |
| Village and round FE | Y | Y | Y | Y |
| State-by-round FE | N | Y | N | Y |
| Observations | 713 | 713 | 713 | 713 |

Notes: The unit of observation is the village-by-round. Each coefficient relates to an individual regression. Each regression controls for village and round of interview fixed effects. Cols. (2) and (4) add state-by-round of interview fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile Bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.22: Effect of Rainfall on Outcomes of Interest – Cash Savings

| | Household | Maternal time | |
|---|----------------------|-------------------|-------------------|
| | Cons. exp. | Household chores | Labor market |
| | (1) | (2) | (3) |
| Precipitation, current year (log mms) | 0.488*** (0.124) | 1.475 (0.936) | -0.895 (0.971) |
| Cash savings (per capita, log) | 0.432*** (0.088) | 0.969* (0.585) | -0.175 (0.627) |
| Precipitation \times Cash savings | -0.042*** (0.012) | -0.118 (0.083) | 0.007 (0.090) |
| Irrigated area (share of village agr. area) | 0.469 (0.361) | 0.678 (2.918) | -4.728 (3.146) |
| Precipitation \times Irrigated area | -0.056 (0.051) | -0.045 (0.421) | 0.720 (0.451) |
| Observations | 15693 | 27009 | 27009 |
| Adjusted R^2 | 0.592 | 0.499 | 0.569 |

Notes: Consumption expenditures are in log INR per capita. Female time allocation is in hours per day for three seasons. Each regression controls for household fixed effects, for state-by-round of interview fixed effects, and for lagged rainfall (in logs) in the last 3 years. Columns (2) and (3) also control for season-by-round, and age fixed effects. Conley (1999) standard errors in parentheses (cut-off 300km, Bartlett weights). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.23: Effect of Risk (calculated at household-level) on School Attendance

| | (1) | (2) | (3) | (4) |
|----------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Panel A: | | | | |
| IQR of predicted log consumption | -5.392*** (1.032) [0.000] | -5.343*** (1.035) [0.001] | -4.295*** (1.122) [0.000] | -4.470*** (1.137) [0.000] |
| Adjusted R^2 | 0.573 | 0.573 | 0.573 | 0.571 |
| Panel B: | | | | |
| SD of predicted log consumption | -7.477*** (1.268) [0.000] | -7.412*** (1.274) [0.000] | -6.132*** (1.403) [0.001] | -6.357*** (1.415) [0.001] |
| Adjusted R^2 | 0.573 | 0.573 | 0.573 | 0.572 |
| Rainfall controls | Y | Y | Y | Y |
| Rainfall by schoolage | N | Y | Y | Y |
| Income, wealth, area, hh size | N | N | Y | Y |
| Initial savings-by-round | N | N | N | Y |
| Dep. var. mean | 0.762 | 0.762 | 0.762 | 0.762 |
| Observations | 55455 | 55455 | 55455 | 48546 |

Notes: Each regression controls for household, village-by-round of interview, season-by-round, and age-by-gender fixed effects, and for household-specific linear time trends. Rainfall by schoolage interacts each rainfall lag with a dummy that equals 1 if the child was 6 years or older in that year. Income, wealth and area are measured in per capita terms (and are in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the household by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.24: Effect of Risk (calculated at household-level) on Economic Outcomes

| | Income | Land | Wealth | Cons. exp | Agr. prod. (=1) | Mat. labor supply |
|----------------------------------|------------------------------|--------------------------------|---------------------------------|-------------------------------|--------------------------------|--------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| PANEL A: | | | | | | |
| IQR of predicted log consumption | -7.156 (2.730) [0.587] | -36.168* (4.631) [0.058] | -43.931** (1.829) [0.015] | -9.223* (0.825) [0.062] | -6.583** (0.819) [0.038] | 21.128* (12.082) [0.144] |
| Observations | 13701 | 13701 | 13701 | 13701 | 13701 | 23472 |
| Adjusted R^2 | 0.337 | 0.654 | 0.733 | 0.723 | 0.424 | 0.730 |
| PANEL B: | | | | | | |
| SD of predicted log consumption | -9.369 (3.417) [0.557] | -44.862* (5.842) [0.076] | -56.254** (2.324) [0.019] | -11.661 (1.093) [0.109] | -8.364** (1.068) [0.048] | 22.794* (15.014) [0.087] |
| Observations | 13701 | 13701 | 13701 | 13701 | 13701 | 23472 |
| Adjusted R^2 | 0.337 | 0.653 | 0.733 | 0.723 | 0.424 | 0.730 |

