NBER WORKING PAPER SERIES

CONSUMPTION RISK AND HUMAN CAPITAL ACCUMULATION IN INDIA

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

Partial support provided by the Population Studies and Training Center at Brown University through an NIH Population Center Grant (R24 HD041020) and by the German Federal Ministry for Economic Development and Cooperation. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Consumption Risk and Human Capital Accumulation in India
Andrew D. Foster and Esther Gehrke
NBER Working Paper No. 24041
November 2017
JEL No. I25,O13,O15

ABSTRACT

We study the consequences of dynamic complementarities in the production of child human capital for the relationship between risk and schooling investment in a low income setting. In contrast to previous literature, we explore the ex ante response of schooling to risk. We develop a model that incorporates, dynamic complementarity in the education production function, ex-post labor market responses by mothers to income shocks, and substitutability between maternal and child work in the household. We test the model using data from rural India, focusing particularly on the schooling of girls. We find that risk substantially reduces the probability that girls attend school. We then 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
University of Goettingen
Department of Agriculture and Rural Economics
Platz der Goettingen Sieben 5
37073 Goettingen, Germany
esther.gehrke@uni-goettingen.de

1 Introduction

Accumulating human capital is an important path to participating in a country's growth process, escaping poverty, and achieving higher income and prosperity. However, we still lack conclusive knowledge of how human capital is formed, and what determines the investment in human capital, and schooling in particular, within the household.

It is by now well established that economic shocks affect children's time in school in the absence of well-functioning markets. Jacoby and Skoufias (1997) show in the context of India that children have to drop-out of school and work in case of severe economic shocks. Similar evidence was provided by Beegle et al. (2006) in the context of Tanzania and by Duryea et al. (2007) in the context of Brazil. A related strand of literature has established that children's school time falls in the aftermath of economic shocks because households lack the financial resources to send children to school (see e.g. Jensen, 2000; Skoufias and Parker, 2006; Gubert and Robilliard, 2008).

Much less attention has been devoted to the possible *ex ante* effects of risk on schooling decisions and human capital accumulation. Recent works make a strong argument for dynamic complementarities in the production of human capital (c.f. Cunha and Heckman, 2007; Cunha et al., 2010; Attanasio et al., 2015). Dynamic complementarity arises when school investments for a given child in different periods are complements in the production of human capital. In such a scenario, the returns to each year of school depend on past as well as on future investments in school. This relationship also implies that early stage investments in education have to be followed up by investments at later stages in order to make that early stage investment productive. However, the ability to follow-up on this investment may depend on the realization of shocks in the future. If future school time is not known, then 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. Under such circumstances, the cost of risk in terms of human capital accumulation would be considerably higher than the sum of the direct effects of shocks.

A reduction in school time as response to shocks might not only arise from child labor or lack in financial resources. Evidence suggests that households tend to increase their labor supply in response to adverse economic shocks (Kochar, 1999; Rose, 2001; Jayachandran, 2006). If women work more in the aftermath of economic shocks, they might have less time available for household activities. In the presence of intra-household substitution, for instance in the performance of household chores, increased female labor supply could have negative effects on the time allocation of their children. As women have less time available for home production and childcare, and such activities can only be foregone at high cost, they might be forced to take older children out of school or to cut down on the time they study at home in order for them to fill in for these tasks (Skoufias, 1993; Ilahi, 2000).[1] Given that women in India are responsible for most home-production activities, a labor supply response of mothers might in fact have larger intra-household repercussions than a

[1] Shah and Steinberg (2015) for example show that girls in rural India are more likely to drop out of school if their mothers' experience an increase in wages.

labor supply response of fathers.[2] This also implies that risk might affect the time children spend in school even in the absence of a child labor market for girls and in contexts of no school fees.

We study this question in the context in rural India and for girls. We focus on girls for three reasons. First, previous literature has emphasized the effect of shocks on boys' labor supply and time in school, as they are more likely to be working outside the home than girls. But while girls are not often found to be working outside the home, they are more likely to be performing household chores. Second, our data suggest that mothers respond to shocks by supplying more labor *ex post*, but not fathers. And given the gendered nature of work in India in general, and of housework in particular, girl and mother time are likely to be more closely linked than are boy and mother time, giving us an opportunity to look more closely at underlying mechanisms for any observed responses. Finally, and perhaps most importantly, we have no clear way, within the village, of distinguishing contemporary effects of risk and future effects of risk. Because of norms of patrilocal exogamy we cannot, for men, separately identify the current consumption-smoothing effects of risk on schooling from longer-term effects of risk on the returns to human capital. For women, however, who will typically leave the village for marriage, we can control for the latter effects by conditioning on risk at the level of aggregation of the marriage market (as in Foster and Rosenzweig, 1999) and thus isolate the contemporaneous effect.

This paper develops a model of household time allocation and human capital accumulation that highlights the effect of uncertainty regarding future consumption on a child's school time. Following Jacoby and Skoufias (1997), we model schooling decisions in a unitary household. In order to incorporate time allocation to home production, we assume that the household derives utility from two consumption goods, one of which is produced at home and the other on the market. We also assume that women increase their labor supply in order to cope with income shocks *ex post*. As both the mother and the daughter can perform household chores, shocks not only affect labor supply decisions of adult household members, but also reduce the time girls spend in school.

By incorporating two periods into the model we are able to show that the presence of risk reduces the time girls spend in school to the extent that marginal utility of second period schooling is increasing in first-period schooling. Whether this is true depends both on the degree of complementarity in the education production function and the extent of concavity in the preferences for child human capital. In particular, if household utility is linear in child-human capital then negative effects are observed except when child school times in different periods are perfect substitutes in the production of human capital. In contrast, if household utility is log-linear in human capital, the predicted effect of risk on school time is negative at high levels of dynamic complementarity but turns positive for elasticities of substitution greater than one.

[2] The India Time Use survey from 1998/99 revealed that women spend more time than men on activities other than leisure. While women spend 53.4 hours per week on salaried activities and household work combined, men only spend 45.6 hours per week on these activities (Government of India, 2000). Women spend on average 18.7 hours per week on wage work and 34.6 hours per week on household-related activities. These activities are classified as extended SNA activities in the Time Use Survey and include household maintenance, as well as care for children, sick and elderly.

We therefore explore the case for dynamic complementarity in study hours among children in rural India directly by estimating a production function of cognitive skills. In the absence of skills data in the Rural Economic and Demographic Survey (REDS), we use the Young Lives Survey (YLS) data from Andhra Pradesh, a state in the south of India. While Attanasio et al. (2015) have used these data for a similar study, that paper considers educational expenditures rather than study time. Using a non-linear instrumental variables procedure, we cannot reject an elasticity of substitution of one between lagged cognitive skills and current study hours in the production of current cognitive skills. This elasticity is sufficient, given our model, to produce a negative *ex ante* effect of risk on schooling if parental utility is linear in human capital but not, for example, if parental utility is log-linear in human capital.

We then test the predictions of our model with the REDS data from 2006. Our unit of observation is the mother-daughter pair for girls aged 6 to 16 who live with their mothers. We observe time allocation decision over three different agricultural seasons, which have a length of four months. This gives a final sample of 12,918 observations.

We predict consumption risk at the village level as a function of the historical rainfall distribution and a village's share of land that is under irrigation. Irrigation can mitigate some of the effects of adverse rainfall shocks on consumption, such that an increase in the availability of irrigation systems in a village contributes to a reduction in the variability of consumption. Specifically, we estimate the relationship between consumption, 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. Consumption risk, finally, is defined as the variability in the distribution of predicted consumption. This concept of risk explores the fact that labor markets in rural India are largely dominated by agriculture at the time of the survey. 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 find that consumption risk substantially reduces the probability that girls attend school. This effect is robust to controlling for average risk in sample villages belonging to the same marriage market. Consistent with our theoretical model, we also find that risk increases the probability that girls perform household chores, and stronger effects for older girls. Our findings are also robust to controlling flexibly for household income and wealth, as well as for the share of area irrigated on a household's own farm for the sub-sample of farm-households. We also test if our results are driven by a savings motive at the household level, acknowledging that one alternative explanation could be that households in more risky environments may wish to save more. This would lead both men and women to work more, such that the observed effect on girls would really be driven by an increase in maternal labor supply, not dynamic complementarity in the education production function. We find no evidence that household savings is higher in risk-prone areas. Finally, we test if returns to education or the supply of education vary with risk, again finding no evidence that these factors may explain our results.

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 accumulation. We argue that the policy makes an *ex post* labor supply response to shocks more effective by mitigating the negative effect of shocks on wages. We estimate the extent to which NREGS reduces variability in consumption and use this reduction in consumption risk to simulate the effect of the program on girls' probability to attend school. We find that a similar program, that held the level of wages fixed, would increase school attendance of girls by 2.9 percentage points.

The wage effects if such programs should not be ignored, however. Shah and Steinberg (2015) and Li and Sekhri (2013) find negative effects of the NREGS on the time girls spend in school using the rolled phase-in of NREGS as a source of variation. It is argued that the NREGS increased the opportunity costs of time of adolescents and therefore led them to drop out of school at younger ages.[3] 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.

This paper links two significant strands in the development literature. The first, focuses on the role of risk as a constraint to productive investment in low-income countries (see e.g. Rosenzweig and Binswanger, 1993; Fafchamps and Pender, 1997; Dercon and Christiaensen, 2011; Bianchi and Bobba, 2013; Karlan et al., 2014). Consistent with much of this literature we find that risk reduces investment. The second strand considers the determinants of human capital investment in low-income countries (see e.g. Foster and Rosenzweig, 1996; Glewwe and Jacoby, 2004; Jensen, 2010, 2012; Oster and Steinberg, 2013). As noted previously, there is other work that focuses on how income realizations affect schooling investment in a setting where school investments at different points in time are complementary. However, we are not aware of other work that considers the question from an *ex ante* perspective.

The remainder of the paper proceeds as follows. Section 2 discusses the theoretical model. Section 3 presents evidence for dynamic complementarity in the production of human capital in India. The data are introduced in Section 4, and the estimation strategy in Section 5. Results are discussed in Section 6. Section 7 simulates the effect of the NREGS on the outcome of interest, and Section 8 concludes.

[3] This effect seems to work through intra-household substitution in household chores for girls, and through the labor market for boys.

2 A model of household time allocation, risk and human capital investment

Our model helps to illustrate the basic mechanisms of an *ex ante* schooling response to risk that is particularly relevant to the circumstances of girls in rural India. It captures, in particular, three important features that may lead to a relationship between consumption risk and child schooling. The first is dynamic complementarity in the education production function as in Jacoby and Skoufias (1997); Cunha and Heckman (2007); Cunha et al. (2010). As a consequence of this complementarity households may be reluctant to make initial investments in schooling if they are uncertain about the ability to capture the returns to this investment when schooling may be interrupted as a consequence of a negative shock. The second feature is an *ex post* labor supply response on the part of the mother. The idea is that a mother will enter the labor market in the face of an adverse shock (Kochar, 1999; Rose, 2001; Jayachandran, 2006). The third feature is a substitution between the mother's and her daughter's time in home production. The idea here is that the opportunity cost of time for a daughter must rise during bad periods even though she is unlikely to be active in the labor market (Skoufias, 1993; Ilahi, 2000).

We consider a two period model where utility u in each of the two periods is defined over market c^m and home c^h goods, with an additional second period term that is defined as a function v of the eventual stock of human capital H_2 at the end of the second period:[4]

$$L = u(c_1^m, c_1^h) + \delta\,\mathbb{E}[u(c_2^m, c_2^h) + v(H_2)]. \tag{1}$$

The human capital stock in period 2 increases in school time in periods 1 and 2 according to a constant elasticity of substitution (CES) production function:

$$H_2 = A[\nu s_1^\rho + (1-\nu)s_2^\rho]^{1/\rho}, \tag{2}$$

for $\nu \in [0,1]$ and $\rho \in [-\infty, 1]$. A is the total factor productivity, and ν the share parameter, also described as *skill multiplier* in Cunha and Heckman (2007). The degree of dynamic complementarity is entirely governed by ρ, where $\sigma = 1/(1-\rho)$ is the elasticity of substitution between school time in periods 1 and 2. The single-period utility function u is assumed to be log Cobb-Douglas with share parameter α:

$$u(c^m, c^h) = \alpha \ln c^m + (1-\alpha)\ln c^h. \tag{3}$$

Market goods are purchased using cash income that is earned by the mother 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 mother

[4] Because we focus on girls who are likely to leave the village for marriage, we assume that the returns to human capital are unaffected by village-level risk.

in home production,

$$c_t^m = I_t + \theta_t + w_t(T^a - h_t^a). \tag{4}$$

I_t is the non-stochastic part of income (the father's income, where the father is assumed to be full time in the market or any 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 mother h_t^a and the daughter h_t^c,

$$c_t^h = h_t^a + h_t^c, \tag{5}$$

and school time is the difference between the daughter's time endowment T^c and the amount of time spent in home production:

$$s_t = T^c - h_t^c. \tag{6}$$

Finally, the shock θ_t is zero in the first period, and takes on two values $[-r, r]$ with equal probability in the second period. Note that we do not allow the shock to affect the wage; adding such an effect substantially increases the analytic complexity as the price and income effects of wages move in opposite directions. Our empirical results suggest that the income effects dominate, i.e. mothers tend to work more when facing adverse shocks.

We first solve for second period maternal home time by equating the marginal return in the market to the marginal time at home:

$$h_2^a = (1-\alpha)T^a - \alpha h_2^c + (1-\alpha)(I_2 + \theta_2)/w_2. \tag{7}$$

The result is decreasing in the daughter's home time, increasing in the shock, and decreasing in the wage. With a more favorable shock or a lower wage, mothers spend less time in the labor market and more time at home. Similarly first period maternal time at home is:

$$h_1^a = (1-\alpha)T^a - \alpha h_1^c + (1-\alpha)I_1/w_1. \tag{8}$$

To develop an intuition for how risk affects first-period school time in the presence of dynamic complementarity, we first consider the extreme of a production function with perfect complements ($\rho \to -\infty$) and where $v(H_2) = H_2/\zeta$ is linear in the stock of human capital. We will subsequently numerically illustrate the effects for intermediate values of ρ and alternative specifications of $v()$.

When first and second period school times are perfect complements, second period home time of the daughter will depend on whether or not school investment in the second period is constrained by first period school time. By "constrained" we mean that, due to the complementarity in production, if first period schooling were exogenously higher, then *ceteris paribus* the household would also choose higher second period schooling. In this sense the household will be constrained if:

$$\theta_2 > w_2(\zeta - T^a - h_1^c) - I_2, \tag{9}$$

in which case

$$h_2^c = h_1^c. \tag{10}$$

Otherwise,

$$h_2^c = \zeta - T^a - (I_2 + \theta_2)/w_2. \tag{11}$$

Note that in the former, constrained case, child home time does not depend on the shock but in the second unconstrained case, child home time is decreasing in the shock. Under either condition maternal home time is decreasing in the shock.

