Remember When It Rained: Gender Discrimination in Elementary School Enrollment in India

Laura Zimmermann*
University of Michigan

March 2011

Abstract

Girls in India have significantly lower school enrollment rates than boys. Anecdotal evidence suggests that intra-household gender discrimination is the most important reason for this gap, but empirical support in most of the previous literature is rare and seemingly inconsistent with patterns in related economic research. I propose that this may be due to a combination of endogeneity issues and inadequate attention to age-specific forms of discrimination. I analyze school enrollment using a plausibly exogenous source of income variation for rural households: rainfall shocks. The results show that girls' school enrollment is more vulnerable to rainfall shocks than that of boys, with 6-10 year old children driving these effects. This empirical pattern is consistent with young girls being out of school because of credit constraints, and older girls being out of school because of low perceived net benefits of education.

*email:lvzimmer@umich.edu. I thank Dean Yang, David Lam, Raj Arunachalam, Ashwini Deshpande, Suresh Tendulkar, Miguel Urquiola and participants of the University of Michigan Informal Development Seminar for valuable comments and feedback. Sulabha Parasuraman provided important information on district analysis in the NFHS datasets.

1 Introduction

In a number of developing countries, girls have worse education outcomes than boys. In India, for example, girls are significantly less likely to be enrolled in school and show worse school attainment, suggesting that they are disadvantaged in access to education. Gender differences persist in most Indian states, at basically all ages and along a number of other dimensions like caste and religion. This is especially true for rural areas and at older ages (see e.g. Jayachandran 2002, Kingdon 2007, Nambissan 1996). The gender gap has decreased, however, in the light of major policy programs by the Indian government, including free meal programs, the construction of new schools and the distribution of free textbooks to girls (Government of India 2006).

Largely descriptive work on education outcomes in India suggests a role for a broad variety of explanations, including school-based discrimination by teachers, sample selection of households with girls and boys, and intra-household gender discrimination[1].

Intra-household gender discrimination is usually identified as the most important channel and defined as parents systematically allocating resources differently between sons and daughters. A number of predominantly descriptive papers find that households allocate limited resources to the education of boys rather than girls across areas and income levels, or that school expenditures on girls are lower than those for boys, especially in private schools (see e.g. Filmer and Pritchett 1999, Jejeebhoy 1993, Ramachandran 2002, Tilak 1996, Tilak 2002) [2].

The literature usually focuses on two main reasons why parents may practice intra-household discrimination against girls in education decisions. First, the opportunity costs of attending school differ by gender and may be higher for girls than for boys. Girls in India are found to start helping in the household from early childhood on, for example by fetching water, cooking, cleaning or looking after their younger siblings. They spend two to three times the amount of time on domestic duties as boys. Boys,

[1]There is suggestive evidence, for example, that teachers systematically treat girls differently in the classroom and expect them to adhere to traditional gender stereotypes by requiring them to be quiet and non-participating (see e.g. Chanana 1990, Jethwani-Keyser 2008, Nambissan 1995). Research also suggests that sample selection in households with girls and boys is a valid concern. In a country like India, where son preference is pronounced, the gender of a child is no longer random and sex ratios are becoming more and more skewed towards males (Das Gupta, Chung and Shuzhuo 2009, Agnihotri 2003, Srinivasan and Bedi 2009). In such an environment, households that end up having girls may therefore have very different characteristics than households that have boys. Jensen (2002), for example, argues that girls may live in larger and poorer households than boys on average because parents continue to have children until the desired number of sons is reached. This could significantly impact enrollment for girls if parents are unable to afford the monetary expenses of school education. Some literature documents that poor children, and especially girls, are significantly disadvantaged in their access to education (Bhatty 1998, Filmer and Pritchett 1999, Jha and Jhingran 2002, Sipahimalani 1996). Jejeebhoy (1993), however, reports that in rural Maharashtra girls in small families are not more likely to be enrolled in school than girls in large families.

[2]Other literature on the topic of gender discrimination in areas such as food, health expenditures, time allocation or mortality includes Asfaw, Klasen and Lamanna (2007), Klasen (1996), Messer (1997), Miller (1997), Sauerborn, Berman and Nougtara (1996).

on the other hand, are more likely to help on the farm or work for pay, which tends to be more seasonal than household chores and to only become productive at older ages (Burra 2001, Caldwell et al. 1985, Jejeebhoy 1993). This suggests that sending a girl to school may be perceived as more costly than the school attendance of a boy, at least at younger ages.