Notes: Income, land owned, wealth (value of all assets owned), and consumptions expenditures are in per capita terms and logs. Maternal labor supply is in hours per day for three seasons. Each regression controls for household and village-by-round of interview fixed effects, and for household-specific linear time trends. Col. (6) additionally controls for age and for season-by-round of interview fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the household by round of interview. Percentile Bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.25: Effect of Risk on School Attendance - Panel Households

| | (1) | (2) | (3) | (4) | (5) |
|----------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|-------------------------------|
| Panel A: | | | | | |
| IQR of predicted log consumption | -1.067** (0.285) [0.030] | -1.119** (0.278) [0.023] | -1.104** (0.279) [0.024] | -1.075** (0.271) [0.021] | -0.967* (0.311) [0.062] |
| Adjusted R^2 | 0.434 | 0.435 | 0.435 | 0.437 | 0.443 |
| Panel B: | | | | | |
| SD of predicted log consumption | -1.287** (0.367) [0.045] | -1.344** (0.359) [0.035] | -1.323** (0.359) [0.039] | -1.282** (0.352) [0.033] | -1.155* (0.397) [0.077] |
| Adjusted R^2 | 0.434 | 0.435 | 0.435 | 0.437 | 0.443 |
| Rainfall controls | N | Y | Y | Y | Y |
| Rainfall by schoolage | N | N | Y | Y | Y |
| Income, wealth, area, hh size | N | N | N | Y | Y |
| State-by-round FE | N | N | N | N | Y |
| Dep. Var. mean | 0.762 | 0.762 | 0.762 | 0.762 | 0.762 |
| Observations | 56922 | 56922 | 56922 | 56922 | 56922 |

Notes: Each regression controls for household, season-by-round of interview, and age-by-gender fixed effects. Rainfall controls are current and lagged rainfall in past 3 years (in log mms), and the standard deviation of rainfall over the last three years. Rainfall by schoolage additionally interacts each rainfall lag with a dummy that equals 1 if the child was 6 years or older in that year. Income, wealth and area are measured in per capita terms (and are in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.26: Robustness of Effects to Differential Time Trends

| | (1) | (2) | (3) | (4) | (5) |
|---|------------------------------|-------------------------------|--------------------------------|--------------------------------|--------------------------------|
| Panel A: | | | | | |
| IQR of predicted log consumption | -0.613 (0.319) [0.149] | -0.729* (0.283) [0.071] | -0.929** (0.279) [0.030] | -0.834** (0.283) [0.043] | -0.893** (0.293) [0.050] |
| Observations | 66708 | 66708 | 66708 | 66708 | 66708 |
| Adjusted R^2 | 0.209 | 0.220 | 0.211 | 0.212 | 0.210 |
| Panel B: | | | | | |
| SD of predicted log consumption | -0.635 (0.401) [0.234] | -0.886* (0.350) [0.077] | -1.074* (0.357) [0.052] | -0.954* (0.360) [0.063] | -1.053* (0.367) [0.058] |
| Observations | 66708 | 66708 | 66708 | 66708 | 66708 |
| Adjusted R^2 | 0.209 | 0.220 | 0.211 | 0.212 | 0.210 |
| Irrigation (village mean, 1982) \times round FE | Y | N | N | N | N |
| Hh head literate (village mean, 1982) \times round FE | N | Y | N | N | N |
| Log hh income p.c. (village mean, 1982) \times round FE | N | N | Y | N | N |
| Log hh wealth p.c. (village mean, 1982) \times round FE | N | N | N | Y | N |
| Log hh area p.c. (village mean, 1982) \times round FE | N | N | N | N | Y |

Notes: Each regression controls for village, season-by-round of interview, and age-by-gender fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the grid-cell by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.27: Effect of Risk on School Attendance - Placebo Tests

| | (1) | (2) | (3) | (4) |
|---|------------------------------|-----------------------------|------------------------------|-----------------------------|
| IQR of predicted log consumption, current | -0.901 (0.418) [0.159] | | | |
| IQR of predicted log consumption, lead | | 0.442 (0.418) [0.452] | | |
| SD of predicted log consumption, current | | | -1.042 (0.509) [0.205] | |
| SD of predicted log consumption, lead | | | | 0.539 (0.573) [0.536] |
| Observations | 41886 | 41886 | 41886 | 41886 |
| Adjusted R^2 | 0.241 | 0.240 | 0.241 | 0.240 |