By substituting h_2^a, h_1^a and h_1^c into L and taking expectations over the assumed distribution of shocks, we can then write the maximand as a function of child home time in the first period only. This expression is straightforward in the case that the variance of the shock is small enough that the household's second period school time is always constrained by the first period time – that is as long as things cannot be not so bad in the second period that the household chooses to provide less schooling than was chosen in the first period. In particular, assuming we have an interior solution we can differentiate with respect to h_1^c. While the resulting expression cannot be solved in closed form, we can write a series approximation in the size of the shock for small values of r:

$$h_1^c = \zeta(1 + \frac{1}{\delta}) - T^a - \frac{I}{w} + \frac{\delta^2 r^2}{w^2 \zeta (1+\delta)^2} + O(r^3). \tag{12}$$

where that we assume that $I_1 = I_2$ and $w_1 = w_2$ to keep the solution simple.

Taking values for the parameters: $T^a = T^c = 4$, $\alpha = \delta = 1/2$, $\zeta = 4$, $I = 6$, and $w = 1$, this expression becomes:

$$h_1^c = 2 + r^2/36 + O(r^3). \tag{13}$$

The condition for second period schooling to be constrained by the first period schooling is $r \leq 12$; however, this implies home-time of 6, which exceeds the time allocation for the child. Thus for an interior solution and these parameter values, we always have the constrained case and second period schooling is equal to first period schooling. The exact numerical value for h_1^c is 2 (corresponding to schooling of 2) when there is no variance, and 3 (schooling of 1) when $r = 6$.

This result shows clearly that – in the constrained case – first period home time is increasing in the variance of the shock and thus conversely first period schooling is decreasing in the variance of the shock. This result also implies that second period school time and thus the eventual stock of schooling is decreasing in the variance of the shock. The intuition is simple. If risk increases, then the marginal cost of second-period school time – in terms of foregone consumption of the home produced good – increases more in the event of a bad shock than it decreases in the event of a good shock. Because, in the constrained case, second-period schooling is set equal to first period schooling, an increase in risk lowers first period schooling.

Note that this is the opposite of 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. Here in effect first period investment is a complement with second period investment, so the opposite occurs when faced with greater risk.

If the variance is so high that condition (9) is not met in the bad state of the world, then variance can have the opposite effect on first period time allocation. In particular, in the good state of the world one obtains (10) and in the bad state of the world one obtains (11). Thus the second period downside risk is mitigated because it is no longer tied to the first-period level of schooling. On the other hand, the upside risk can fully be realized by setting first period equal to second period schooling. In particular, the series approximation to first-period home time in the unconstrained case is:

$$h_1^c = \zeta(1+\frac{2}{\delta}) - T^a - \frac{I}{w} - \frac{\delta\, r}{w\ (\delta+2)} + 2\,\frac{\delta^2 r^2}{w^2\zeta\ (\delta+2)^3} + O(r^3). \tag{14}$$

Note that first period home time falls in r and increases in the variance of r. The negative term may dominate if the ratio δ/w is sufficiently small. This result has an analogy with respect to borrowing in the case of investment risk: increasing risk 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 one simply defaults.

More generally, the effect of variance on first-period investment depends on whether human capital demand exhibits dynamic complementarity, that is whether increases in first period schooling lead to an increase in the marginal utility of second period schooling. This depends, in turn, both on the elasticity of substitution of the human capital production function and on the extent to which parents exhibit diminishing marginal utility with respect to human capital through the function $v(H_2)$. In particular, let:

$$v(H_2) = \frac{1}{1-\phi}(H_2)^{1-\phi}, \tag{15}$$

with $\phi >= 0$. By substituting (2) into (15) and differentiating with respect to s_1 and s_2 we find:

$$sign(\frac{d^2 v(H_2)}{ds_1 ds_2}) = sign(\phi+\rho-1). \tag{16}$$

Thus for $\phi = 0$, i.e. linear utility with respect to child human capital, there will always be dynamic complementarity in demand, except the extreme case where the human capital production function exhibits perfect substitutability ($\rho = 1$). On the other hand for $\phi = 1$, logarithmic utility, time in school at different points will be dynamic complements if and only if the production function exhibits complementarity ($\rho < 0$). Finally, if $\phi = 2$ then school investments at different points in time will be dynamic substitutes for $\rho > -1$.

Although we are unable to derive analytic solutions for the effects of risk on first-period schooling in the general CES case, numerical solutions suggest that the sign of the effect of variance on first-period human capital investment maps 1:1 with the presence of dynamic complementarity in demand for human capital as described above.

Figure 1 plots the derivative of first period schooling with respect to the magnitude of the shock against the log of the elasticity of substitution $\sigma = 1/(1-\rho)$.[5] It shows outcomes for three functional forms, linear utility in human capital, and two alternative functions that allow for risk-aversion in human capital: $v(H_2) = ln(H_2)$ and $v(H_2) = -1/H_2$. 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 predicted effect of risk is strictly negative as long as the human capital production function exhibits strong complementarity. For more moderate values of ρ, the picture is somewhat more complex. With linear v 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 v, this is no longer the case. Second period risk will then affect the expected marginal utility of the human capital stock, and households will rather invest more in schooling in order to avoid very low stocks of human capital for their children. If $v(H_2) = ln(H_2)$, for example, first period schooling will increase in risk for all $\rho > 0$. If $v(H_2) = -1/H_2$, the predicted effect of risk becomes positive already at negative values of ρ.

Whether variance increases or decreases first-period investment thus depends on the presence of dynamic complementarity in the demand for human capital, which depends in turn, at least in part, on parental preferences for child human capital. While we can (and do, below) directly estimate a production function, the shape of human capital preferences is unclear. If we think about equation (1) as overlapping-generations problem, we would expect parents to exhibit diminishing marginal utility with respect to the future consumption of their child, such as in Becker (1991). But we are not maximizing household utility over the daughter's future consumption, but her final stock of human capital. How the stock of human capital then translates into future earnings and consumption is essentially an empirical question. While returns to schooling seem to be rather linear in reality (Banerjee and Duflo, 2005), increasing evidence suggests that households have inconsistent expectations regarding the exact relationship between years of schooling and future income. Many households seem to expect that marginal returns to schooling increase with years of schooling (Banerjee and Duflo, 2011; Barrera-Osorio et al., 2011; Akresh et al., 2012). A linear utility in human capital might thus be a reasonable approximation of household decision making

[5] Similarly to above we choose the parameter values of $T^a = T^c = 4$, $\alpha = \delta = 1/2$, $I = 6$, $\zeta = 4$, and $w = 1$. Furthermore we set $r = 6$ and the skill multiplier $\nu = 1/2$.

if we combine the expectations of increasing returns to human capital with decreasing marginal utility of the child's future consumption.

3 Estimating the elasticity of substitution between school times

One important component of the predicted effect of risk is the degree of complementarity between each period's school times. An exact estimate of ρ would allow us to make more sensible predictions about the effect risk on time in school.

In order to estimate ρ, we need measures of skills and investment. Given that the data used in this paper does not contain information about skills, we have to draw on a different data set from India, the Young Lives Survey. The Young Lives project focuses on child development and follows children from two different cohorts in four different countries since 2002. For this study, we use the India data from rounds two and three.[6]

We follow Cunha and Heckman (2007), in specifying a constant elasticity of substitution production function, in which the stock of human capital, more precisely of cognitive skills, increases in the initial stock of human capital H_1, current investments s_2, and the parental stock of human capital H_p:

$$H_2 = A[\nu_1 H_1^{\rho} + \nu_2 s_2^{\rho} + \nu_3 H_p^{\rho}]^{1/\rho}. \tag{17}$$

A is the total factor productivity. $\nu_i \in [0, 1]$ and $\sum_1 \nu_i = 1$ for $i \in \{1, 2, 3\}$. $\sigma = 1/(1 - \rho)$ is the elasticity of substitution between inputs and skills in the production of outcomes, and measures how easily previously accumulated stocks of skills can be compensated by current investments.

Skills are likely measured with error. Therefore, we follow Cunha et al. (2010) in their approach of estimating a skills production function by first estimating a set of latent factors (H_1, H_2 and H_p) from multiple measurements of skills (Z_j). This allows to identify the structural parameters of eq. (17) even if skills are measured with error. To keep the problem simple, we specify linear factor models, and as in Cunha et al. (2010) we assume that the measurements Z_j proxy the natural logarithms of the factors, such that for H_1:

$$Z_{1j} = \mu_{1j} + \alpha_{1j} \ln H_1 + \epsilon_{1j}. \tag{18}$$

We also allow for endogeneity of investment and skills in (17) by instrumenting both variables with rainfall, land ownership and the interaction of the two. We find that rainfall raises wages, which draws children out of school, but this pressure falls with wealth, here captured by land ownership. The identifying assumption is that conditional on parental skills, neither land ownership nor its interaction with rainfall has a direct effect on skills or investment. Assuming that the identifying

[6] More details on the data and estimation strategy are available in Appendix A.

assumption is valid, eq. (17) can be identified in a control function approach,

$$\ln(H_2) = A + 1/\rho \ln(\nu_1 H_1^\rho + \nu_2 s_2^\rho + \nu_3 H_p^\rho) + \kappa_1 \gamma_h + \kappa_2 \gamma_s, \tag{19}$$

where γ_h is the residual of the investment equation and γ_s is the residual of the skills equation. Results can be found in Table 1. Columns (1) - (3) report results for the younger cohort, aged 5 in round 2 (2007) and aged 8 in round 3 (2009-10). Columns (4)-(6) report results for the older cohort, aged 12 in 2007 and 15 in 2009-10.[7] As can be seen the results are fairly similar for both cohorts and irrespectively of whether we allow investments, or investments and initial skills to be endogenous. The self-productivity parameter ν_1 is 0.26 in the younger cohort and 0.64 is the older cohort. The elasticity of substitution σ is close to one for both cohorts: it varies between 0.91 and 1.06 for the younger cohort and between 0.92 and 1.16 for the older cohort. The results thus suggest that the human capital production function can best be approximated by a Cobb-Douglas production function, while we can strongly reject perfect substitutability in all cases.

These results are well in line with Cunha et al. (2010) who use data from the US, as well as with Attanasio et al. (2015), who also use the Young Lives Survey data from India. While we focus on children's study time per day as key investment variable, Attanasio et al. (2015) use spending on school as well as the number of meals per day to proxy investment. Not surprisingly, the results closely overlap.[8]

4 Context and data

We test if risk affects school time with the 2006 round of the Rural Economic and Demographic Survey data. The REDS is the follow-up survey of the Additional Rural Incomes Survey (ARIS), which was first collected in 1971. The sample was designed to represent the rural population of India across 17 major states. The ARIS covers 4,527 households in 259 villages. Three follow-up rounds were collected in 1982, 1999, and 2006 to re-visit these households. The sample was increased over time by randomly sampling additional households from the same villages. The sample in 2006 consists of roughly 9,500 households in 242 villages.[9] We can only use the 2006 round because of changes in the questionnaire over time.

There are three reasons for using the REDS data to study the questions outlined above. First, the geographic coverage of almost the entire country allows a comparison of households in very different agro-climatic regions and economic conditions. Second, the REDS 2006 survey collects detailed information about time allocated to different activities for all household members. This information is collected for three seasons of the year (each of which lasts for four months), which are

[7] The estimation is restricted to children who are still in school in the follow-up round. Summary statistics are reported in Table C.1 in the Appendix.

[8] Attanasio et al. (2015) find a elasticity of substitution of 0.871 for the younger cohort and of 0.952 for the older cohort, and self-productivity parameters around 0.283 and 0.605 for the younger and the older cohorts, respectively. In Cunha et al. (2010), the estimated elasticity of substitution varies between 0.421 and 0.547 in the age group 5-14.

[9] Due to armed conflict no data were collected in Jammu & Kashmir and in Assam in the 2006 round of interviews.

also marked by very different levels of agricultural activity. Seasons 1 and 2 are the agricultural peak seasons in which most of the agricultural production takes place. Season 3 is the dry season, during which only very few crops are cultivated and agricultural activity is considerably lower. We restrict the sample to households and individuals with complete information on time allocation, income and consumption, and who live in rural areas.[10] In order to use the available information efficiently, we first generate a sample of mothers with children aged 6 to 16. This sample consists of 4,356 households and 4,651 mothers.[11] Table 2 reports some general household summary statistics for this sample. We then restrict the sample to matched daughter-mother information, for all girls aged 6 to 16, who still live with their mothers.[12] This final sample consists of 2,934 households, 3,040 mothers and 4,306 daughters.[13] The unit of observation in the main estimation is the daughter-mother match across three agricultural seasons, resulting in a final sample size of 12,918 observations. Detailed summary statistics for mothers and daughters are reported in the Appendix, Table C.2.

4.1 Risk in rural India

In rural India, agriculture is the dominant source of income for the majority of households, either through own agricultural production or through casual agricultural employment. Agriculture employs 67% of all male workers and 83% of all female workers in 2004/05 (National Sample Survey Office, 2006). In the REDS sample, 63% of households cultivate their own land, and labor income consists to 59% of income from casual employment in agriculture or own agricultural production (c.f. Table 2). In addition, employment in agriculture (own agricultural production and agricultural casual labor) makes up for 79% of the labor market time of mothers.[14]

Agricultural income is inherently risky. As previous literature has shown, wages and employment levels in the agricultural sector are strongly influenced by rainfall conditions (see e.g. Jayachandran, 2006; Kaur, 2014). 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. The variability of rainfall combined with the village-level availability of irrigation systems should therefore be good proxies for consumption risk in this context.

In order to predict consumption risk, we merge the REDS data with historical rainfall data.[15] We then compute consumption risk as follows. First, we estimate in the cross-section how strongly current rainfall levels determine a household's consumption per capita (in logs) given a village's

[10] We drop all households living in cities with more than 10,000 inhabitants, based on the Census in 2001.

[11] We use this sample to calculate our risk variable and to predict wages.

[12] We focus on the age group 6 to 16 because most girls still live at home at this age.

[13] The final sample is considerably smaller than the original sample. By far most households had to be dropped because they have no children in the relevant age group ($\approx 2,500$), or only boys ($\approx 1,500$). Another 1,000 households had to be dropped because they live in cities with more than 10,000 inhabitants, and for roughly 1,000 households it was not possible to match mothers and daughters by their identifiers.

[14] As reported in Table C.2 in the Appendix.