Second, lower returns to education for girls, broadly defined, also seem to be an important reason for intra-household discrimination in education. Male labor force participation in India is significantly higher than that of women, so having an educated son may lead to large wage increases and better old-age support for the parents. Research also suggests that the income returns from being literate are substantial and that an educated son increases the family's social status (Bhatty 1998, Caldwell et al. 1985, Rosenzweig and Schultz 1982). Especially in North India, on the other hand, girls are traditionally seen as ceasing to be part of the family when they marry, so that benefits of education accrue to the husband's family instead. If they are to marry into a family of higher social status, educating a daughter may actually limit her chances in the marriage market or make the payment of a higher dowry necessary, unless potential grooms are educated and prefer a literate wife (Caldwell et al. 1985, Chanana 1990, Dreze and Saran 1995).

Overall, intra-household discrimination in the allocation of resources is therefore identified as a major reason for poor education outcomes of girls in most anecdotal and descriptive evidence. Research in economics, however, has only had limited success in verifying this importance: Empirical analyses usually do not find strong evidence of discrimination within the household. In the rare cases where they do, results often show lower education expenditures for girls only in the 10-16 age range (see e.g. Subramanian and Deaton 1990, Lancaster et al. 2003).

In contrast, a growing literature on the impact of exogenous shocks like rainfall or price shocks documents the existence of gender differences in a number of outcomes. While only few papers focus directly on the impact of these shocks on children's education, research suggests important gender-differential impacts especially early in life. The usual explanation given for this pattern is that resources are allocated differently to sons and daughters (see e.g. Maccini and Yang 2009, Qian 2008).

Taken together, the empirical results therefore present a puzzle: The literature that directly tries to identify whether intra-household allocation matters rarely finds results and if so, finds that results are concentrated among older children. The literature that looks at the reduced form impact of shocks, on the other hand, identifies young girls as being most vulnerable, and often attributes this to gender discrimination. This raises the question whether these differences can be explained by methodological or data problems or whether both patterns somehow reflect reality.

In this paper, I propose that two factors may hamper a good understanding of gender-differential outcomes. Combined, they may contribute to reconciling the different empirical patterns in the literature. First, as has been suggested in previous research, the intra-household gender discrimination literature may suffer from endogeneity and aggregation problems, given that explanatory variables like household income or composition widely used in that literature are likely to be correlated with the error term. An empirical identification strategy that focuses on exogenous income shocks may therefore give us more credible results.

Second, I suggest that interpreting empirical results may be complicated by different types of gender discrimination that are concentrated at different age groups. In the context of school enrollment, for example, a child that is not enrolled in school may be out of school because of credit constraints or because the net benefits of school enrollment are lower than those of the best alternative option. I call a child an 'involuntary 0' in the first case and a 'true 0' in the second case. This distinction matters since only 'involuntary 0s' can be affected by an income shock: Higher household income will alleviate credit constraints and will therefore lead to more 'involuntary 0s' being enrolled in school. But pure income shocks will in general not change net benefits or opportunity costs, so an income shock will not affect the school enrollment of 'true 0s'.

There is reason to believe that younger girls will tend to be 'involuntary 0s', whereas older girls will often be 'true 0s'. Finding that there are gender differences in the impact of income shocks at young ages therefore does not necessarily imply that there is no gender discrimination at older ages if fewer older girls can ever be affected by income shocks. This may lead to a pattern where gender differences in outcomes are biggest for older children, but differences in the impact of shocks are most pronounced for younger children. This would explain the patterns found in the different strands of literature: The gap in education expenditures may be most pronounced at older ages because few girls are still attending school, so an additional girl of that age group will raise household education expenditures by significantly less than another boy. This is consistent with the few results in the gender discrimination literature. At the same time, income shocks may only matter for young girls since most older girls will not be affected by income changes, leading to a pattern consistent with the shock literature.

The empirical analysis in this paper focuses on the impact of rainfall shocks on school enrollment for 6-14 year old children in India. In rural areas, deviations of rainfall from long-term averages will lead to income shocks for households as agricultural output is influenced by the weather. Rainfall shocks therefore give us a proxy for exogenous income shocks [3], and allow me to analyze intra-household allocation of resources in reaction to unanticipated changes in income.

I find that girls are more vulnerable to rainfall shocks than boys with respect to school enrollment. The results are driven by 8-10 year olds, whereas there are no significant differences for older children. Higher than average rainfall leads to an increase in school enrollment for both boys and girls, but the effect is monotonically decreasing in age. There is little evidence that this is driven by extreme weather phenomena like floods or droughts, which may have effects that go beyond mere household income shocks. The findings are robust to a number of alternative specifications. Overall, my analysis supports the idea that intra-household gender discrimination is a driving

[3]Rainfall shocks may also have additional impacts on schooling that are not directly related to household income: Heavy rainfalls may, for example, make schools inaccessible or change opportunity costs. I pay attention to these other potentially confounding influences in my empirical analysis.

factor in school enrollment differences and that discrimination manifests itself differently empirically for children of different ages.