Sample: REDS 1981-82 and 1998-99. Notes: Each regression controls for village, season-by-round of interview, and age-by-gender fixed effects, and for current rainfall and lagged rainfall in the past 3 years (in log mms), and the standard deviation of rainfall over the last three years. Each lag of rainfall is additionally interacted with a dummy that equals 1 if the child was 6 years or older in that year. Additional controls are income, wealth and area are measured in per capita terms (and in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.28: Effect of Risk on School Attendance – Weather Controls

| | (1) | (2) | (3) | (4) | (5) |
|--|----------|---------|----------|----------|---------|
| Panel A: | | | | | |
| IQR of predicted log consumption | -0.976** | -0.821* | -0.960** | -0.947** | -1.007* |
| | (0.279) | (0.285) | (0.286) | (0.296) | (0.291) |
| | [0.028] | [0.061] | [0.040] | [0.046] | [0.058] |
| Observations | 68301 | 68301 | 68301 | 68301 | 68301 |
| Adjusted R^2 | 0.227 | 0.229 | 0.231 | 0.233 | 0.234 |
| Panel B: | | | | | |
| SD of predicted log consumption | -1.106** | -0.916* | -1.123* | -1.088* | -1.162* |
| | (0.358) | (0.359) | (0.363) | (0.374) | (0.371) |
| | [0.044] | [0.087] | [0.056] | [0.057] | [0.076] |
| Observations | 68301 | 68301 | 68301 | 68301 | 68301 |
| Adjusted R^2 | 0.227 | 0.229 | 0.231 | 0.232 | 0.234 |
| Rainfall controls | Y | Y | Y | Y | Y |
| Rainfall controls (square) | N | Y | Y | Y | Y |
| Temperature controls | N | N | Y | Y | Y |
| Temperature controls (square) | N | N | N | Y | Y |
| Rainfall \times Temperature controls | N | N | N | N | Y |

Notes: Each regression controls for village, season-by-round of interview, and age-by-gender fixed effects. Rainfall controls are current rainfall and lagged rainfall (in log mms) in the past 3 years. Temperature controls are current temperature and lagged temperature (Degree Celsius) in the past 3 years. Each regression also controls for income, wealth and area (all measured in per capita terms, in logs) and for household size. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile Bootstrap-t p-values (null-imposed, 1,000 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.29: Effect of Risk on Rainfall Outcomes

| TREATMENT VARIABLE: | IQR of pred. cons | | SD of pred. cons | |
|---------------------------------------|-----------------------------|------------------------------|-----------------------------|------------------------------|
| | (1) | (2) | (3) | (4) |
| Dep. var.: Annual Rainfall, log (t+3) | 0.572 (0.754) [0.493] | -0.399 (0.554) [0.474] | 0.696 (0.976) [0.519] | -0.595 (0.694) [0.414] |
| Dep var.: Annual Rainfall, log (t+2) | 0.549 (0.712) [0.453] | 0.011 (0.513) [0.989] | 0.875 (0.893) [0.375] | -0.073 (0.664) [0.912] |
| Dep var.: Annual Rainfall, log (t+1) | 0.890 (0.659) [0.218] | 1.008 (0.529) [0.119] | 1.428 (0.876) [0.176] | 1.448 (0.719) [0.114] |
| Dep var.: Annual Rainfall, log (t) | 0.902 (0.583) [0.136] | -0.061 (0.481) [0.899] | 1.132 (0.726) [0.222] | -0.121 (0.598) [0.837] |
| Dep var.: Annual Rainfall, log (t-1) | 0.022 (0.619) [0.965] | -0.354 (0.527) [0.535] | 0.143 (0.757) [0.867] | -0.496 (0.661) [0.503] |
| Dep var.: Annual Rainfall, log (t-2) | 0.460 (0.723) [0.540] | -0.689 (0.485) [0.186] | 0.752 (0.977) [0.513] | -0.895 (0.601) [0.159] |
| Dep var.: Annual Rainfall, log (t-3) | 0.282 (1.068) [0.795] | 0.141 (0.508) [0.780] | 0.584 (1.518) [0.713] | 0.188 (0.694) [0.780] |
| Village and round FE | Y | Y | Y | Y |
| State-by-round FE | N | Y | N | Y |
| Observations | 718 | 718 | 718 | 718 |