[15] We use precipitation data compiled by the University of Delaware for the period 1960 to 2010. Data are available for 1900 onwards, but the data quality improved a lot over the time period, which is why we prefer working only with more recent data. We merge the data with the geocode of the village center. Since the data are available for grids of 0.5 degrees in latitude and longitude (approximately 50 km), some of the villages fall in the same cell.

share of area that is irrigated.[16] Results are reported in Table 3. Rainfall is interacted with the share of agricultural land under irrigation in the village, to capture differential risk exposure. The assumption here is that rainfall shocks will translate less strongly into consumption outcomes the higher the share of irrigated land in a village is. This can be because households are more likely to have irrigation on their own land in villages with a high share of area under irrigation or because casual agricultural employment will be less affected by current rainfall levels (since most farmers in the village do not depend on current rainfall levels for their agricultural output). We also control for irrigation levels, and average annual rainfall separately. State fixed effects and a linear time trend are also included. To gauge the robustness of these estimates, we add a number of controls that could proxy for permanent income such as education, caste, religion and land ownership in column (2). As can be seen in Table 3, our results are robust to the inclusion of these controls.[17]

Given these estimates, we then simulate the amount of risk faced by every household in a village using the historical rainfall distribution and the current share of area that is irrigated in each village. We use historical rainfall data to calculate the probability of each rainfall outcome. We then predict a household's log consumption per capita for each rainfall outcome given the current availability of irrigation in a village. By multiplying the probability of each rainfall outcome with the predicted log consumption at that rainfall level, we obtain a probability distribution of consumption outcomes for each household in a particular village.[18] Finally, we calculate the interquartile range as well as the standard deviation of a household's predicted log consumption per capita as measures of consumption risk.

4.2 Human capital accumulation in rural India

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. By 2010/11 the gross enrollment ratio reached 114.9% for boys and 116.3% for girls in the classes 1 to 5. The gender gap has also been closing in middle school enrollment: by 2010/11, 87.5% of the boys were enrolled in the classes 6 to 10, as compared to 82.9% of the girls (also in gross figures).[19]

Still, literacy rates remain low, particularly for women and in rural areas. According to the Census in 2011, only 50.6% of the women aged 15 and above living in rural India are literate. This is not only an issue among the adult population: according to the Annual Status of Education

[16] This approach is inspired by Dercon and Christiaensen (2011).

[17] Still we have to worry that these results are driven by differences in average consumption across villages that correlate with average rainfall levels rather than current shocks. We therefore employ two additional robustness checks. First, we test if using the deviation in rainfall from the long-term average instead of current precipitation changes the results. Second, we pool the 2006 data with the 2014 follow-up round and test if our results are sensitive to the inclusion of village fixed effects. In both approaches, the results are qualitatively similar to those reported here, and are available from the authors upon request.

[18] The probability weights 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.

[19] Ministry of Human Resource Development, Government of India. Data retrieved from Indiastat.

Report (ASER) for Rural India, only 48.2% of the children in grade 5 could read a grade 2 level text in 2011 (Pratham, 2012).[20]

While the quality of education is an often-cited reason for low learning outcomes in India (Kremer et al., 2005; Banerjee et al., 2010), demand-driven factors also play a role, particularly in the inequality of learning outcomes between boys and girls. According to the Status of Education and Vocational Training in India survey conducted in 2011/12, the ratio of children not attending school in the age group 10 to 14 is highest for girls in rural areas (10.1%), as compared to 6.4% for girls in urban areas and 6.7% for boys in rural areas. This inequality increases further if we consider the age group 15 to 19: in this group, 44.8% of the rural girls are out of school as compared to 29.7% of the urban girls and to 34.3% of the rural boys. In the same survey, more than half of the girls aged 5 to 29 years who were currently not in school but ever attended an educational institution stated that attending domestic chores was the single reason for not being enrolled in any educational institution (National Sample Survey Office, 2015).

In the REDS sample, 79% of the girls aged 6 to 16 are currently enrolled in school, as compared to 86% of the boys in the same age group.[21]

4.3 Time allocation of mothers and daughters in rural India

The REDS collects two different types of time allocation data. The demographic section records the number of hours per day allocated to different activities on a typical day for all household members in each of three agricultural seasons. This section explicitly differentiates between time in school, time allocated to household chores, market work and leisure. Consistent with the ILO definition, we define labor supply as the total number of hours per season worked in paid employment or in self-employment. However, we exclude hours worked in own-agricultural production from this variable. The reason for this is that labor supply to own-agricultural production will be affected by rainfall risk in different ways and we want to avoid mixing up different causal mechanisms.[22] Household chores, then, include all activities related to the household: construction and maintenance of the house, farm and other assets, household work and other household-related activities (collecting fuel, herding cattle, fishing, cutting grass). We also count the time allocated to livestock production as home production, arguing that livestock production is mostly a household duty, even though it could also be undertaken for profit.

As reported in Table 4, girls in our sample spend on average 6.2 hours per day in school or studying in season 1 (one of the two agricultural peak seasons). Girls also spend 1.6 hours per day on household chores, but only 0.1 hours per day working on the labor market. This exemplifies why

[20] ASER is based on an annual survey that assesses children's schooling status and basic learning levels throughout all rural districts of India. It is facilitated by the Indian NGO Pratham and interviews are conducted by volunteers, which has raised a number of doubts regarding data quality. Still, it is the only India-wide assessment of learning levels currently available.

[21] Summary statistics are reported in the Appendix: Table C.2 for girls, and Table C.3 for boys.

[22] Labor supply to own-agricultural production should fall with increasing rainfall risk, as this income source becomes increasingly risky. Labor has to be allocated to agricultural production partly before the rainfall realizes, hence before the household can assess how the harvest, and therefore the returns to that labor, will be.

we need to take into account intra-household substitution effects when thinking about the effect of shocks on the time allocation of girls.

For all household members aged 10 to 65, the questionnaire also collects information about the actual number of days and hours per day allocated to a number of different activities per season.[23] We use this section to compute time allocation variables for mothers. The labor force participation in our sample of mothers is 21% in the agricultural peak seasons and 15% in the lean season (c.f. Table 4).[24] In all seasons, labor supply is on average slightly higher than time allocated to own-agricultural production. By far the majority of maternal time is spent on household chores: 701 hours in season 1, which corresponds to roughly 43.8 hours per week. Hours allocated to labor supply and own-agricultural production together make up for about 9.4 hours per week on average in season 1. Total time spent on activities other than leisure is thus 53.2 hours per week in season 1, which is very close to the time spent on those activities reported in the Time Use Survey of 1998/99 (53.4).[25]

Because this section does not collect time in school or studying, we cannot use this information to test for the effects of risk on time in school. However having information on time allocated to household chores from two different sections, allows to test if our results are robust to using alternative time measures. With respect to time spent on household chores by girls, we find correlation coefficients of 0.83, 0.81 and 0.81 between the two time measures for seasons 1, 2 and 3, respectively.[26]

5 Estimation strategy

The structural hours function to be estimated is:

$$h_{ijt} = \beta_0 + \beta_1 R_{jt} + \beta_2 \ln w_{ijt} + \beta_3 N_{ijt} + \beta_4 X_{ijt} + \epsilon_{ijt}, \tag{20}$$

where the dependent variable is the amount of time child i, living in village j, allocates to schooling (or to home production) at time t. We are particularly interested in estimating β_1, i.e. the effect of consumption risk R_{jt} on hours allocated to different activities. As discussed previously, this variable is predicted at the village level. The identifying assumption is that, conditional on the control variables, no other (unobserved) characteristics that determine hours in school are correlated with our regressor of interest, i.e. consumption risk. This assumption is quite strong and we will

[23] The full list of activities are salaried work, agricultural casual labor, own-crop production, own-livestock production, work for public works programs, non-agricultural casual labor, migration, self-employment in non-farming, construction and maintenance of house, farm and other assets, household work, other household-related activities (collecting fuel, herding cattle, fishing, cutting grass) and other activities (schooling, unemployment, leisure).

[24] This is slightly lower than the 24.5% reported for rural India by the NSSO in 2009/10. But the NSSO classifies own-agricultural production as labor supply, while we do not.

[25] There is a substantial time lag between the two surveys, however. Also our sample consists only of mothers, while the TUS surveys women aged 15 to 65.

[26] We also find qualitatively similar results when using actual hours worked in the household rather than usual hours per day as dependent variable when estimating the effect of risk. Results are reported in the Appendix, Table C.4.

discuss how we address some possible threats to identification at the end of this section.

In the absence of a labor market for girls and given intra-household substitution effects, maternal wage w_{ijt} should be a good proxy for the opportunity cost of time of girls and needs to be included in the structural hours equation. Furthermore, we expect time allocation to depend on non-labor income and all other household members' labor income N_{ijt} (which excludes maternal labor income). These two variables are expected to capture income effects on time in school. We assume that preferences for school can be captured by individual and household characteristics X_{ijt}, i.e. age, caste, religion, the number of children (in the household), household size, and land ownership, as measure of general household wealth. Finally, the error term can be decomposed into a permanent and two transitory components such that $\epsilon_{ijt} = \mu_j + \sigma_{jt} + \eta_{it}$, where μ_j is a state-level fixed effect, σ_{jt} are village-level shocks and η_{it} is a mean zero, strictly exogenous, idiosyncratic shock.

As discussed above, the dependent variable will depend on maternal wages, if there is substitution between mother's time and daughter's time at home. But adequately controlling for maternal wages is challenging for a number of reasons. First, observed wages are potentially endogenous to labor supply if wages reflect work effort as well as skills or ability, such that individuals who work longer hours are likely to get higher wages. This is of concern particularly in the labor markets of developed economies (see e.g. Keane, 2011). In the context of low-skill agricultural wage work, however, we feel that this is less of an issue, because wage rates are determined by local conditions rather than individual abilities (Rosenzweig, 1978; Rose, 2001; Jayachandran, 2006). Second, both wages and other household members' labor income are endogenous to shocks. There might be general equilibrium effects of shocks on wages (see e.g. Jayachandran, 2006): if all households in the village supply more labor due to a shock, equilibrium wages should be lower than in the absence of a shock. We need to be aware of this when testing for effects of rainfall shocks on labor supply: our predicted wages already account for general equilibrium effects of shocks on wages in the village economy, and this might mask some of the effect of shocks on time allocation decisions. Third, wages are only observed for women who are currently in the labor market, such that we have to deal with missing wages for all women who are not participating in the labor market. Potentially, the sample of workers is not a random sub-sample of all women, such that we have to account for selection bias when imputing wages. Fourth, most women in our sample reported working in different activities, at different wages. Fifth, we have to deal with classical measurement error since wages are mostly measured with considerable error in micro data (Keane, 2011).

We address the last three issues by predicting wages for all women in our sample. Following Blundell et al. (2007), we predict wages using a Heckman selection correction. Since women report working in a number of different activities at different wage levels, we calculate weighted average wages, i.e. calculate total income from all activities and divide it by the total number of hours worked.[27] Assume wages are determined as follows:

$$\ln w_{ijt} = \delta_0 + \delta_1 X_{ijt} + \delta_2 R_{jt} + \mu_j + \sigma_{jt} + \upsilon_{it}. \tag{21}$$

In eq. (21), X_{ijt} captures age, age squared, caste, number of children in the household, religion, education, land ownership, and non-labor income. R_{jt}, μ_j and σ_{jt} again capture consumption risk, state fixed effects, and village-level rainfall shocks respectively. The structural participation equation is similar to the wage equation, except that we add wages $\ln w_{ijt}$, other household members' labor income N_{ijt}, and augment X_{ijt} by marital status:

$$p_{ijt} = \alpha_0 + \alpha_1 \ln w_{ijt} + \alpha_2 X_{ijt} + \alpha_3 R_{jt} + \alpha_4 N_{ijt} + \mu_j + \sigma_{jt} + \upsilon_{it}. \tag{22}$$

Substitution eq. (21) into the structural participation equation, gives the reduced-form participation equation, which is the selection equation for the Heckman selection model,

$$p_{ijt} = a_0 + a_1 X_{ijt} + a_2 R_{jt} + a_3 N_{ijt} + \mu_j + \sigma_{jt} + \upsilon_{it}. \tag{23}$$

We estimate this reduced-form participation equation first and then estimate the log hourly wage equation (21) including the inverse Mills ratio obtained from the estimated participation equation (23). In this approach, the wage equation is identified from the exclusion of other household members' labor income, and marital status from the wage equation, as well as from the normality assumption.[28] The structural hours equation (20) can then be estimated with imputed mothers' wages.

Since our measure of consumption risk is not directly observed but predicted for each village, robust standard errors clustered at the village level probably overstate the estimator's precision. We therefore additionally conduct a percentile t-bootstrap method with cluster re-sampling at the village level, and 4,999 replications. Predicted consumption risk is re-calculated for every bootstrap sample, and the distribution of t-statistics is obtained from regressing the outcome of interest in every bootstrap sample on a newly predicted risk variable. We report bootstrap p-values in addition to cluster-robust standard errors where appropriate.[29]

As mentioned previously, the identifying assumption warrants some discussion. Obviously, irrigation systems are not placed randomly and one would expect wealthier villages to have more irrigation and lower consumption variability. If wealthier villages have also higher school invest-

[27] Alternatively one could use marginal wages, hence the lowest wages observed at which women supply a positive number of hours. But we are worried that this measure cannot account for necessity-driven labor supply, i.e. cases in which labor supply is high because wages are low in all activities that are performed.

[28] Results are reported in the Appendix, Tables C.5 and C.6. In our sample, 49 women report having a permanent job. We assume that the wages of these women are determined by a different process than the casual wages observed for the rest of the sample. Therefore, we predict wages in permanent employment separately for this group. As the participation in permanent jobs cannot be predicted by other household members' income, marital status or the number of children in the household, we predict these wages without adjusting for self-selection as reported in Table C.6.

[29] The percentile-t cluster bootstrap has a better asymptotic refinement than the simple pairs cluster bootstrap, and seems more appropriate in a sample with 221 clusters. For more details see Cameron and Miller (2015).

ments, a regression of school time on predicted consumption risk would yield biased estimates. However, we control for different sources of household income and wealth in all our specifications and allowing for more flexibility in the relationship between income and hours in school does not affect our results. Additionally, for the subset of farm-households, we control for the share of area that is irrigated on the farm, to avoid attributing the effect of reduced female irrigation work on the farm to a reduction in consumption risk. Again, our results do not change. The full set of robustness checks is discussed in Section 6.3.

6 Results

6.1 *Ex post* responses to shocks

In the following we want to test explicitly for the validity of two crucial assumptions made in the theoretical model. First, an *ex post* labor supply response by mothers to economic shocks. And second, substitutability in maternal and daughter time in the performance of household chores.

We start by estimating the effect of rainfall shocks on maternal labor supply and time allocated to households chores (Table 5). As before, our sample consists of all mothers in our data to use the available information most efficiently. We control for age, caste, religion, education, household size, number of children in the household, land ownership, non-labor income, other member's labor income, and state fixed effects. We do not control for risk, nor for wages, because we do not want to mask the effect of shocks by conditioning on these variables.[30] We interact rainfall with a lean season dummy, to test explicitly for the timing of the labor supply response.[31]

Column (1) presents the effect of shocks on extensive margin labor supply. We do not find that rainfall affects extensive margin labor supply in either season. The intensive margin labor supply response to shocks is shown in column (2). While rainfall seems to have a positive (albeit not statistically significant) effect on labor supply in the agricultural peak seasons, we find that negative rainfall shocks increase labor supply in the agricultural lean season, thus *ex post* to the full realization of rainfall outcomes. As can be seen in column (3), the time allocation to household chores seems to mirror this, although the coefficient on the interaction term in not statistically significant.