The rest of this article is structured as follows: Section 2 provides background information on education in India and presents an overview of the related literature. Section 3 sets up a simple conceptual framework for analyzing school enrollment. Section 4 discusses rainfall in India, introduces the data and discusses the empirical strategy. Section 5 presents the main results as well as a number of robustness checks. Section 6 concludes.

2 Education in India and Related Literature

2.1 Education in India

Compulsory schooling in India applies to all children who are between 6 and 14 years old, and spans their elementary education years in school, which consist of primary and upper primary school [4]. This requirement often is not binding, however. Most Indian states have not nearly achieved universal school enrollment. Table 1 demonstrates that school enrollment in rural areas in India is significantly lower for girls than for boys at all ages and that the disparity is increasing in age. It also shows, however, that while enrollment rates for both boys and girls have increased substantially over time, they have improved faster for girls, leading to a narrowing in the gender gap over time. Similar patterns hold at the sub-national level, even though levels and gender gaps differ markedly between states: In the 1991 Indian census, for example, primary school enrollment for Indian children was 35 percent in Bihar but 85 percent in Kerala (Jayachandran 2002, Kingdon 2007). Significant gender differences also hold up for Hindus and Muslims as well as for Scheduled Castes (SC) and Scheduled Tribes (ST) [5]. Ramachandran (2002) estimates that about 70 percent of out-of-school children are girls.

Table 1 also reveals an inverse U-shaped pattern of school enrollment with age. While many children are enrolled in school when they are six years old, enrollment rises until about age nine, and then starts to decline as children drop out of school. This pattern is consistent with school enrollment in other developing countries, for example in a number of Latin American countries, although children in India tend to start dropping out much earlier (Urquiola and Calderon 2006).

Lower school enrollment for girls than for boys is driven by two phenomena: Girls

[4]Elementary education is roughly equivalent to primary and middle school. In general, primary schooling includes classes I through V, and upper primary school classes VI through VIII, although there is state variation in the exact classes which fall under primary and upper primary education (Tilak 2002).

[5]Scheduled Castes (SCs) is the official name for untouchables or dalits. Scheduled Tribes (STs), also often called adivasis, are not part of the usual caste hierarchy and often live in secluded hilly or forest areas. Both SCs and STs are traditionally disadvantaged groups within Indian society and still tend to have much poorer outcomes than members of other castes.

are more likely to never attend school at all, and are also more likely to drop out at any age than boys, especially when moving to upper primary and secondary levels of schooling. While the number of children who have never been enrolled in school has decreased markedly over the years, dropouts still remain high. Many children drop out as early as grades I and II, especially among low-caste children (Caldwell et al. 1985, Nambissan 1996, Ramachandran 2002, Seetharamu and Devi 2007). This pattern suggests that both entry and exit are important margins at which school enrollment for girls and boys differ [6].

Since the second half of the 1980s, the Indian government has shown efforts to actively increase enrollment rates for children, and girls specifically, through a variety of programs. The national government's New National Policy of Education (1986) and the Programme of Action (1992) were meant to target learning achievements, access and dropouts of 6-14 year olds. A number of specific interventions were made, like distributing free textbooks and uniforms, giving scholarships to SC and ST students, or providing mid-day meal schemes in various Indian states. The meal programs were especially successful in raising enrollment rates, particularly among girls [7].

Public elementary school education in India is supposed to be free. Some evidence suggests, however, that this may not always be the case. Looking at data from the second half of the 1980s which distinguishes between private and public school enrollment, Tilak (1996) finds that less than half of the students in rural India receive free primary school education in government schools. Expenditures can include tuition or examination fees or a variety of other fees. Newer evidence suggests that school expenditures for a child, including tuition fees as well as expenditures on uniforms or stationary, are around 340 rupees in rural India per year; this implies that for an agricultural laborer living in Bihar with three children, his complete earnings from more than 40 days of work would go towards sending his children to primary school (Dreze 2003, Kingdon 2005).

Another problem that has garnered an increasing amount of attention in recent years revolves around the issue of school quality. Especially public schools often do not manage to teach children basic reading and writing skills and are marked by very high levels of teacher absenteeism (see e.g. Duflo and Hanna 2006, Kremer et al. 2005). Many children still cannot properly read and write by grade III. Children primarily advance by practices of non-detention and often do not complete the grade for which they are enrolled. Schools also tend to have large class sizes and poor facilities. In reaction to this perceived low quality of government schools, a variety of private schools

[6] It also suggests that while there is a role for sheepskin effects occurring when children for example complete primary school, dropout decisions especially at young ages are not driven by such considerations.