Notes: The unit of observation is the village-by-round. Each coefficient relates to an individual regression. Each regression controls for village and round of interview fixed effects. Cols. (2) and (4) add state-by-round of interview fixed effects. Standard errors (in parentheses) are corrected to allow for clustering by grid cell by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.30: Effect of Risk on School Attendance by Season

| | Season 1 (Oct./Nov.) | | Season 2 (February) | | Season 3 (April/May) | |
|----------------------------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A: | | | | | | |
| IQR of predicted log consumption | -1.063** (0.289) [0.021] | -0.953* (0.309) [0.055] | -0.881** (0.295) [0.047] | -0.836* (0.313) [0.079] | -0.966** (0.297) [0.044] | -0.884* (0.314) [0.069] |
| Adjusted R^2 | 0.220 | 0.232 | 0.230 | 0.244 | 0.224 | 0.236 |
| Panel B: | | | | | | |
| SD of predicted log consumption | -1.193** (0.374) [0.046] | -1.083* (0.393) [0.066] | -0.985* (0.380) [0.068] | -0.937 (0.402) [0.123] | -1.115** (0.379) [0.046] | -1.003* (0.397) [0.075] |
| Adjusted R^2 | 0.220 | 0.232 | 0.230 | 0.244 | 0.223 | 0.236 |
| Rainfall by schoolage | Y | Y | Y | Y | Y | Y |
| Income, wealth, area, hh size | Y | Y | Y | Y | Y | Y |
| State-by-round FE | N | Y | N | Y | N | Y |
| Dep. Var. mean | 0.769 | 0.769 | 0.762 | 0.762 | 0.755 | 0.755 |
| Observations | 22767 | 22767 | 22767 | 22767 | 22767 | 22767 |

Notes: Each regression controls for village, season-by-round of interview, and age-by-gender fixed effects, for current and lagged rainfall in past 3 years (in log mms), and the standard deviation of rainfall over the last three years. Rainfall controls are additionally interacted with a dummy that equals 1 if the child was 6 years or older in that year. Each regression also controls for income, wealth and area, measured in per capita terms (and in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.31: Effect of Risk on School Attendance by Quartiles of Consumption, Income and Wealth

| Quartiles of: | Cons.exp. (1) | Income (2) | Wealth (3) |
|--|--------------------------------|--------------------------------|--------------------------------|
| PANEL A: | | | |
| IQR of predicted log consumption× 1st Quartile | -1.005* (0.386) [0.055] | -0.996* (0.353) [0.052] | -1.323* (0.408) [0.027] |
| IQR of predicted log consumption× 2nd Quartile | -0.889** (0.349) [0.082] | -1.195** (0.363) [0.029] | -1.064* (0.363) [0.063] |
| IQR of predicted log consumption× 3rd Quartile | -0.891* (0.347) [0.077] | -0.776 (0.341) [0.129] | -0.955* (0.352) [0.070] |
| IQR of predicted log consumption× 4th Quartile | -0.686 (0.398) [0.214] | -0.785 (0.381) [0.150] | -0.460 (0.359) [0.357] |
| Adjusted R^2 | 0.241 | 0.240 | 0.240 |
| PANEL B: | | | |
| SD of predicted log consumption× 1st Quartile | -1.138 (0.500) [0.103] | -1.101* (0.455) [0.075] | -1.379** (0.527) [0.049] |
| SD of predicted log consumption× 2nd Quartile | -0.858 (0.442) [0.148] | -1.263* (0.467) [0.066] | -1.189* (0.469) [0.090] |
| SD of predicted log consumption× 3rd Quartile | -1.088* (0.441) [0.080] | -0.958 (0.438) [0.123] | -1.148* (0.438) [0.075] |
| SD of predicted log consumption× 4th Quartile | -0.846 (0.498) [0.209] | -0.925 (0.474) [0.156] | -0.593 (0.446) [0.320] |
| Adjusted R^2 | 0.241 | 0.240 | 0.239 |
| Educ. exp. (per child)/ income p.c. (1st Quartile) | 15.18% | 31.45% | 13.62% |
| Educ. exp. (per child)/ income p.c. (2nd Quartile) | 9.93% | 12.96% | 11.38% |
| Educ. exp. (per child)/ income p.c. (3rd Quartile) | 9.45% | 10.08% | 8.92% |
| Educ. exp. (per child)/ income p.c. (4th Quartile) | 6.73% | 5.03% | 6.66% |
| Observations | 68301 | 68301 | 68301 |