If maternal and daughter time are substitutes in the production of the home produced good, then any labor supply response of mothers to shocks should translate into the time their daughters allocate to household chores and spend in school. We therefore test if shocks affect the time girls spend on those activities in different seasons in Table 6. Again, we also control for age, caste, religion, household size, number of children in the household, land ownership, non-labor income, other member's labor income, and state fixed effects.

As shown in columns (1) and (2), rainfall shocks, seem to affect time in school only at the intensive margin – not at the extensive margin–, and only in the lean season, but not in the peak

[30] Remember that we allow predicted wages to vary with rainfall levels.

[31] For a more direct interpretation of the interaction term, log rainfall is normalized to a mean of zero.

seasons. The seasonality of the effect is consistent with the theoretical model: *ex post* to shocks, mothers have to work in order to generate additional income, and this puts greater pressure on the daughter to reduce time in school and to perform household duties. The fact that an *ex post* response to shocks only materializes at the intensive margin is consistent with the notion of fixed costs of school attendance: since those costs had to be incurred already at the time of shock realization, an intensive margin response is "less costly" than an extensive margin response. Again, the time allocation to household chores mirrors the effect on school time, as reported in column (3). On average, girls spend more time on household chores in the lean season than in the peak season, and only at very high rainfall outcomes that difference approaches zero and is no longer statistically significant.[32]

6.2 The effect of consumption risk on girls' time in school

We continue by examining how risk affects the time girls spend in school. We have seasonal data for 4,306 girls in the age group 6 to 16. Our model predicts a negative effect of consumption risk on time in school if there is a high degree of dynamic complementarity in the education production function. At substitution elasticities of one, as observed in the YLS data from India, our model predicts a negative effect of risk only if household utility is increasing linearly in the stock of human capital.[33]

Table 7 reports OLS estimates of the effect of risk on time in school at the extensive margin in columns (1) and (2), and at the intensive margin in columns (3) and (4). We use two variables to proxy consumption risk, i.e. the interquartile range as well as the standard deviation of predicted log consumption. As discussed above, we control for predicted wages of mothers and for individual and household-level socio-economic characteristics in all our specifications. In order to account for the effect of non-labor income on time allocation, we include a measure of all other household members' labor income as well as total household non-labor income per capita.[34] We also control for rainfall shocks, state fixed effects, and a lean-season fixed effect. Standard errors are clustered at the village level; i.e. the level of variation of the main explanatory variable. Bootstrap p-values are reported at the bottom of the table.

We find that risk considerably reduces time in school at the extensive margin, but not at the intensive margin. This suggests that girls drop out of school temporarily or permanently in the presence of risk, while there is little scope for adjusting the time in school at the intensive margin. This result suggests the presence of an (unmodeled) fixed cost of school attendance, that causes households to withdraw a child from school entirely when faced with higher risk. According to our estimates, a one standard deviation increase in the interquartile range of predicted log consumption

[32] Note that the coefficient of the interaction term is not statistically significant, still marginal effects of the lean season are negative and significant at low rainfall levels, and close to zero (and not significant) and high rainfall levels.

[33] Remember, a positive effect of risk on time in school could either stem from high risk aversion with respect to the final stock of human capital or from perfect substitutability of school times in the production of human capital at moderate levels of risk aversion with respect to H_2.

[34] Other household members' labor income excludes the labor income of the mother.

(0.033) reduces the probability that girls attend school by 5.8 percentage points in every season.[35] This effect is statistically significant at the 1% level for both risk variables (c.f. bootstrap p-values at the bottom of Table 7).

Mothers' wages seem to be positively associated with time in school at the extensive margin, and negatively at the intensive margin, although the coefficient is statistically significant only at the extensive margin. This suggests that the income effect, rather than the intra-household substitution effect dominates at the extensive margin. However, due to the fact that we have to predict wages, we caution against a causal interpretation of the coefficient on mothers' wages. The remaining coefficients have the expected sign: as girls get older, they are less likely to attend school. Girls belonging to scheduled castes and tribes are less likely to attend school, while Hindu girls are more likely to attend school than girls from other religions. As non-labor income increases, the probability of attending school, as well as the time spent in school increases, again pointing to income effects.

A potential concern about these measured effects of risk is that risk affects the future returns to human capital as well as the opportunity cost of time, which is the focus of our model. As noted above, this concern is one of the reasons why we focus on daughters, who are likely to leave the village following marriage. Nonetheless, because risk is likely correlated across space, marriage market risk may well be correlated with consumption risk in the village of origin. We address this problem by adapting a strategy from Foster and Rosenzweig (1999). 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 case, 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. Table 8 reports the resulting estimates. The coefficient on marriage-market risk is close to zero and its inclusion does not substantially affect the point estimate on the coefficient of interest. Thus, as posited in our model the risk estimates appear to reflect the consequences of risk for the opportunity cost of time of daughters while they are children.

6.3 Robustness checks

As girls spend less time in school in the presence of risk, overall household utility can be increased if the daughter spends her available time on household activities. We test this prediction in Table 9. As before columns (1) and (2) report the extensive margin effects, and columns (3) and (4) the intensive margin effects. Consistent with the theoretical model, we find evidence that risk increases the time girls spend on household activities. Again the effect is only statistically significant at the extensive margin. The magnitude of the coefficient suggests that a one standard deviation increase in the interquartile range of predicted log consumption (0.033) increases the probability to perform

[35] Likewise, an increase in the standard deviation of predicted log consumption by one standard deviation (0.025) reduces the probability that girls are in school by 5.2 percentage points.

household chores by 2.9 percentage points. A one standard deviation increase in our alternative risk variable, the standard deviation of predicted log consumption, increases the probability to perform household chores by 3.4 percentage points. The effects are statistically significant at the 10% and at the 5% level, respectively, according to the bootstrap p-values.

The hypothesis that risk reduces time in school rests on the assumption that girls have to fill in for their mothers in household duties when faced with shocks. As the probability to perform household work strongly increases with age, this also suggests that the effect of risk should be more pronounced for older girls.[36] And this is precisely what we find, when we split the sample by age group. The coefficient on both risk variables almost doubles in magnitude when moving from the group of girls below the age of 11 to girls aged 11 and above.[37]

One concern could be that the observed effect of risk is not really a forward looking risk effect. Suppose past school time increases the marginal return to current school time, and past school time is lower in riskier areas due to more regularly occurring shocks. Then current school time should also be lower given that there was more variation in school time previously. In order to address this concern, we condition current school time on the stock of human capital – measured as the highest level of schooling achieved so far. Given concerns about endogeneity, we instrument the stock of human capital by rainfall shocks occurring in the school history of a child. Our results are robust to the inclusion of these controls.[38]

We also address the concern that our risk variables might capture wealth rather than variability in consumption. First, we test if dropping wealth and income controls, controlling more flexibly for these variables, or allowing for the effect of risk to vary with land ownership (as proxy for wealth) affects our results.[39] None of these specifications alters our main results. Second, we test if there is a direct relationship between area under irrigation and the probability of attending school for the subset of farm-households, and if this is driving our results. Again, controlling for the share of area that is irrigated on the farm, its square or the interaction with consumption risk does not affect our results.[40]

Finally, if boys also have to fill in for mothers in household chores, and preferences for boys' human capital are similar to preferences for girls' human capital, then we would expect to see a similar effect of risk on boys. On the other hand, as noted above, boys are more likely than girls to remain in the village as adults and risk may affect the returns to human capital in the future. Such effects might be either positive or negative, depending, among other things, on whether education is seen as way to diversify income risk. We therefore test if shocks affect the time boys spend in school, on household chores and on the labor market and whether risk also affects the schooling of boys. We find similar effects of rainfall shocks on boys to those on girls: conditional on being

[36]Figure B.1 in the Appendix plots the probability of performing household chores and of being in school by age. The probability that girls perform household duties increases continuously with age, while the probability of being in school starts to decline at the age of 11.

[37]Results are reported in the Appendix, Table C.7.

[38]Results are reported in the Appendix, Table C.8.

[39]Results are reported in the Appendix, Table C.9.

[40]Results are reported in the Appendix, Table C.10.

in school, a negative rainfall shock reduces the time boys spend in school in the agricultural lean season, and increases the time they spend on household activities. Interestingly, boys do not seem to spend more time, but rather less time on the labor market in the lean season when the household is faced with a negative rainfall shock, perhaps reflecting a wage effect. However, when we test for the effect of risk on boys, we fail to find a statistically significant effect on school attendance.[41] This result could reflect differences in the concavity of household preferences in the stock of human capital of their sons versus their daughters or, as posited above, differences in the return to human capital by risk.

To summarize, our results suggest a strong and robust relationship between risk at the household level and girls' time allocation, both to school and to household chores, which seem to be driven by *ex post* labor supply responses of mothers to income shocks and the intra-household substitution in household work. These effects are very harmful to the accumulation of human capital, and potentially for future earnings of these girls. We will discuss the potential role for public policy later and now look at some alternative explanations.

6.4 Alternative explanations

An alternative explanation to the effect of consumption risk on daughters' time in school might be that her mother is simply trying to work more in order to increase savings, and therefore spends less time on household chores. That households tend to save more in the presence of uninsured risk is by now well established. However, it is 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 (Zeldes, 1989; Deaton, 1991). On the other side, one might think that women would tend to work harder in order to accumulate savings more easily. If this is the case, then we must worry that the observed effect on child outcomes is only an artifact of maternal labor supply decisions.

Table 10 shows the effect of risk on maternal time allocation. In columns (1) and (2), we test if risk affects maternal labor supply at the extensive margin, and do not find any statistically significant effect. In columns (3) and (4), and (5) and (6) we test if risk affects hours worked on the labor market and on household chores, conditional on being in the labor force. We do find some – albeit weak – evidence that mothers, conditional on being in the labor force, spend more time on the labor market and less time on household chores in the presence of risk. However, only the effect of risk on time allocated to household chores is statistically significant at the 10% level, when inference is based on the bootstrap p-values. The bootstrap p-values on intensive margin labor supply are 15% and 22%, respectively.

An alternative explanation for an effect of risk on maternal time allocation is that risk reduces the time girls spend in school and increases their time in household activities, combined with the intra-household substitution in home production. As a household can increase utility by increasing

[41] Because the main substitution seems to work through household chores, we again control for maternal wages in the school time regression. Results are reported in the Appendix, Tables C.11 and C.12.

consumption of the home produced as well as the market produced good, mothers might decrease the time they spend on household activities and work more on the labor market, when their daughter spend more time on household chores. We can thus not conclusively identify the underlying causal mechanisms by looking at the time allocation of mothers alone.

Another, admittedly imperfect, way to approach this issue is to test if households tend to save more in the presence of uninsured risk. Since households in this context typically have no access to good liquid savings, our very crude measure of savings is the difference between household income and household consumption, which is probably measured with considerable error and also becomes negative as households dissave when faced with a shock. We therefore restrict the sample to households with positive savings. As reported in Table 11, however, we do not find any evidence that households tend to save more as risk increases. If anything savings seem to decrease with risk.

More importantly, in a multi-period model with known risk distribution, the effect of risk on incremental savings should be small in any given period because savings accumulation is substitutable across time. In contrast, any child human capital response is complementary over time and tied to child age. We thus have both theoretical and empirical arguments against the explanation that observed effects on girls' school time are driven by precautionary savings motives.

Above we used marriage market data to explore the possibility that the returns to education vary with risk. An alternative approach is to test directly whether the wage return to education varies with risk. We do not find any evidence that risk in fact reduces the returns to education either for men or for women. Of 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.[42]

Finally, one might worry that any schooling response of girls is driven by supply side constraints rather than demand side factors. For example, it could be that our risk measure is picking up general differences in village wealth which correlates with irrigation as well as with the supply of education. We therefore test if our results are affected by controlling for the number of primary or secondary schools, by the degree of teacher absenteeism or the availability of the mid-day meal scheme in schools. None of these variables affect our results.[43]

7 Simulating the effect of the NREGS on consumption risk, labor supply and human capital accumulation

Given the magnitude of the effects of risk observed above, it seems worthwhile to explore potential policy tools to mediate these. Obviously, any policy tool 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 (consumption) risk in village economies as farmers become more risk taking in their production decisions. This would then be

[42] C.f. with Tables C.13 and C.14 in the Appendix.

[43] As reported in Table C.15 in the Appendix.

particularly harmful for the poorest households with no own land and no access to agricultural insurance.

An alternative could be an income-smoothing policy, such as the Indian National Rural Employment Guarantee Scheme, that was introduced in rural India in 2006. The 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.

This scheme can affect schooling decisions through two effects: first, the NREGS provides employment at higher wages than casual agricultural wages, which affected both total labor supply and the amount of labor supplied to the private sector, and therewith equilibrium wages in the private sector (Berg et al., 2012; Imbert and Papp, 2015). As documented by Shah and Steinberg (2015) the resulting increase in wages did pull more women into the labor force and reduced the time women spent on household chores. Due to intra-household substitution in home production, this also raised the opportunity costs of time for girls and reduced the time they spent in school.[44] Second, the NREGS reduces consumption risk in rural areas, because it provides a minimum amount of employment at a fixed wage level independently of rainfall shocks. Assuming income effects dominate maternal wage allocation, the fact that the wage does not fall during periods of need is likely to reduce the magnitude of the labor supply response of women relative to case in which there are general equilibrium wage effects. This effect should in turn reduce the extent of the response among girls *ex post*. This lower response, in turn, increases the return to first-period school time.

The following simulation of the effects of the NREGS seeks to shed light on the importance of this second mechanism and is intended to complement existing empirical evidence on the short-term impact of the NREGS on schooling outcomes. We 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 short-time effects of the program.[45] This simulation is intended to provide insights about potential long-run effects associated with public works programs such as the NREGS.

To estimate the importance of the risk-reduction effect of the NREGS on time in school, we use the Socio-Economic Profiles of Rural Households in India (SEPRI) data of 2014. It is a follow-up survey to the REDS that was collected in 8 states of India: Andhra Pradesh, Bihar, Chhattisgarh, Haryana, Jharkhand, Madhya Pradesh, Tamil Nadu and Rajasthan. It samples the entire population of the REDS survey villages, but applies a questionnaire that is considerably shorter than in the REDS.[46]

Using the SEPRI data, we estimate the extent to which the presence of the NREGS, or more

[44] Li and Sekhri (2013) also find evidence that the NREGS reduced school attendance.

[45] Given that we estimate the wage elasticity of different activities, we could also simulate a potential wage effect of the NREGS, and attempt to predict a net effect. But due to the cross-sectional nature of our data we do not trust the wage coefficients to be precise enough to perform such an analysis and restrict ourselves to an analysis of the risk mitigating effect of the NREGS.