[7] Problems with a number of these programs persist, however, with textbooks and uniforms for example not reaching beneficiaries, or not on time. There is also large inter-state variation in the scale of the programs: In a number of states, beneficiaries amount to less than 7 percent of students, while in states like Tamil Nadu or Karnataka about 70 percent of children receive free textbooks (Bhatty 1998, Ramachandran et al. 2003, Sinha 2003, Tilak 2002).

and government-aided schools have opened, and there is some anecdotal evidence that even relatively poor families prefer sending their children to these schools (Dreze and Kingdon 2001, Kingdon 1996, Ramachandran et al. 2003, Sinha 2003).

Overall, it is therefore likely that households that send their children to school incur education expenditures and that at least some households are credit constrained because they find themselves unable to borrow against future income to enroll their children in school. We would therefore expect that a positive income shock raises school enrollment.

2.2 Related Literature

Previous literature on gender discrimination in economics has often found it difficult to document systematic gender differences in the treatment of children within a household, both generally and in the Indian context. Some recent papers on India find evidence of differential treatment of girls and boys: Rose (2000) and Barcellos et al. (2010), for example, demonstrate differences in time allocation of mothers in Indian households with and without sons. Jayachandran and Kuziemko (2009) identify gender differences in the duration of breastfeeding of young children.

In general, however, literature on gender bias in intra-household allocation often does not find evidence of differential treatment of children. Research predominantly from the early 1990s usually estimates an extended Engel curve where the budget share of a good is regressed on per-capita household expenditures and a variety of household characteristics and composition variables. Researchers then test whether the estimated coefficient for a female household member of a given age range is significantly different from that for a male household member of the same age. If that is the case, this is taken as evidence of intra-household gender discrimination. Unfortunately, strong results have been rare. The typical analysis is not able to detect strong gender bias even in countries where other indicators show severe discrimination against girls (see e.g. Ahmad and Murdoch 1993, Deaton 1989, Fuwa et al. 2006, Gibson 1997, Gibson and Rozelle 2004, Haddad and Reardon 1993, Himaz 2008, Lee 2008, Liu and Hsu 2004, Murdoch and Stern 1997, Subramaniam 1996). Individual level gender bias seems to somehow disappear at the household level.

The case of education expenditures is not very different from the general picture: Subramanian and Deaton (1990) only find weak evidence of gender discrimination in rural Maharashtra for 10-14 year olds, but not for other age groups, while Lancaster, Maitra and Ray (2003) detect significant gender bias in their sample of rural Bihar and rural Maharashtra only for the age group of 10-16 year olds. Kingdon (2005) uses Indian survey data on 16 states but finds that the Engel curve approach often does not pick up existing individual level bias. Zimmermann (2010) finds stronger evidence of gender discrimination than most other papers, but the results are again concentrated among 10-19 year olds.

Some recent papers have suggested various possible explanations for this pattern (see e.g. Aslam and Kingdon 2008, Bhalotra and Attfield 1998, Gibson and Rozelle 2004, Gong et al. 2005, Kingdon 2005). Especially threats to identification and aggregation issues could potentially explain the widespread failure to find results. Data is often only available at the household level rather than for each individual child, which in general requires goods that are only consumed by a certain group of household members and the assumption that an additional child in the household acts as a negative income shock but induces no substitution effects. Household size and composition especially in countries with son preference are not exogenous, however, although it is not clear whether this will bias coefficients of interest upwards or downwards [8].

In contrast to this, a growing literature in economics has been able to demonstrate important differential impacts of exogenous shocks by gender. One prevalent type of such shocks are rainfall shocks. They are taken as proxies for income shocks in agricultural areas, although research often looks at the reduced form impact, which may also include other factors like general equilibrium or spillover effects. The advantage of rainfall shocks is that they are defined as deviations of rainfall from long-term averages, which are arguably exogenous. The analysis of the impact of rainfall shocks on various outcome variables is therefore better identified than the older literature on gender discrimination, which could be a reason for finding gender-differential effects. Rose (1999), for example, finds that positive deviations of rainfall from average rainfall significantly increase the survival chances of girls in India. Maccini and Yang (2009) find that higher rainfall in a person's birth year in Indonesia has positive effects on women's long-run socioeconomic outcomes, whereas there is no effect for men. Their results are consistent with intra-household gender discrimination in the allocation of nutrition and other resources.

The pattern is similar for other types of shocks. Qian (2008) finds that in China sex-specific agricultural income shocks occuring to women raise girls' survival chances whereas a positive male-specific income shock worsens them. Jayachandran (2006) documents that negative shocks to air quality in Indonesia have bigger effects on female than on male infant mortality. Duflo (2003) finds that a grandmother's pension eligibility in South Africa has an impact on the anthropometric status of young girls, but not that of boys.

To the best of my knowledge, however, so far no one has specifically looked at the effect of rainfall on the school enrollment of children. In my analysis, I combine the advantages in identification strategy provided by using exogenous shocks with a stronger focus on analyzing whether intra-household gender discrimination can be seen as an important channel through which rainfall shocks lead to gender-differential outcomes.