Notes: Each regression controls for village, for state-by-round of interview, for season-by-round of interview, and age-by-gender fixed effects, and for current and lagged rainfall in past 3 years (in log mms), as well as for income, wealth and area are measured in per capita terms (and are in logs). Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.32: Effect of Risk on Returns to Education for Men and Women

| | Wages (men) | | Wages (women) | |
|----------------------------------|----------------------|----------------------|---------------------|--------------------------------|
| | (1) | (2) | (3) | (4) |
| PANEL A: | | | | |
| IQR of predicted log consumption | 0.259 (0.782) | 0.368 (0.910) | -0.113 (1.032) | 2.573 (1.991) |
| Highest degree: Illiterate | -0.229*** (0.022) | -0.215*** (0.028) | -0.094* (0.046) | -0.005 (0.079) |
| Highest degree: Primary | -0.105*** (0.021) | -0.090** (0.028) | -0.038 (0.047) | 0.015 (0.080) |
| Highest degree: Higher secondary | 0.278*** (0.035) | 0.241*** (0.047) | 0.176* (0.082) | 0.155 (0.125) |
| Highest degree: University | 0.792*** (0.063) | 0.759*** (0.071) | 1.093*** (0.142) | 1.150*** (0.223) |
| Illiterate \times IQR | | -0.433 (0.757) | | -3.275 ⁺ (1.908) |
| Primary \times IQR | | -0.500 (0.690) | | -2.206 (2.063) |
| Higher secondary \times IQR | | 1.300 (1.368) | | 0.571 (3.908) |
| University \times IQR | | 1.177 (2.183) | | -2.162 (7.582) |
| R^2 | 0.403 | 0.404 | 0.455 | 0.457 |
| PANEL B: | | | | |
| SD of predicted log consumption | 0.509 (0.991) | 0.798 (1.187) | -0.168 (1.376) | 3.972 (2.791) |
| Highest degree: Illiterate | -0.230*** (0.021) | -0.211*** (0.028) | -0.094* (0.046) | 0.010 (0.081) |
| Highest degree: Primary | -0.105*** (0.021) | -0.085** (0.029) | -0.038 (0.047) | 0.021 (0.083) |
| Highest degree: Higher secondary | 0.277*** (0.035) | 0.235*** (0.048) | 0.176* (0.082) | 0.158 (0.123) |
| Highest degree: University | 0.791*** (0.063) | 0.757*** (0.073) | 1.093*** (0.142) | 1.244*** (0.233) |
| Illiterate \times SD | | -0.778 (0.972) | | -4.977 ⁺ (2.623) |
| Primary \times SD | | -0.884 (0.923) | | -3.171 (2.804) |
| Higher secondary \times SD | | 1.902 (1.884) | | 0.540 (5.195) |
| University \times SD | | 1.563 (2.929) | | -7.348 (10.275) |
| Observations | 16547 | 16547 | 5180 | 5180 |
| R^2 | 0.403 | 0.404 | 0.455 | 0.458 |

Notes: Sample consists of men and women aged 15 to 64, with non-missing wages (2007-08 round of interviews only). Dep. var.: Log daily wage (averaged over all activities). Omitted category: Lower secondary. Each regression controls for state and season fixed effects, for age fixed effects, marital status, and for current rainfall and lagged rainfall in the past 3 years (in log mms), and the standard deviation of rainfall over the last three years. Standard errors (in parentheses) are corrected to allow for clustering at the village by round of interview. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.33: Effect of Risk on School Attendance of Girls – Marriage Market Risk

| | (1) | (2) | (3) | (4) |
|----------------------------------|--------------------------------|--------------------------------|-------------------------------|-------------------------------|
| PANEL A: | | | | |
| IQR of predicted log consumption | -1.238** (0.314) [0.017] | -1.138** (0.314) [0.020] | -1.094* (0.335) [0.068] | -1.063* (0.332) [0.063] |
| Average IQR in marriage market | | -0.635 (0.388) | | -0.664* (0.403) |
| Adjusted R^2 | 0.279 | 0.280 | 0.293 | 0.293 |
| PANEL B: | | | | |
| SD of predicted log consumption | -1.368** (0.401) [0.043] | -1.246** (0.402) [0.050] | -1.208* (0.420) [0.091] | -1.163* (0.419) [0.094] |
| Average SD in marriage market | | -0.732 (0.495) | | -0.772 (0.511) |
| Adjusted R^2 | 0.279 | 0.279 | 0.293 | 0.293 |
| State-by-round FE | N | N | Y | Y |
| Observations | 30909 | 30909 | 30909 | 30909 |