[46] Summary statistics of the SEPRI data are shown in Table C.16 in the Appendix.

specifically, the amount of employment generated within the NREGS in a given year, mediates the effect of rainfall on household consumption. Formally, we estimate the same equation as in Table 3, but now add employment per capita (in person-days) generated in a given village within the NREGS and its interaction with rainfall to the estimation. Results are reported in Table 12.[47] We find that a one unit increase in employment per capita generated, reduces the effect of rainfall on consumption by 4.7 percentage points.[48] We use these results to predict the standard deviation of predicted log consumption at different levels of employment per capita generated.

To make our simulation comparable to existing studies, we need an estimate of average employment-days created within the NREGS for the time period of these studies. The most well documented paper on the effect of the NREGS is by Imbert and Papp (2015). Using data from 2004/05 and 2007/08, the authors estimate the effects of the NREGS on casual wages. In the financial year 2007/08, the NREGS was implemented in 330 (out of 625) rural districts of India, covering a population of 642 Million.[49] In that year, the total amount of employment-days created within the NREGS was 1,435.9 Million (Ministry of Rural Development, 2012). The average number of person-days of employment generated per capita was thus 2.24 in the implementing districts.

In the SEPRI data, average employment creation per capita within the NREGS is 0.76, with a standard deviation of 1.23, in 2011/12. At the mean of irrigation and the minimum of NREGS employment creation (i.e. zero employment), an increase in NREGS employment by 2.24 days per capita would reduce the standard deviation of log consumption by 0.029 on average. This corresponds to a 43.0% reduction in risk as compared to the baseline of no NREGS employment in 2014 (c.f. Table C.16).

With these estimates we can simulate the risk-mitigating effect of the NREGS on the probability that girls attend school. Using the estimates presented in Table 7, column (2), the reduction in risk exposure by 43% of the mean – as can be attributed to the NREGS – increases the probability of attending school by 2.9 percentage points. Our simulation thus suggests that an employment guarantee such as the NREGS can have positive effects on girls' human capital accumulation by reducing consumption risk. However, from previous literature we know that the extent to which such programs affect adult's wages is also critical. 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 as the NREGS decreases school attendance not only in the first years of program implementation, but permanently.

[47] The rainfall data covers the agricultural year 2012/13, which is presumably the period that determines consumption outcomes in 2014. Due to data limitations we have to use the amount of employment generated in the financial year 2011/12. The NREGS data of 2012/13 is incomplete in the survey.

[48] In both regressions, we drop 1,786 observations in order to reduce the influence of outliers. Outliers are detected using the DFBETA statistic for the share of irrigated land and the usual cutoff value of 2/sqrt(N), with N being the number of observations.

[49] This is the projected population as of March 2007 of Phase I and Phase II districts, thus all districts that implemented the NREGS by May 2008. To get the population estimates, we apply the India-wide projected population growth rate to the Census 2001 population count of these districts. The projected growth rate was obtained from National Commission on Population (2006).

8 Conclusions

This paper develops a model of human capital investment that highlights the joint effect of consumption risk and dynamic complementarity in the production of human capital. The model highlights that risk affects the time girls spend in school as long as three conditions are met. First, there is dynamic complementarity in the education production function. Second, there is substitutability of maternal and daughter time in home production, and third, mother's respond to income shocks by increasing their labor supply *ex post*.

Under these three conditions the effect of risk is consistently negative as long as household utility is assumed to increase linearly in the daughter's final stock of human capital. If the household's marginal utility decreases with increasing stocks of human capital – in other words, if the household is risk-averse with respect to the stock of human capital – the effect of risk on school investment is negative at very high levels of dynamic complementarity, and turns positive as school investments of different time periods become more substitutable.

Using the Young Lives Data of India, we estimate the elasticity of substitution in school times between periods. Using a non-linear instrumental variables approach, we cannot reject substitution elasticities of one. At this level, the model predicts a negative effect of risk only if household utility increases linearly in human capital. In contrast, if household utility is log-linear in human capital, the predicted effect of risk on school time is zero.

We test the model's assumptions and implications in the context of rural India, and find strong evidence for a negative effect of risk on the probability that girls attend school. While the main focus of this paper lies on school attendance, we also find evidence that girls who are more exposed to risk also spend more time on household activities. Taken together, these findings point to the detrimental consequences of risk on human capital accumulation.

This effect can only be explained by dynamic complementarity in the production of human capital, and by a household utility that increases linearly (or close to linearly) in the human capital stock of the child. Such a utility function is consistent with overlapping-generations models as long as we account for households expectations of a convex relationship between years of schooling and future income.

These findings contribute to a better understanding of the consequences of dynamic complementarity in the production of human capital. Dynamic complementarity not only calls for early lifetime interventions in order to remedy the consequences of shocks, and to reduce inequality of opportunities. We also need to think of better ways to support households, particularly in developing countries, in dealing with various sources of consumption risk.

Arguably, a public works program could be a policy that supports households in contexts of high exposure to climate risk, because it off-sets the negative effects of rainfall shocks on wages and incomes. We simulate the effect of the National Rural Employment Guarantee Scheme on the time allocation of girls. We estimate the extent to which NREGS reduces variability in consumption and use this reduction to simulate the effect of the program on girls' time in school. We find that the program would increase girls' school attendance by 2.9 percentage points, holding the level of wages fixed. These results suggest that income transfers, either in the form of cash transfers or public works programs could increase investments in human capital by reducing consumption risk.

However such programs need to be carefully designed. As shown in the case of the NREGS, the fact that public works wages were above market wages let to increasing wages and to declines in the enrollment rates of adolescents. While we cannot assess if the risk effect could have later mitigated the negative wage effect on human capital accumulation that was evident during the program's phase-in, our results do support the idea that policies that support households in smoothing incomes without raising wages, will also increase school investments. Our results also speak to the design of such programs: welfare effects can be maximized by making access to work transparent and reliable, and by differentially targeting such programs at villages that are more regularly affected by adverse shocks.

References

Akresh, R., E. Bagby, D. de Walque, and H. Kazianga (2012). Child ability and household human capital investment decisions in burkina faso. *Economic Development and Cultural Change 61*(1), 157–186.

Attanasio, O., C. Meghir, and E. Nix (2015). *Human capital development and parental investment in India.* NBER Working Paper No. 21740.

Banerjee, A. V., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in india. *American Economic Journal: Economic Policy 2*(1), 1–30.

Banerjee, A. V. and E. Duflo (2005). Growth theory through the lens of development economics. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth* (1 ed.), Volume 1, Part A, Chapter 07, pp. 473–552. Elsevier.

Banerjee, A. V. and E. Duflo (2011). *Poor economics: a radical rethinking of the way to fight global poverty.* PublicAffairs.

Barrera-Osorio, F., M. Bertrand, L. L. Linden, and F. Perez-Calle (2011, April). Improving the design of conditional transfer programs: Evidence from a randomized education experiment in colombia. *American Economic Journal: Applied Economics 3*(2), 167–95.

Becker, G. (1991). *A treatise on the family (enlarged edition).* Cambridge, Mass.: Harvard University Press.

Beegle, K., R. H. Dehejia, and R. Gatti (2006). Child labor and agricultural shocks. *Journal of Development Economics 81*(1), 80–96.

Berg, E., S. Bhattacharyya, R. Durgam, and M. Ramachandra (2012). *Can rural public works affect agricultural wages? Evidence from India.* CSAE Working Paper 2012-05.

Bianchi, M. and M. Bobba (2013). Liquidity, risk, and occupational choices. *Review of Economic Studies 80*(2), 491–511.

Blundell, R., P.-A. Chiappori, T. Magnac, and C. Meghir (2007). Collective labour supply: Heterogeneity and non-participation. *Review of Economic Studies 74*(2), 417–445.

Cameron, C. A. and D. L. Miller (2015). A practitioners guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

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.

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.

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.

Fafchamps, M. and J. Pender (1997). Precautionary Saving, Credit Constraints, and Irreversible Investment: Theory and Evidence From Semiarid India. *Journal of Business & Economic Statistics 15*(2), 180–194.

Foster, A. and M. Rosenzweig (1999). *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), 931–953.

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.

Government of India (2000). *Report of the Time Use Survey 1998-99*. Central Statistical Organization. Ministry of Statistics and Programme Implementation.

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.

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.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry (2014). Agricultural Decisions after Relaxing Credit and Risk Constraints. *Quarterly Journal of Economics 129*(2), 597–652.

Kaur, S. (2014). *Nominal Wage Rigidity in Village Labor Markets*. Working Paper. New York: Columbia University.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature 49*(4), 961–1075.

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.

Kremer, M., N. Chaudhury, F. H. Rogers, K. Muralidharan, and J. Hammer (2005). Teacher Absence in India: A snapshot. *Journal of the European Economic Association 3*(2-3), 658–667.

Li, T. and S. Sekhri (2013). *The unintended consequences of employment based safety net programs.* Manuscript. University of Virginia.

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.

National Commission on Population (2006). *Population Projections for India and States 2001-2026.* Office of the Registrar General & Census Commissioner. Government of India.

National Sample Survey Office (2006). *Employment and Unemployment Situation in India (2004–05).* Ministry of Statistics and Programme Implementation. Government of India.

National Sample Survey Office (2015). *Status of Education and Vocational Trainingi in India (July 2011 - June 2012).* Ministry of Statistics and Programme Implementation. Government of India.

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 – 135.

Pratham (2012). *Annual Status of Education Report (Rural) 2011.*

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.

Rosenzweig, M. R. (1978). Rural Wages, Labor Supply, and Land Reform: A Theoretical and Empirical Analysis. *American Economic Review 68*(5), 847–861.

Rosenzweig, M. R. and H. P. Binswanger (1993). Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments. *Economic Journal 103*(416), 56–78.

Shah, M. and B. M. Steinberg (2015). *Workfare and Human Capital Investment: Evidence from India.* NBER Working Paper No. 21543.

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.

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

Figures

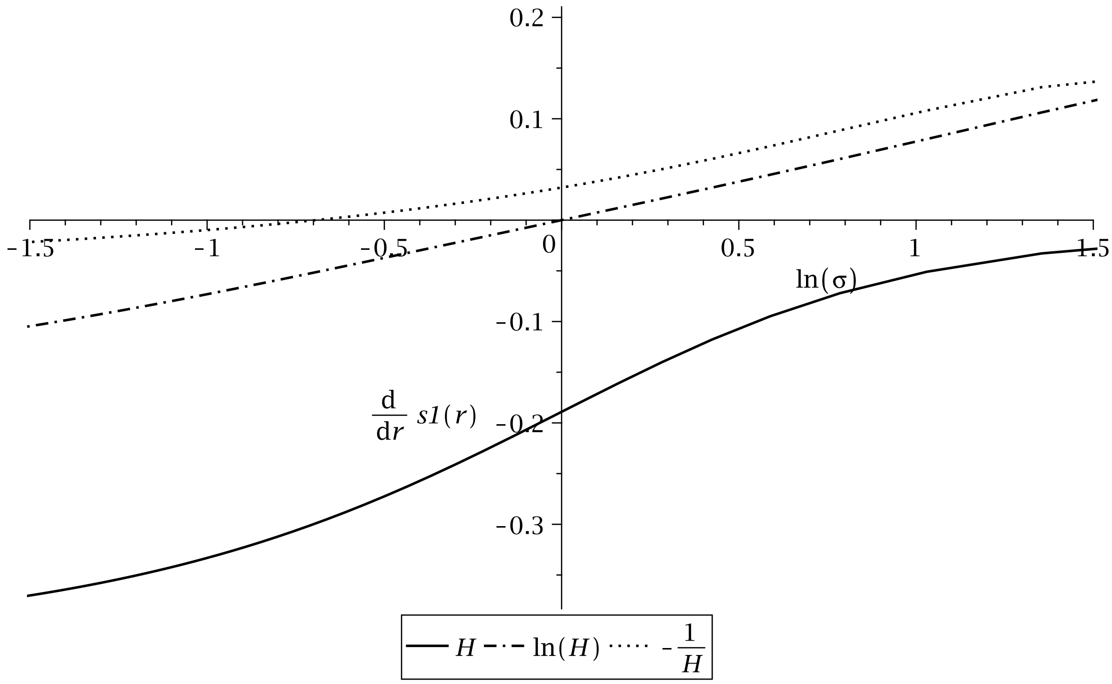

Figure 1: Numeric approximation of the effect of risk

Tables

Table 1: Technology of skill formation

	Younger cohor			Older cohort		
	(1)	(2)	(3)	(4)	(5)	(6)
Scale parameter	0.377***	0.394***	0.364***	0.533***	0.689***	0.683***
	(0.041)	(0.043)	(0.064)	(0.121)	(0.106)	(0.149)
Cognitive skills (t-1)	0.317***	0.315***	0.262*	0.537***	0.641***	0.635***
	(0.071)	(0.073)	(0.107)	(0.159)	(0.113)	(0.156)
School time (t)	0.415***	0.417***	0.479**	0.290	0.133	0.138
	(0.118)	(0.120)	(0.152)	(0.206)	(0.138)	(0.176)
Parental cognition	0.268***	0.268***	0.259***	0.172**	0.226***	0.227***
	(0.062)	(0.061)	(0.071)	(0.060)	(0.050)	(0.057)
Complementarity parameter (ρ)	0.058	-0.010	-0.094	-0.085	0.135	0.126
	(0.270)	(0.278)	(0.314)	(0.430)	(0.379)	(0.458)
Implied elasticity of substitution $\sigma = 1/(1-\rho)$	1.061	1.000	0.914	0.922	1.156	1.144
First stage residual (school time)		0.103***	0.102***		0.114***	0.114***
		(0.018)	(0.019)		(0.019)	(0.020)
First stage residual (skills)			0.020			0.004
			(0.032)			(0.073)
Observations	1197	1197	1197	518	518	518

Estimation in Non-linear Least Squares. Bootstrap standard errors (4999 replications) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Household characteristics

	Mean	SD
Household size	6.14	(2.56)
No of children in household	2.98	(1.52)
Age	47.6	(12.6)
Male	0.92	(0.27)
Married	0.91	(0.29)
Caste: SC/ST	0.27	(0.44)
Religion: Hindu	0.90	(0.30)
Education: no grade	0.39	(0.49)
Education: primary	0.24	(0.43)
Education: secondary	0.23	(0.42)
Education: tertiary	0.14	(0.34)
Hh cultivated any land	0.63	(0.48)
Area cultivated p.c. (acres)	0.42	(0.83)
Annual p.c. income: labor	8710.4	(11560.7)
Annual p.c. income: agriculture	5178.1	(10758.0)
Annual p.c. income: labor (log)	8.31	(1.83)
Annual p.c. income: non-labor	1309.1	(3640.5)
Annual p.c. income: non-labor (log)	5.74	(1.82)
Consumption expenditure per capita	7738.2	(4912.9)
Consumption expenditure per capita (log)	8.82	(0.49)
Total annual precipitation (log)	6.78	(0.51)
Std dev. of log annual rainfall (1960 - 2010)	0.28	(0.096)
Share of irrigated land, village average	0.64	(0.36)
Predicted log consumption	8.81	(0.22)
Interquartile range of predicted log consumption	0.043	(0.033)
SD of predicted log consumption	0.032	(0.025)
Observations	4356	

Notes: All values in current INR. Age, male, married, caste, religion and education refer to the household head.