[8] If son preference is positively correlated with education spending, the omitted variable bias will tend to overstate pro-male bias. If families continue to have children until the desired number of sons is achieved, however, the coefficients may actually understate pro-male bias as long as son preference and education spending are still positively correlated and son preference is also highly correlated with the share of girls in the household. Therefore, the bias in standard regressions is ambiguous even assuming that all households prefer sons to daughters (Zimmermann 2010).

3 Conceptual Framework and Predictions

The empirical patterns in the existing literature suggest a puzzle where two strands of literature are concerned with the analysis of the same channel of household behavior but where one strand of literature finds gender differences among young children whereas another one finds no results or results concentrated among older children. In addition to concerns about identification, one reason for this may be that empirical results need to be interpreted carefully. Some simple considerations about school enrollment demonstrate the importance of this.

One feature of the decision about the school enrollment of a child is that it is an either-or decision: A child is either enrolled in school, or he/she is not. In order for school enrollment to be the optimal decision for a household, the net benefit of school enrollment of the child needs to exceed the level of utility of any alternative. These alternatives are mainly doing household chores, working on the farm or working for pay. A household will therefore enroll a child in school if the following inequality holds:

$$B_S(a, g, e, p(a, g), y_H) \text{ - } C(a, g, e, p(a, g), y_H) \geq$$
$$\max[B_H(a, g, e, p(a, g), y_H), B_F(a, g, e, p(a, g), y_H), B_W(a, g, e, p(a, g), y_H)]$$

where $B_S(.)$ is the benefit derived from sending the child to school, C(.) is the monetary cost required for the child's school enrollment, $B_H(.)$ is the benefit of letting the child do household chores, $B_F(.)$ the benefit of having the child work on the farm, and $B_W(.)$ the benefit of paid work of the child. The benefit functions here can be seen as indirect utility functions coming out of a household utility maximisation problem. All benefits and costs can, in this general form, depend on the child's age a, his or her gender g, the education level already achieved e, parental preferences p (which can again change with the child's age or sex), and household income y_H. Especially for young children, these costs and benefits are just reflecting parental costs and benefits since parents make education decisions for their children [9].

Households that are credit-constrained, however, may find themselves in a situation where the inequality above holds, but the child nevertheless does not attend school because parents cannot afford it. For a full characterization of the situation we therefore also need the following inequality:

$$C(a, g, e, p(a, g), y_H) \leq y(a, g, e, p(a, g), y_H)$$

Here, C(.) is the same monetary cost function from above, whereas y(.) is the amount of household income allocated for education expenditures for this particular

[9] There is some debate in the literature about who makes education decisions. While some research finds that parents make the decisions and do not consult their children, other literature has argued that education decisions cannot always be made against the child's will (see e.g. Dreze and Kingdon 2001). Especially when child and parents do not agree about the costs and benefits of different options, this introduces issues of household decision making behavior and bargaining power. I here abstract away from this issue by assuming that parents are the sole decision makers in education issues which, especially for young children, seems to be a good approximation.

child. Therefore, a child is only enrolled in school if the monetary costs of enrollment are lower than the education budget for the child in question [10].

Figure 1 illustrates the situation. It is an adaptation of the Roy model to the particular situation of school enrollment and is somewhat related to generalized Roy models (see e.g. Heckman and Vytlacil 2005, d'Haultfoeuille and Maurel 2009). The figure has the maximum (net) benefit of all alternative options to school enrollment on the horizontal axis, and the benefit of enrolling the child in school on the vertical axis. These signify the payoffs from the two options open to every child: Being enrolled in school or not. Any point in the diagram presents a payoff combination of the benefit a given child would receive when going to school and when engaging in the best possible alternative.

If there were no costs to school enrollment in inequality 1, then school enrollment would simply be determined by whichever benefit was higher: For any payoff pair to the southeast of the 45 degree line, the child would not attend school since the benefit from an alternative is higher than the benefit of going to school. The child would be enrolled in school for all points to the northwest. Since there are costs to school enrollment however, the net benefit of school enrollment (B_S-C) needs to exceed the benefit of the best alternative option. The cost function C is plotted as a dashed line [11]. It is therefore now optimal for a child to be enrolled in school for all payoff combinations that lie to the northwest of the cost curve, and to not be enrolled otherwise. This represents graphically the condition given by the first inequality above.

In order to be enrolled in school, however, the second inequality also needs to hold: Costs C need to be smaller than the education budget allocated to a particular child. Since this inequality is in monetary cost terms but the axes in the diagram are in benefit units, depicting this inequality explicitly is difficult. For clarity, in the diagram credit constraints are therefore incorporated in the benefit combinations for a child. A child living in a credit-constrained household is assumed to have a lower benefit from enrollment than an identical child without such constraints.