Notes: Each regression controls for village, for season-by-round of interview, and age-by-gender fixed effects, and for current rainfall and lagged rainfall (in log mms) in the past 3 years. Each lag is additionally interacted with a dummy that equals 1 of the child was 6 years or older in that year. Each regression also controls for income, wealth and area (all measured in per capita terms, in logs), and for household size. Cols. (3) and (4) add state-by-round of interview fixed effects. Villages belonging to the same marriage market are defined as villages located in a 67 km radius around village of interest. When calculating average risk in these villages, we exclude the village of interest itself. Sample: Girls aged 6 to 15 living in a village that has at least one more REDS village in a 67 km radius. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile Bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.34: Effect of Risk on Fertility and Gender-Composition

| | No. of children born | | Boys per girl | |
|----------------------------------|-----------------------------|-----------------------------|------------------------------|------------------------------|
| | (1) | (2) | (3) | (4) |
| PANEL A: | | | | |
| IQR of predicted log consumption | 0.014 (1.337) [0.999] | 1.907 (1.314) [0.317] | -0.283 (1.642) [0.849] | -1.936 (1.675) [0.243] |
| Adjusted R^2 | 0.429 | 0.446 | 0.055 | 0.093 |
| PANEL B: | | | | |
| SD of predicted log consumption | 0.715 (1.681) [0.782] | 2.930 (1.658) [0.264] | -0.257 (2.132) [0.894] | -2.319 (2.176) [0.288] |
| Adjusted R^2 | 0.429 | 0.446 | 0.055 | 0.093 |
| State-by-round FE | N | Y | N | Y |
| Observations | 11125 | 11125 | 710 | 710 |

Notes: The unit of observation is the mother in cols. (1) and (2), and the village-by-round in cols. (3) and (4). Each regression controls for village and round of interview fixed effects, and for current rainfall and lagged rainfall (in log mms) in the past 3 years, and the standard deviation of rainfall over the past 3 years. Cols. (1) and (2) also control for age (of the mother) fixed effects, and cols. (2) and (4) add state-by-round of interview fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the village by round of interview. Percentile bootstrap-t p-values (null-imposed, 999 replications) are in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ according to bootstrap p-values.

Table D.35: Summary statistics (SEPRI)

| | Mean | SD |
|---|--------|----------|
| Consumption expenditures p.c. (log) | 9.34 | (0.58) |
| Precipitation, current year | 6.90 | (0.35) |
| Irrigated area (share of village agr. area) | 0.79 | (0.30) |
| Employment generated per capita, NREGS (lag) | 0.74 | (1.46) |
| SD of predicted log consumption | 0.013 | (0.014) |
| SD of predicted log consumption (Av. NREGS employment = 0) | 0.014 | (0.0036) |
| SD of predicted log consumption (Av. NREGS employment = 2.24) | 0.0085 | (0.0022) |
| Observations | 77,938 | |

Notes: Consumption expenditures are in constant July 2010 INR.

Table D.36: Effect of Rainfall on Consumption Expenditures per Capita (1982-2016)

| | (1) | (2) | (3) | (4) |
|--|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Precipitation, current year (log mms) | 0.192*** (0.063) [0.066] | 0.196*** (0.063) [0.066] | 0.136** (0.062) [0.059] | 0.136** (0.061) [0.059] |
| Irrigated area (share of village agr. area)) | 1.377*** (0.381) [0.394] | 1.367*** (0.381) [0.395] | 1.013*** (0.352) [0.329] | 0.998*** (0.349) [0.324] |
| Precipitation \times Irrigated area | -0.174*** (0.055) [0.058] | -0.173*** (0.055) [0.058] | -0.130*** (0.050) [0.046] | -0.127*** (0.049) [0.046] |
| Employment generated per capita, NREGS (lag) | | 0.306** (0.149) [0.124] | | 0.270* (0.149) [0.143] |
| Precipitation \times Employment generated, NREGS | | -0.042** (0.021) [0.017] | | -0.037* (0.021) [0.020] |
| Village FE | N | N | Y | Y |
| Observations | 98383 | 98383 | 98383 | 98383 |
| Adjusted R^2 | 0.296 | 0.296 | 0.349 | 0.349 |

Notes: Each regression controls for state-by-round of interview, and for lagged rainfall (in logs) in the last 3 years. Columns (3) and (4) also control for village fixed effects. Standard errors (in parentheses) are corrected to allow for clustering at the level of the grid cell by round of interview. Conley (1999) standard errors in brackets (cut-off 300km, Bartlett weights). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, according to Conley standard errors.