Table 3: Determinants of consumption per capita (log)

	(1)	(2)
Total annual precipitation (log)	0.272**	0.313***
	(0.091)	(0.079)
Share of irrigated land, village average	1.723**	1.882***
	(0.551)	(0.493)
Total annual precipitation × Share of irrigated land	-0.231**	-0.251***
	(0.081)	(0.073)
Average of annual rainfall (1960 - 2010), log	-0.105	-0.127
	(0.082)	(0.077)
Religion: Hindu		0.097
		(0.062)
Caste: SC/ST		-0.158***
		(0.020)
Education: primary		0.044*
		(0.019)
Education: secondary		0.110***
		(0.019)
Education: tertiary and higher		0.238***
		(0.024)
Area cultivated per capita (acres, log)		0.388***
		(0.030)
Observations	4356	4356
R^2	0.213	0.357

Notes: OLS estimation. State fixed effects included in all specifications but not reported. Caste, religion, and education refer to household head. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Time allocation of daughters and mothers

	Daughters			Mothers		
	N	Mean	SD	N	Mean	SD
Season 1						
Hours per day: chores	4306	1.59	(2.20)			
Hours per day: studying	4306	6.15	(3.54)			
Hours per day: labor supply	4306	0.12	(0.88)			
Average wage (per hour)				914	7.13	(6.11)
Predicted wage (per hour)				4651	9.14	(3.56)
Predicted wage (log)				4651	2.15	(0.34)
Labor force participation				4651	0.21	(0.41)
Household chores (hours per season)				4651	701.4	(279.9)
Labor supply (hours per season)				4651	84.2	(185.2)
Own agr. production (hours per season)				4651	65.5	(135.8)
Season 2						
Hours per day: chores	4306	1.63	(2.27)			
Hours per day: studying	4306	6.19	(3.54)			
Hours per day: labor supply	4306	0.11	(0.88)			
Average wage (per hour)				905	7.27	(6.15)
Predicted wage (per hour)				4651	7.53	(2.63)
Predicted wage (log)				4651	1.97	(0.29)
Labor force participation				4651	0.21	(0.41)
Household chores (hours per season)				4651	700.4	(285.3)
Labor supply (hours per season)				4651	82.9	(185.3)
Own agr. production (hours per season)				4651	49.8	(123.6)
Season 3						
Hours per day: chores	4306	1.68	(2.31)			
Hours per day: studying	4306	5.79	(3.51)			
Hours per day: labor supply	4306	0.079	(0.78)			
Average wage (per hour)				631	7.55	(7.40)
Predicted wage (per hour)				4651	6.98	(2.12)
Predicted wage (log)				4651	1.91	(0.26)
Labor force participation				4651	0.15	(0.36)
Household chores (hours per season)				4651	711.7	(300.5)
Labor supply (hours per season)				4651	57.4	(154.5)
Own agr. production (hours per season)				4651	12.5	(66.7)

Notes: Season 1 & 2 are agricultural peak seasons, season 3 is lean season. Unit of observation is the mother-daughter match. Sample of daughters consists of all girls aged 6 to 16, who live with their mothers.

Table 5: Mothers' labor supply response to shocks

Outcome:	Labor supply		Household chores
Margin:	Extensive	Intensive	Intensive
	(1)	(2)	(3)
Total annual precipitation (log, normalized)	-0.034	25.531	-6.436
	(0.026)	(20.788)	(21.256)
Lean season	-0.059***	-27.014**	-4.435
	(0.006)	(9.167)	(7.115)
Lean season × Total annual precipitation	0.009	-37.547+	12.032
	(0.012)	(20.292)	(14.162)
Observations	13953	2679	2679
R^2	0.190	0.221	0.321

Notes: OLS estimation. Controls: Age, caste, religion, education, household size, number of children in the household, land ownership, non-labor income and other member's labor income, and state fixed effects. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Daughters' response to shocks

Outcome:	School attendance		Household chores
Margin:	Extensive	Intensive	Intensive
	(1)	(2)	(3)
Total annual precipitation (log, normalized)	0.017	0.118	-0.123
	(0.032)	(0.231)	(0.104)
Lean season	-0.008	-0.410***	0.074***
	(0.005)	(0.040)	(0.014)
Lean season × Total annual precipitation	0.001	0.333***	-0.033
	(0.007)	(0.077)	(0.021)
Observations	12918	10232	10232
R^2	0.060	0.217	0.321

Notes: OLS estimation. Controls: Age, caste, religion, household size, number of children in the household, land ownership, non-labor income and other member's labor income, and state fixed effects. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Daughters' time in school in hours per day

	Extensive margin		Intensive margin	
	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	-1.753***		3.488	
	(0.433)		(2.842)	
SD of predicted log consumption		-2.098***		6.408+
		(0.542)		(3.724)
Mother's predicted wage	0.123***	0.116***	-0.159	-0.190
	(0.023)	(0.023)	(0.207)	(0.205)
Age	-0.016***	-0.016***	0.120***	0.120***
	(0.003)	(0.003)	(0.013)	(0.013)
Household size	0.009*	0.009*	0.021	0.021
	(0.004)	(0.004)	(0.025)	(0.025)
No of children in household	-0.009	-0.010	-0.007	-0.008
	(0.009)	(0.009)	(0.043)	(0.043)
Caste: SC / ST	-0.055**	-0.057**	0.003	0.007
	(0.020)	(0.020)	(0.115)	(0.115)
Religion: Hindu	0.089**	0.093**	0.574*	0.579*
	(0.030)	(0.030)	(0.233)	(0.233)
Area cultivated per capita (acres, log)	0.041+	0.038+	-0.007	-0.016
	(0.023)	(0.023)	(0.146)	(0.146)
Annual p.c. income: other hh members' labor (log)	-0.005	-0.005	0.006	0.007
	(0.003)	(0.003)	(0.019)	(0.019)
Annual p.c. income: non-labor (log)	0.015***	0.016***	0.045+	0.045+
	(0.004)	(0.004)	(0.024)	(0.024)
Total annual precipitation (log)	-0.046	-0.038	0.344	0.382
	(0.030)	(0.031)	(0.232)	(0.235)
Bootstrap p-value of main treatment variable	0.004	0.008	0.306	0.246
Observations	12918	12918	10232	10232
R^2	0.073	0.070	0.218	0.219

Notes: OLS estimation. State and season fixed effects included, but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Daughters' probability of attending school and marriage market risk

	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	-2.226***	-2.187***		
	(0.463)	(0.491)		
Average IQR in villages belonging to same marriage market		-0.162		
		(0.506)		
SD of predicted log consumption			-2.603***	-2.633***
			(0.599)	(0.649)
Average SD in villages belonging to same marriage market				0.125
				(0.572)
Mother's predicted wage	0.118***	0.117***	0.111***	0.112***
	(0.025)	(0.025)	(0.025)	(0.025)
Age	-0.015***	-0.015***	-0.015***	-0.015***
	(0.003)	(0.003)	(0.003)	(0.003)
Household size	0.010*	0.010*	0.010*	0.010*
	(0.005)	(0.005)	(0.005)	(0.005)
No of children in household	-0.014	-0.014	-0.015	-0.015
	(0.010)	(0.010)	(0.011)	(0.011)
Caste: SC / ST	-0.048*	-0.048*	-0.050*	-0.050*
	(0.021)	(0.021)	(0.022)	(0.022)
Religion: Hindu	0.069*	0.068*	0.076*	0.076*
	(0.033)	(0.033)	(0.033)	(0.033)
Area cultivated per capita (acres, log)	0.059*	0.060*	0.055*	0.054*
	(0.025)	(0.025)	(0.025)	(0.025)
Annual p.c. income: other hh members' labor (log)	-0.005	-0.005	-0.005	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)
Annual p.c. income: non-labor (log)	0.014**	0.014**	0.015**	0.015**
	(0.005)	(0.005)	(0.005)	(0.005)
Total annual precipitation (log)	-0.070*	-0.072*	-0.056+	-0.055+
	(0.033)	(0.032)	(0.034)	(0.033)
Observations	10905	10905	10905	10905
R^2	0.081	0.081	0.077	0.077

Notes: OLS estimation. 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. State and season fixed effects included, but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Daughters' time in home production in hours per day

	Extensive margin		Intensive margin	
	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	0.881*		2.124	
	(0.379)		(1.964)	
SD of predicted log consumption		1.379**		2.790
		(0.524)		(2.567)
Mother's predicted wage	-0.003	-0.007	-0.417*	-0.413*
	(0.029)	(0.029)	(0.194)	(0.193)
Age	0.071***	0.071***	0.295***	0.295***
	(0.003)	(0.003)	(0.020)	(0.019)
Household size	-0.012*	-0.012*	-0.098**	-0.098**
	(0.005)	(0.005)	(0.030)	(0.030)
No of children in household	0.004	0.004	0.100*	0.101*
	(0.008)	(0.008)	(0.049)	(0.049)
Caste: SC / ST	-0.006	-0.005	0.048	0.051
	(0.019)	(0.019)	(0.120)	(0.121)
Religion: Hindu	0.064	0.064	0.103	0.100
	(0.039)	(0.039)	(0.204)	(0.204)
Area cultivated per capita (acres, log)	-0.042	-0.043	0.227	0.230
	(0.026)	(0.026)	(0.173)	(0.173)
Annual p.c. income: other hh members' labor (log)	-0.002	-0.001	-0.038	-0.038
	(0.004)	(0.004)	(0.026)	(0.026)
Annual p.c. income: non-labor (log)	0.005	0.005	0.056	0.056
	(0.005)	(0.005)	(0.041)	(0.041)
Total annual precipitation (log)	-0.054	-0.050	0.250	0.249
	(0.035)	(0.035)	(0.259)	(0.259)
Bootstrap p-value of main treatment variable	0.058	0.044	0.325	0.340
Observations	12918	12918	6817	6817
R^2	0.384	0.385	0.225	0.225

Notes: OLS estimation. State and season fixed effects included, but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Mothers' time allocation

Outcome:	Labor supply				Household chores	
Margin:	Extensive		Intensive		Intensive	
	(1)	(2)	(3)	(4)	(5)	(6)
Interquartile range of predicted log consumption	-0.538		677.390*		-1022.691*	
	(0.353)		(313.069)		(428.208)	
SD of predicted log consumption		-0.377		650.546^{+}		-1254.573*
		(0.473)		(383.838)		(535.414)
Predicted wage (log)	0.152***	0.144***	51.391^{+}	54.025^{+}	-61.176^{+}	-62.538^{+}
	(0.041)	(0.041)	(29.152)	(29.366)	(32.626)	(32.501)
No of children in household	0.021***	0.021***	5.842	5.484	10.662	10.819
	(0.006)	(0.006)	(7.419)	(7.437)	(10.074)	(10.145)
Household size	-0.023***	-0.023***	4.085	4.397	-22.453***	-22.683***
	(0.003)	(0.003)	(4.906)	(4.934)	(6.037)	(6.113)
Married	-0.142***	-0.142***	-60.577*	-60.976*	73.408**	74.540**
	(0.032)	(0.032)	(27.385)	(27.399)	(27.073)	(27.126)
Mother's age	0.001	0.001	0.426	0.413	-2.637**	-2.604**
	(0.001)	(0.001)	(0.752)	(0.747)	(0.962)	(0.963)
Caste: SC / ST	0.110***	0.109***	20.032	20.361	-29.479^{+}	-30.186^{+}
	(0.021)	(0.021)	(12.940)	(13.006)	(17.045)	(16.990)
Religion: Hindu	0.025	0.027	12.092	10.955	-31.976	-31.118
	(0.022)	(0.023)	(17.857)	(17.929)	(23.413)	(23.652)
Education: primary	-0.052**	-0.051**	18.681	17.780	28.166	28.572
	(0.016)	(0.016)	(19.932)	(20.017)	(22.446)	(22.457)
Education: secondary	-0.078***	-0.077***	69.804**	68.092**	32.145	33.072
	(0.016)	(0.016)	(21.775)	(21.982)	(24.879)	(25.005)
Education: tertiary and higher	-0.088***	-0.086***	113.852**	111.410**	-23.171	-22.481
	(0.025)	(0.024)	(39.894)	(39.835)	(39.566)	(39.525)
Annual p.c. income: non-labor (log)	-0.004	-0.004	-7.078^{+}	-7.084^{+}	2.578	2.418
	(0.004)	(0.004)	(3.682)	(3.649)	(4.223)	(4.264)
Annual p.c. income: other hh members' labor (log)	0.005	0.005	-5.897^{+}	-6.080^{+}	-2.326	-2.225
	(0.003)	(0.003)	(3.366)	(3.353)	(3.290)	(3.314)
Area cultivated p.c. (acres, log)	-0.249***	-0.251***	-75.515	-73.717	-8.844	-9.673
	(0.026)	(0.026)	(55.286)	(55.655)	(40.831)	(40.016)
Total annual precipitation (log)	-0.061*	-0.053^{+}	30.828	24.741	-28.178	-25.275
	(0.028)	(0.028)	(21.584)	(21.977)	(25.843)	(25.002)
Bootstrap p-value of main treatment variable	0.213	0.491	0.150	0.215	0.082	0.071
Observations	13953	13953	2679	2679	2679	2679
R^2	0.194	0.193	0.228	0.225	0.331	0.330

Notes: OLS estimation. State and season fixed effects included, but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Savings

	Savingsrate		Savings to cons.	
	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	-0.114		-1.128	
	(0.244)		(1.260)	
SD of predicted log consumption		-0.318		-2.129
		(0.318)		(1.807)
No of children in household	0.010***	0.010***	0.049**	0.049**
	(0.003)	(0.003)	(0.017)	(0.017)
Married	-0.023	-0.023	0.004	0.004
	(0.015)	(0.015)	(0.073)	(0.072)
Age	0.001***	0.001***	0.006**	0.006**
	(0.000)	(0.000)	(0.002)	(0.002)
Caste: SC/ST	0.010	0.009	0.058	0.056
	(0.011)	(0.011)	(0.061)	(0.060)
Religion: Hindu	-0.029	-0.030+	-0.104	-0.106
	(0.018)	(0.018)	(0.084)	(0.084)
Education: primary	0.017	0.016	0.013	0.013
	(0.011)	(0.011)	(0.056)	(0.056)
Education: secondary	0.011	0.011	0.064	0.065
	(0.011)	(0.011)	(0.060)	(0.060)
Education: tertiary and higher	0.068***	0.068***	0.413***	0.413***
	(0.013)	(0.013)	(0.100)	(0.100)
Annual p.c. income: non-labor (log)	0.016***	0.016***	0.074***	0.074***
	(0.003)	(0.003)	(0.014)	(0.014)
Annual p.c. income: labor (log)	0.048***	0.048***	0.231***	0.230***
	(0.007)	(0.007)	(0.038)	(0.038)
Area cultivated per capita (acres, log)	0.118***	0.118***	0.774***	0.776***
	(0.017)	(0.017)	(0.132)	(0.131)
Total annual precipitation (log)	0.003	-0.001	-0.008	-0.023
	(0.017)	(0.017)	(0.109)	(0.108)
Observations	2316	2316	2316	2316
R^2	0.264	0.264	0.269	0.270