Figure 1 shows the effects from an income shock. The situation before the income shock is depicted by points A, B and C. Child A will be enrolled in school since the payoff combination is to the north-west of the cost line, whereas children B and C will not attend school. With a positive income shock, a household has now more resources to spend on a child's education, which will work to shift up the benefit combination for children from credit-constrained households like child B. In the case of a large enough income shock, child B's payoff combination is shifted up enough to cross the cost line

[10] This general version is agnostic about how a household ends up with an education budget for an individual child. Households could decide on the household education budget first, and then decide how to allocate the money among the children. Alternatively, households could have a budget for total expenditures for every individual child, and then decide how much of this money to spend on education rather than on other things like health care or food for that child.

[11] The cost function C can be envisioned to have many different functional forms. One possible version is plotted as the dashed line in Figure 1, which assumes that the monetary costs of school enrollment are constant across opportunity cost levels. Functional form considerations do not affect the general points made here.

and the child will therefore now attend school.

Points A, B and C are the three possible situations a child may be in. A child with combination A is what I call a 'true 1': Both inequalities are satisfied for this child and he/she is therefore enrolled in school, both before and after the positive income shock. Child B is an 'involuntary 0' as the income shock is big enough to push the child across the cost line. Before the shock, the child was not enrolled in school, but not because it was not optimal for the child to be enrolled, but because of credit constraints. The income shock alleviates these constraints and the child therefore enrolls in school. Lastly, child C is a 'true 0', where school enrollment is not optimal even with a higher household income.

These considerations need to be taken into account when thinking about the empirical predictions we expect from the impact of rainfall shocks on school enrollment. Unfortunately, the actual benefit that households derive from school enrollment and all possible alternatives is in general unobservable. Only the outcome, being enrolled in school or not, is observed in the data.

In the setup of the diagram, gender differences in school enrollment may arise out of different patterns of benefit combinations. Significantly more boys in India are enrolled in school at any given point in time, so a bigger percentage of boys than girls will be 'true 1s' and more girls will be out of school. Out-of-school children can either be 'true 0s' or 'involuntary 0s', and we do not observe which of the two categories a child belongs to. Previous research shows, however, that by now few children have never attended school which suggests that at least for low grades the net benefits of school enrollment exceed alternative options for most children (see e.g. Ramachandran 2002). Therefore, it is plausible to assume that at young ages most out-of-school children are 'involuntary 0s'. Since more girls than boys are out of school, a positive income shock will therefore have a bigger impact on school enrollment for girls than boys. Intra-household gender discrimination here therefore demonstrates itself through a larger impact of the rainfall shock on female school enrollment, which is driven by the fact that most boys are already attending school so that a positive income shock can only affect girls.

At older ages, however, more and more girls and boys may turn into 'true 0s', for example because of low returns to education or high opportunity costs of children's time, for example in the form of farm or household work. In consequence, we should expect the magnitude of the increase in school enrollment to decline with age, since fewer children can now be affected by a given income shock. If this effect is more severe for girls than for boys, for example because returns are perceived to be lower for girls, gender differences in the impact of the shock will disappear. In an extreme case, boys' increase in school enrollment could then even be bigger than that of girls since so few girls can now ever be affected by a given income shock.

This suggests that regression results of the impact of rainfall shocks need to be interpreted carefully since it is not true that the absence of significant gender differences in the impact of the shocks implies that there is no intra-household gender discrimination. However, the above considerations lead us to expect that we should find bigger effects of rainfall shocks on school enrollment for younger children and for girls than boys, which should decrease in age. For older children, gender differences may disappear or even flip.

4 Weather in India, Data and Empirical Strategy

4.1 Weather in India and the Effect of Rainfall Shocks on Agricultural Output

Weather in India varies widely by region and often also at the sub-regional level. India has six different climatic regions, ranging from humid tropical areas to desert-like dry regions. While there are local differences, India in general experiences four seasons: winter (January to February), summer (March to May), the monsoon season (June to September) and the post-monsoon season (October to December) (De et al. 2005, Ribot et al. 1996). About 80 percent of annual rainfall occurs during the monsoon months. In general, rainfall during these months is smooth, with rainfall increasing in intensity until July, and then decreasing afterwards (Department of Agriculture and Cooperation 2007, Webster and Hoyos 2004).

As about 60 percent of the agricultural sector depends on rain as the only source of water, agricultural output should be heavily influenced by precipitation levels and timing (Directorate of Economics and Statistics India). Unfortunately, the dataset that I will be using for my analysis does not ask households about their crop yield or household income, so we cannot verify directly whether rainfall shocks do indeed feed through to changes in household income [12]. Figures 2 and 3 therefore show the relationship between standardized rainfall shocks (actual rainfall - mean rainfall divided by standard deviation) and analogously standardized district output of rice and wheat, two major food grains, with data from the Indian Department of Agriculture and Cooperation for Indian districts for the period 1999-2007 [13].

Positive numbers signify more rainfall or more agricultural output than usual, whereas negative numbers mean less rainfall or output than usual. A scatter plot point gives the rainfall shock and agricultural output shock in a given year for a certain district. The figures also show the non-parametrically estimated relationship between rainfall and output shocks. The estimation in Figure 2 suggests that the relationship between rainfall shocks and output of rice is close to concave[14]. While more rainfall

[12]Most collected information is related to assset ownership, which is a better proxy for permanent income rather than transitory income changes induced by rainfall shocks.

[13]In India, food grain production accounts for about two thirds of total agricultural production and cropped area. About 37 percent of food grain area was irrigated in the early 1990s, which the time of the first household survey that I am using (Directorate of Economics and Statistics India). This percentage is increasing: In 2007-08, about 47 percent of the food grain area was irrigated. To the extent that irrigation insulates a household from rainfall shocks, this will reduce the observed impact of rainfall shocks.

[14]This pattern differs somewhat from Levine and Yang (2006) who estimate an approximately linear relationship of rainfall shocks and rice output in Indonesia.

is always beneficial for rice output, the impact of rainfall shocks on output shocks is lower for positive shocks than for negative shocks. This suggests that in my analysis the impact of negative rainfall shocks should be especially pronounced.

Previous literature has usually ignored the potential issue that the impact of rainfall shocks on agricultural output may vary by crop. Papers such as Levine and Yang (2006) focus on the relationship between rainfall and rice to motivate the use of rainfall shocks as income shocks. Rice is a very special kind of crop, however, as rice fields are often flooded during the planting process. In dry parts of India without extensive irrigation systems, for example, rice is usually sown immediately after the onset of the monsoon (National Portal of India). Other crops like wheat, on the other hand, are much more vulnerable to excessive rainfall, so that non-linearities of the impact of rainfall on wheat output are an important potential concern (Sreenivasan and Banerjee 1973). Figure 3 presents the results for the relationship between rainfall and wheat output shocks and demonstrates that at least in India even relatively large positive rainfall shocks do not seem to have a negative impact on wheat output. There is a positive relationship between more rain and agricultural output for negative rainfall shocks, but the non-parametrically estimated curve quickly flatens out for positive rainfall shocks[15].

Overall, the figures suggest that rainfall shocks affect household income through agricultural yield and that therefore one channel through which rainfall affects school enrollment is through the change in household income. The impact of rainfall shocks on agricultural output is always positive but may be non-linear, especially for negative rainfall shocks.

4.2 Data and Variable Construction

The primary datasets used in this paper are the National Fertility and Health Surveys (NFHS) of 1992/93 and 1998/99. The NFHS is a cross-sectional dataset that is nationally representative of India's population. It includes current school enrollment for children and other family background characteristics as well as information about the district in which households currently live, which will be used to match household survey information to rainfall shocks[16]. Unfortunately, a district variable is missing for the most recent NFHS of 2005/06, so this data cannot be exploited in this article [17].

[15]The overall less steep relationship between rainfall and wheat is probably heavily influenced by the prevalence of irrigation. Wheat tends to be much more heavily grown in highly irrigated areas than rice, which insulates output from rainfall shocks to some degree (National Portal of India).

[16]India's state and district boundaries change dramatically over time. Between 1971 and 2001, for example, the number of districts rose from 356 to 593. New districts are not only created by partitioning existing districts, but may also arise from more complicated rearrangements (Kumar and Somanathan 2009). This often makes district comparisons across time difficult. Both rounds of NFHS surveys that I use base the district boundaries on the boundaries of the Indian Census of 1991, however, and are thus comparable between surveys.

[17]More detailed geographical information is excluded in the NFHS of 2005/06 to ensure confidentiality since HIV test results were collected simultaneously. See NFHS Supplemental Documentation (2010) for more details

Rainfall data comes from the Center of Climatic Research at the University of Delaware, which includes monthly precipitation values on 0.5 intervals in longitude and latitude. This grid is achieved by using data from various weather stations and a number of different sources of rainfall data to interpolate rainfall in these intervals[18]. In order to get rainfall at the district level, I match the closest grid point to the latitude and longitude of the district capital. Rainfall shocks are constructed by subtracting the long-term average rainfall from rainfall in a given period and then dividing this term by the standard deviation of rainfall in that district. A positive number therefore signifies more than average rainfall, whereas a negative number implies less rainfall than usual. The long-term average includes rainfall from 1970 to 2008, but does not include the current period. Such a construction of the rainfall shock has the advantage of being readily interpretable in terms of standard deviations and ensures that geographic differences in rainfall are taken into account: 200mm of additional rainfall, for example, will be a major positive deviation in very dry parts of the country like Rajasthan, but may not be a large deviation from the average in wetter regions like Assam.