Notes: OLS estimation. State fixed effects included, but not reported. Age, marital status, caste, religion, and education refer to household head. Standard errors (clustered at the village level) in parentheses. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Determinants of consumption per capita (log) SEPRI

	(1)	(2)
Total annual rainfall (log)	0.277^{+}	0.339^{*}
	(0.148)	(0.157)
Share of irrigated land, village average	0.682	1.025
	(0.804)	(0.781)
Total annual rainfall × Share of irrigated land	-0.076	-0.125
	(0.125)	(0.121)
Average of annual rainfall (1960 - 2010), log	-0.304^{+}	-0.311^{*}
	(0.159)	(0.154)
Employment generated per capita in 2011-12, NREGA		0.313^{**}
		(0.110)
Total annual rainfall (log) × Employment generated per capita		-0.047^{**}
		(0.016)
Observations	46766	46766
R^2	0.377	0.379

OLS estimation. State xed eects and linear time trend included in all specications but not reported. Inuential outliers excluded using the DFBETA statistics on the share of irrigated land and the usual cutos. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

APPENDIX

A Data Appendix

A.1 YLS data

The Young Lives survey was collected between 2002 and 2013-14 in Andhra Pradesh, a state in the south of India. Three rounds of data were available at the time of writing this article, and given our focus on school age children we restrict our analysis to rounds 2 (2007) and 3 (2009-10). Summary statistics for the variables used in this analysis are reported in Table C.1 of this Appendix.

The data contains various features that are crucial to estimating a production function (and specifically the complementarity parameters) of cognitive skills. First of all YLS performs different cognitive tests: a Peabody Picture Vocabulary Test (PPVT), a math test, a reading comprehension test (EGRA), a Cognitive Development Assessment (CDA) and a verbal test (CLOZE). The test scores of all these tests are internally standardized. Over time and across cohorts, different tests are being administered. For the younger cohort we use the PPVT and CDA as measurements of past cognition, and the PPVT, the math test and the EGRA for current cognition. For the older cohort, we use the PPVT and the math test as measurements for past cognition, and the PPVT, the math test and the CLOZE for current cognition. We use reported time in school and studying as the only investment variable in our production function. Finally, we use the years of education completed by the mother and the father of each child as measurements for parental cognition.

B Supplementary Figures

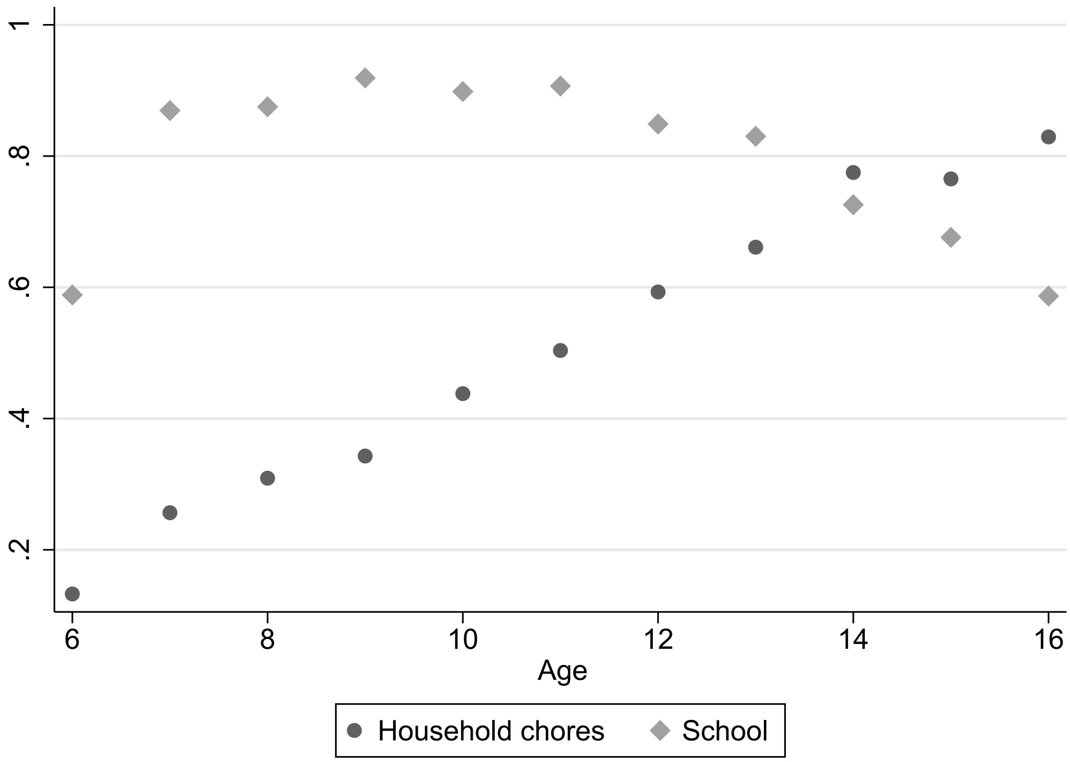

Figure B.1: Probability to perform household chores and being in school

C Supplementary Tables

Table C.1: Household and child characteristics

	Younger cohort		Older cohort	
	Age 5	Age 8	Age 12	Age 15
Owned land (in acres)	1.91	1.91	2.31	2.31
	(3.30)	(3.30)	(5.76)	(5.76)
Precipitation (log)	6.79	6.64	6.79	6.66
	(0.25)	(0.45)	(0.25)	(0.45)
Father's grade	4.65	4.65	4.35	4.35
	(4.66)	(4.66)	(4.70)	(4.70)
Mother's grade	2.81	2.81	2.32	2.32
	(3.96)	(3.96)	(3.49)	(3.49)
Hours in school/ studying	6.63	9.55	8.40	9.16
	(2.58)	(1.31)	(2.73)	(4.04)
Rash score PPVT Test	296.1	300.0	301.5	301.5
	(47.0)	(14.8)	(46.4)	(14.0)
Rash score CDA Test	295.9			
	(48.3)			
Rash score Math Test		299.4	305.0	299.9
		(15.2)	(42.6)	(14.7)
Rasch score EGRA Test		300.3		
		(14.9)		
Rasch score CLOZE Test				300.7
				(14.6)
Observations	1362	1362	605	605

Notes: Standard deviations are reported in parentheses.

Table C.2: Individual characteristics of daughters and mothers

	Daughters			Mothers		
	N	Mean	SD	N	Mean	SD
No of children in household	4306	3.77	(1.79)	4651	3.14	(1.71)
Household size	4306	7.00	(3.00)	4651	6.50	(3.00)
Married	4306	0.00093	(0.030)	4651	0.96	(0.19)
Age	4306	11.2	(3.14)	4651	36.7	(7.93)
Caste: SC / ST	4306	0.28	(0.45)	4651	0.26	(0.44)
Religion: Hindu	4306	0.89	(0.32)	4651	0.90	(0.30)
Education: no grade	4306	0.12	(0.33)	4651	0.59	(0.49)
Education: primary	4306	0.55	(0.50)	4651	0.17	(0.38)
Education: secondary	4306	0.27	(0.45)	4651	0.16	(0.37)
Education: tertiary	4306	0.054	(0.23)	4651	0.071	(0.26)
Years of Schooling	4306	4.58	(3.12)	4651	3.09	(4.20)
Presently enrolled	4306	0.79	(0.40)	4651	0.0084	(0.091)
Hrs p year: agr. casual labor	4306	5.68	(73.6)	4651	147.7	(385.7)
Hrs p year: agr. casual labor (ffw)	4306	0.073	(3.00)	4651	3.15	(29.8)
Hrs p year: own agr. production	4306	2.64	(31.8)	4651	127.8	(292.0)
Hrs p year: own livestock production	4306	45.6	(144.6)	4651	300.9	(369.8)
Hrs p year: public works	4306	0.022	(1.24)	4651	7.27	(57.3)
Hrs p year: non-agr. casual labor	4306	0.80	(38.7)	4651	29.6	(207.0)
Hrs p year: migration	4306	0	(0)	4651	3.24	(66.0)
Hrs p year: self-employed	4306	0.70	(22.5)	4651	17.3	(153.8)
Hrs p year: construction (own)	4306	0.79	(6.88)	4651	17.3	(81.6)
Hrs p year: household work	4306	291.5	(447.3)	4651	1795.3	(710.8)
Hrs p year: CPR	4306	0.025	(1.63)	4651	2.79	(43.5)
Hrs p year: permanent employment	4306	0	(0)	4651	16.3	(162.8)
Hours worked (per year)	4306	1916.0	(828.2)	4651	2840.2	(810.6)
Labor supply (hours per year)	4306	9.92	(94.7)	4651	352.3	(549.7)
Household chores (hours per year)	4306	337.9	(526.4)	4651	2113.5	(826.0)
Annual p.c. income: non-labor	4306	1215.1	(3476.4)	4651	1308.1	(3585.3)
Annual p.c. income: non-labor (log)	4306	5.64	(1.82)	4651	5.74	(1.82)
Annual p.c. income: other hh members' labor	4306	6711.3	(9059.5)	4651	7669.2	(10970.7)
Annual p.c. income: other hh members' labor (log)	4306	7.92	(2.08)	4651	8.02	(2.13)

Notes: All values in current INR. Sample of daughters consists of all girls aged 6 to 16, who live with their mothers. Sample of mothers consist of all mothers, who have cohabiting children in the age group 6 to 16, irrespective of their gender.

Table C.3: Time allocation of boys

	N	Mean	SD
No of children in household	5045	3.42	(1.72)
Household size	5045	6.61	(2.86)
Married	5045	0.00040	(0.020)
Age	5045	11.2	(3.08)
Caste: SC / ST	5045	0.26	(0.44)
Religion: Hindu	5045	0.89	(0.31)
Education: no grade	5045	0.078	(0.27)
Education: primary	5045	0.55	(0.50)
Education: secondary	5045	0.32	(0.46)
Education: tertiary	5045	0.056	(0.23)
Years of Schooling	5045	4.94	(3.08)
Presently enrolled	5045	0.86	(0.34)
Hrs p year: agr. casual labor	5045	8.19	(95.8)
Hrs p year: agr. casual labor (ffw)	5045	0	(0)
Hrs p year: own agr. production	5045	7.39	(64.2)
Hrs p year: own livestock production	5045	47.3	(144.8)
Hrs p year: public works	5045	0	(0)
Hrs p year: non-agr. casual labor	5045	9.78	(131.5)
Hrs p year: migration	5045	2.21	(61.5)
Hrs p year: self-employed	5045	2.45	(47.8)
Hrs p year: construction (own)	5045	0.56	(5.41)
Hrs p year: household work	5045	172.1	(276.0)
Hrs p year: CPR	5045	0.49	(16.3)
Hrs p year: other	5045	1575.9	(788.8)
Hrs p year: permanent employment	5045	1.80	(64.7)
Hours worked (per year)	5045	1828.2	(798.8)
Labor supply (hours per year)	5045	31.8	(213.0)
Household chores (hours per year)	5045	220.5	(361.4)
Annual p.c. income: non-labor	5045	1250.7	(3325.7)
Annual p.c. income: non-labor (log)	5045	5.72	(1.78)
Annual p.c. income: other hh members' labor	5045	7163.9	(10092.3)
Annual p.c. income: other hh members' labor (log)	5045	7.94	(2.15)
Hours per day: chores (season 1)	5045	1.13	(1.66)
Hours per day: studying (season 1)	5045	6.69	(3.18)
Hours per day: labor supply (season 1)	5045	0.27	(1.38)

Notes: Sample of boys consists of all boys aged 6 to 16, who live with their mothers. Season 1 is the agricultural peak season.

Table C.4: Daughters' time in home production in hours per season

	Extensive margin		Intensive margin	
	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	0.874*		196.838	
	(0.352)		(152.826)	
SD of predicted log consumption		1.334**		361.838+
		(0.489)		(199.312)
Mother's predicted wage	-0.058+	-0.061+	-33.794**	-36.427**
	(0.031)	(0.031)	(11.671)	(11.799)
Age	0.066***	0.067***	28.895***	28.967***
	(0.003)	(0.003)	(1.755)	(1.747)
Household size	-0.009+	-0.009+	-9.710***	-9.728***
	(0.005)	(0.005)	(2.505)	(2.496)
No of children in household	0.001	0.002	11.078*	11.143*
	(0.008)	(0.008)	(4.505)	(4.473)
Caste: SC / ST	-0.002	-0.002	-4.784	-4.737
	(0.019)	(0.019)	(10.657)	(10.624)
Religion: Hindu	0.073*	0.073*	5.375	6.122
	(0.036)	(0.035)	(18.446)	(18.446)
Area cultivated per capita (acres, log)	-0.023	-0.024	5.197	4.582
	(0.027)	(0.027)	(16.228)	(16.289)
Annual p.c. income: other hh members' labor (log)	-0.007+	-0.007+	-0.608	-0.564
	(0.004)	(0.004)	(2.109)	(2.119)
Annual p.c. income: non-labor (log)	0.002	0.002	1.187	1.194
	(0.005)	(0.005)	(2.378)	(2.383)
Total annual precipitation (log)	-0.060+	-0.057+	19.925	22.535
	(0.033)	(0.032)	(18.606)	(18.690)
Observations	12918	12918	6383	6383
R^2	0.414	0.414	0.293	0.293

Notes: OLS estimation. State and season fixed effects included, but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.5: Selection equation for casual wages

Season	1	2	3
Age	0.104***	0.115***	0.096***
	(0.024)	(0.024)	(0.027)
Age (squared)	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Caste: SC / ST	0.535***	0.530***	0.446***
	(0.056)	(0.055)	(0.062)
Religion: Hindu	0.078	0.069	0.360**
	(0.103)	(0.102)	(0.128)
No of children in household	-0.088***	-0.081***	-0.036+
	(0.019)	(0.018)	(0.021)
Education: primary	-0.315***	-0.338***	-0.230**
	(0.077)	(0.077)	(0.084)
Education: secondary	-0.661***	-0.662***	-0.642***
	(0.090)	(0.089)	(0.102)
Education: tertiary and higher	-0.985***	-1.027***	-0.939***
	(0.146)	(0.149)	(0.165)
Total annual precipitation (log)	-0.333***	-0.373***	-0.313**
	(0.087)	(0.087)	(0.096)
Area cultivated p.c. (acres, log)	-2.679***	-2.452***	-2.098***
	(0.144)	(0.136)	(0.148)
Annual p.c. income: non-labor (log)	-0.026+	-0.028+	-0.007
	(0.015)	(0.015)	(0.017)
Interquartile range of predicted log consumption	-2.496*	-1.689	-0.494
	(1.140)	(1.118)	(1.244)
Annual p.c. income: other hh members' labor (log)	0.036**	0.042**	0.023
	(0.013)	(0.014)	(0.015)
Married	-0.725***	-0.714***	-0.710***
	(0.129)	(0.128)	(0.137)
Observations	4601	4601	4601

Notes: OLS estimation. State fixed effects included in all specifications but not reported. Standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.6: Hourly wages (log)

	Casual jobs			Permanent jobs
Season	1	2	3	1-3
Age	0.009	0.009	0.027	0.142
	(0.011)	(0.012)	(0.016)	(0.262)
Age (squared)	-0.000	-0.000	-0.000+	-0.002
	(0.000)	(0.000)	(0.000)	(0.003)
Caste: SC / ST	-0.064+	-0.031	-0.030	0.354
	(0.034)	(0.036)	(0.044)	(0.373)
Religion: Hindu	-0.057	-0.035	0.065	0.862
	(0.044)	(0.046)	(0.082)	(0.769)
No of children in household	0.013	0.009	0.002	0.138
	(0.010)	(0.010)	(0.013)	(0.129)
Education: primary	0.064*	0.037	0.025	-0.078
	(0.032)	(0.035)	(0.043)	(0.568)
Education: secondary	0.069	0.019	0.002	0.266
	(0.052)	(0.055)	(0.076)	(0.490)
Education: tertiary and higher	0.151+	0.064	0.012	1.220+
	(0.082)	(0.088)	(0.115)	(0.689)
Total annual precipitation (log)	0.122**	0.088*	0.178**	-0.723
	(0.040)	(0.042)	(0.056)	(0.531)
Area cultivated p.c. (acres, log)	0.141	-0.082	-0.222	0.074
	(0.160)	(0.156)	(0.193)	(0.329)
Annual p.c. income: non-labor (log)	0.011+	0.006	0.008	0.057
	(0.006)	(0.006)	(0.008)	(0.136)
Interquartile range of predicted log consumption	1.039*	0.858+	1.951**	-3.289
	(0.487)	(0.489)	(0.647)	(8.306)
Mills				
Lambda	-0.144+	-0.037	-0.011	
	(0.082)	(0.085)	(0.115)	
Observations	4601	4601	4601	50
R^2				0.484

Notes: OLS estimation. State fixed effects included in all specifications but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.7: Daughters' probability of attending school by age group

	Age below 11		Age 11 and above	
	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	-1.281***		-2.081***	
	(0.357)		(0.555)	
SD of predicted log consumption		-1.480**		-2.525***
		(0.449)		(0.703)
Mother's predicted wage	0.097***	0.091***	0.151***	0.144***
	(0.026)	(0.026)	(0.029)	(0.028)
Age	0.065***	0.065***	-0.066***	-0.066***
	(0.007)	(0.007)	(0.005)	(0.005)
Household size	0.010^{+}	0.010*	0.017**	0.017**
	(0.005)	(0.005)	(0.006)	(0.006)
No of children in household	-0.014	-0.015	-0.023*	-0.024*
	(0.011)	(0.011)	(0.012)	(0.012)
Caste: SC / ST	-0.043^{+}	-0.045^{+}	-0.073**	-0.075**
	(0.025)	(0.026)	(0.024)	(0.024)
Religion: Hindu	0.042	0.046	0.122**	0.126**
	(0.037)	(0.036)	(0.040)	(0.041)
Area cultivated per capita (acres, log)	0.033	0.032	0.048^{+}	0.045
	(0.032)	(0.032)	(0.029)	(0.029)
Annual p.c. income: other hh members' labor (log)	-0.002	-0.002	-0.006	-0.006
	(0.004)	(0.004)	(0.004)	(0.004)
Annual p.c. income: non-labor (log)	0.014**	0.014**	0.019**	0.020**
	(0.005)	(0.005)	(0.006)	(0.006)
Total annual precipitation (log)	-0.041^{+}	-0.035	-0.046	-0.038
	(0.024)	(0.025)	(0.043)	(0.044)
Observations	5646	5646	7272	7272
R^2	0.119	0.117	0.141	0.138

Notes: OLS estimation. State and season fixed effects included, but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.8: Daughter's probability of attending school conditional on school achievement

	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)
IQR of predicted log consumption	-1.543***	-1.359***	-1.677***	-1.710***	-0.944***	-0.689**
	(0.402)	(0.395)	(0.354)	(0.458)	(0.219)	(0.215)
Completed primary	0.145***	0.315***			0.644***	0.875***
	(0.029)	(0.061)			(0.025)	(0.068)
Completed secondary			0.327***	-0.136	0.828***	0.898***
			(0.020)	(0.088)	(0.024)	(0.110)
Observations	12918	12918	12918	12918	12918	12918
R^2	0.158	0.105	0.221	0.025	0.448	0.393
Kleibergen F statistic		15.202		12.049		8.952

Notes: OLS estimation. All columns control for age, age squared, household size, number of children in the household, caste, religion, precipitation, as well as state and season fixed effects. Instruments are five lags of precipitation data interacted with dummies indicating if the child was of school age in that particular year. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.9: Daughters' probability of attending school and wealth controls

	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	-1.738***	-1.753***	-1.724***	-1.657***
	(0.437)	(0.433)	(0.443)	(0.444)
Area cultivated per capita (acres, log)		0.041+	0.035	0.086*
		(0.023)	(0.050)	(0.035)
Annual p.c. income: other hh members' labor (log)		-0.005	-0.022*	-0.005
		(0.003)	(0.011)	(0.003)
Annual p.c. income: non-labor (log)		0.015***	0.006	0.015***
		(0.004)	(0.010)	(0.004)
Area cultivated per capita (squared)			-0.011	
			(0.034)	
Annual p.c. income: other hh members' labor (squared)			0.002+	
			(0.001)	
Annual p.c. income: non-labor (squared)			0.001	
			(0.001)	
IQR of predicted log consumption × Area cultivated				-0.780
				(0.537)
Observations	12918	12918	12918	12918
R^2	0.067	0.073	0.074	0.073

Notes: OLS estimation. All columns control for age, household size, number of children in the household, caste, religion, precipitation, as well as state and season fixed effects. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.10: Daughters' probability of attending school and irrigation controls

	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	-1.927***	-2.095**	-2.081**	-2.218***
	(0.520)	(0.645)	(0.646)	(0.614)
Irrigation (share of own agr. land)		-0.021	-0.048	0.000
		(0.037)	(0.066)	(0.074)
Irrigation (squared)			-0.086	
			(0.158)	
IQR of predicted log consumption × Irrigation				-0.403
				(1.143)
Observations	8274	8274	8274	8274
R^2	0.082	0.082	0.082	0.082

Notes: OLS estimation. All columns control for age, household size, number of children in the household, caste, religion, precipitation, as well as state and season fixed effects. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.11: Sons' response to shocks

Outcome:	School attendance		Household chores	Labor supply
Margin:	Extensive	Intensive	Intensive	Intensive
	(1)	(2)	(3)	(4)
Total annual precipitation (log, normalized)	0.023	-0.069	-0.008	-0.010
	(0.020)	(0.220)	(0.110)	(0.018)
Lean season	-0.007^{+}	-0.434***	0.046***	-0.020**
	(0.004)	(0.040)	(0.011)	(0.006)
Lean season × Total annual precipitation	-0.007	0.281***	-0.037*	0.016^{+}
	(0.006)	(0.078)	(0.017)	(0.010)
Observations	15135	13080	13080	13080
R^2	0.041	0.191	0.300	0.018

Notes: OLS estimation. Controls: Age, caste, religion, household size, number of children in the household, land ownership, non-labor income and other member's labor income, and state fixed effects. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.12: Sons' time in school in hours per day

	Extensive margin		Intensive margin	
	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	-0.459		1.808	
	(0.323)		(3.128)	
SD of predicted log consumption		-0.430		4.827
		(0.378)		(3.960)
Mother's predicted wage	0.056**	0.052**	0.180+	0.127
	(0.018)	(0.017)	(0.096)	(0.100)
Age	-0.007**	-0.007**	0.106***	0.105***
	(0.002)	(0.002)	(0.011)	(0.011)
Household size	0.001	0.002	0.036	0.037
	(0.004)	(0.004)	(0.024)	(0.024)
No of children in household	-0.002	-0.002	-0.063	-0.064
	(0.007)	(0.008)	(0.044)	(0.044)
Caste: SC / ST	-0.059***	-0.059***	-0.021	-0.016
	(0.017)	(0.017)	(0.103)	(0.102)
Religion: Hindu	0.015	0.016	0.563**	0.572**
	(0.023)	(0.022)	(0.201)	(0.199)
Area cultivated per capita (acres, log)	0.053**	0.053**	-0.004	-0.015
	(0.018)	(0.018)	(0.142)	(0.142)
Annual p.c. income: other hh members' labor (log)	-0.001	-0.001	0.003	0.003
	(0.002)	(0.002)	(0.018)	(0.018)
Annual p.c. income: non-labor (log)	0.012***	0.012***	0.055*	0.055*
	(0.003)	(0.003)	(0.023)	(0.023)
Total annual precipitation (log)	0.001	0.005	0.063	0.127
	(0.020)	(0.021)	(0.222)	(0.223)
Observations	15135	15135	13080	13080
R^2	0.043	0.042	0.191	0.192

Notes: OLS estimation. State and season fixed effects included, but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.13: Determinants of male wages (log)

	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	0.399		0.104	
	(0.542)		(0.591)	
SD of predicted log consumption		0.139		-0.378
		(0.744)		(0.823)
Education: primary	0.112***	0.111***	0.098***	0.098***
	(0.018)	(0.018)	(0.027)	(0.027)
Education: secondary	0.176***	0.175***	0.155***	0.146***
	(0.021)	(0.021)	(0.031)	(0.031)
Education: tertiary and higher	0.590***	0.589***	0.556***	0.536***
	(0.042)	(0.042)	(0.058)	(0.057)
Education: primary × IQR of pred. log consumption			0.294	
			(0.539)	
Education: secondary × IQR of pred. log consumption			0.498	
			(0.625)	
Education: tertiary and higher × IQR of pred. log consumption			0.789	
			(1.273)	
Education: primary × SD of pred. log consumption				0.346
				(0.697)
Education: secondary × SD of pred. log consumption				0.909
				(0.825)
Education: tertiary and higher × SD of pred. log consumption				1.667
				(1.675)
Observations	14868	14868	14868	14868
R^2	0.370	0.370	0.370	0.370

Notes: OLS estimation. Controls: Age, age (squared), caste, religion, number of children in the household, area cultivated, non-labor income, precipitation, state fixed effects and season dummies. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.14: Determinants of female wages (log)

	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	1.166		0.808	
	(0.859)		(0.799)	
SD of predicted log consumption		-0.021		-0.409
		(1.092)		(1.105)
Education: primary	0.070*	0.064*	0.031	0.029
	(0.032)	(0.032)	(0.066)	(0.065)
Education: secondary	0.133**	0.128**	0.016	0.003
	(0.049)	(0.049)	(0.087)	(0.100)
Education: tertiary and higher	0.626***	0.621***	0.701***	0.728***
	(0.093)	(0.094)	(0.167)	(0.177)
Education: primary × IQR of pred. log consumption			0.863	
			(1.194)	
Education: secondary × IQR of pred. log consumption			3.085	
			(1.958)	
Education: tertiary and higher × IQR of pred. log consumption			-2.093	
			(2.974)	
Education: primary × SD of pred. log consumption				0.950
				(1.485)
Education: secondary × SD of pred. log consumption				4.273
				(3.157)
Education: tertiary and higher × SD of pred. log consumption				-4.067
				(4.278)
Observations	4828	4828	4828	4828
R^2	0.326	0.324	0.330	0.328

Notes: OLS estimation. Controls: Age, age (squared), caste, religion, number of children in the household, area cultivated, non-labor income, precipitation, state fixed effects and season dummies. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.15: Robustness of results to supply side constraints

	(1)	(2)	(3)	(4)
Interquartile range of predicted log consumption	-1.525***	-1.549***	-2.052***	-1.596***
	(0.400)	(0.411)	(0.475)	(0.418)
Mother's predicted wage	0.097***	0.102***	0.094**	0.097***
	(0.026)	(0.028)	(0.031)	(0.028)
Age	-0.017***	-0.017***	-0.015***	-0.017***
	(0.003)	(0.003)	(0.003)	(0.002)
Household size	0.008+	0.008+	0.010*	0.007
	(0.004)	(0.004)	(0.005)	(0.004)
No of children in household	-0.008	-0.008	-0.009	-0.006
	(0.009)	(0.010)	(0.010)	(0.009)
Caste: SC / ST	-0.051**	-0.052*	-0.057*	-0.054**
	(0.019)	(0.020)	(0.022)	(0.020)
Religion: Hindu	0.091**	0.094**	0.098**	0.094**
	(0.031)	(0.030)	(0.032)	(0.030)
Area cultivated per capita (acres, log)	0.039	0.042+	0.052*	0.045+
	(0.024)	(0.023)	(0.026)	(0.023)
Annual p.c. income: other hh members' labor (log)	-0.003	-0.004	-0.005	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)
Annual p.c. income: non-labor (log)	0.015***	0.015**	0.015**	0.015**
	(0.004)	(0.004)	(0.005)	(0.005)
No. of primary schools in village	-0.009			
	(0.009)			
No. of secondary schools in village		0.003		
		(0.013)		
Average teacher absenteeism			-0.006	
			(0.004)	
Share of schools with Mid-day Meal Scheme				0.123*
				(0.051)
Observations	12750	12750	10803	12636
R^2	0.073	0.072	0.080	0.076

Notes: OLS estimation. State and season fixed effects included, but not reported. Standard errors (clustered at the village level) in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.16: Summary statistics (SEPRI)

	Mean	SD
Consumption expenditure per capita (log)	9.74	(0.61)
Total annual rainfall (log)	6.74	(0.46)
Share of irrigated land, village average	0.77	(0.30)
Employment generated per capita in 2011-12, NREGA	0.76	(1.23)
Average of annual rainfall (1960 - 2010), log	6.72	(0.43)
SD of predicted log consumption	0.059	(0.029)
SD of predicted log consumption (Av. NREGS employment = 0)	0.067	(0.024)
SD of predicted log consumption (Av. NREGS employment = 2.24)	0.038	(0.014)
Observations	46766	

Notes: All values in current INR.