In the main specification of my paper, a rainfall shock is defined as the deviation of rainfall in the 12 months before the interview month and year from typical rainfall in this 12 month interval. For a household interviewed in August 1998, for example, rainfall in the previous 12 months includes all rainfall in the appropriate district from August 1997 to July 1998. The expected rainfall in this period is calculated as average rainfall in a August to July interval between August 1970 and July 2008, excluding August 1997 to July 1998. Analogously, the standard deviation of rainfall in such a period is constructed. The rainfall shock for a household interviewed in August 1998 is therefore the deviation of rainfall in the previous 12 months from expected rainfall in such a period, divided by the standard deviation. In consequence, a rainfall shock is district-, month- and year-specific, where the important date is the interview date [19]

The moving time window, which depends on the household's interview month and year, ensures that rainfall shocks for all households really only refer to past rainfall and capture the same time span before the interview date, which would not work with the use of the same calendar months for all households in the sample. In both NFHS surveys, households were interviewed over the span of almost two whole years, depending on residence within states and districts [20]. In many Indian states, households were interviewed over the span of a couple of months. This feature makes the use of rainfall shocks by calendar year a less sensible choice.

The caveat of this specification is that shocks in the previous 12 months are implicitly

[18]For more details, see the University of Delaware Center of Climatic Research website at $http$: $//climate.geog.udel.edu/\ climate/html_pages/Global2_Ts_2009/README.global_{pt}s_2009.html$.

[19]Lags and leads of these rainfall shocks are also constructed, and the construction of these shocks will be detailed further in the robustness check section.

[20]In the 1992/93 survey only the Jammu districts of the state Jammu and Kashmir were interviewed. Households in the state of Tripura were not interviewed in 1992/93 but were interviewed in 2000 as part of the 1998/99 survey. There is no evidence, however, that in general the interview date is correlated with state characteristics.

taken to have the same impact on school enrollment regardless of the interview month. Yet, a household that reports the enrollment of children at the beginning of the school year may be different from a household where children attend school near the end of the school year: Resources could for example be relatively abundant when children start the school year, so that shocks in the previous 12 months have a limited impact on enrollment, but later on resources become scarcer and shocks may have a larger effect on education outcomes. I will take account of this problem by including month and year fixed effects into my empirical specifications, which control for seasonality as well as general time trends. I also include district fixed effects.

Time and district fixed effects take account of time-invariant regional effects like cultural norms, history or educational priorities and general time trends. Given that I only have two rounds of data, I am unable to look at changes over time. Therefore, the effects that I am identifying here will be those that remain important over the whole time span. As I only know school enrollment in general, but not the school attended, I am also unable to pick up school transfers in reaction to rainfall shocks, for example from government schools to private schools. My analysis thus focuses on the extensive margin of school enrollment.

4.3 Empirical Strategy and Summary Statistics

My main empirical specification is given by the regression equation

$$school_{ijkl} = \beta_0 + \beta_1 shock_{jkl} + \beta_2\, age_{ijkl} + \theta_k + \eta_l + \gamma_j + \epsilon_{ijkl}$$

where the subscripts refer to individual i in district j in month k and year l. *school* is an indicator variable equal to 1 if a child is currently reported to be enrolled in school and *shock* is the standardized rainfall shock variable. *age* is the child's age. In all empirical specifications, I include age dummies for every age between 6 and 14 years, with age 6 being the reference category. θ are month fixed effects, η are year fixed effects and γ are district fixed effects. Standard errors are clustered at the district level. The coefficient of interest is β_1. I run this specification separately for girls and boys but also test for gender differences by fully interacting the model with gender.

In alternative specifications, I interact the rainfall shock with age dummies to get at non-linear effects in age. I also include rainfall shock lags and look at non-linear effects of rainfall shocks on school enrollment by splitting up the rainfall shock variable into a number of indicator variables depending on the magnitude and sign of the rainfall shock. Lags are constructed by moving the time window backwards: The main rainfall shock variable takes into account rainfall deviations in the 12 months prior to the interview months. Lag 1 of this variable, for example, therefore constructs analogously standardized rainfall shocks by focusing on rainfall deviations 13-24 months prior to the interview date.

My identification strategy therefore uses the variation in rainfall shocks within districts across months and years to identify the effect on school enrollment. Summary statistics of the constructed rainfall shocks are presented in Table 2. If deviations of rainfall from long-term averages were perfectly random and we had an infinite sample size, we would expect the mean of the rainfall shock variables to be zero, and the distribution of the variables to be symmetric around the mean since the probability of a positive shock of a given magnitude occurring is the same as the probability of getting a negative rainfall shock of the same size. In reality, however, the mean rainfall shock will often be different from zero because of spatial correlation: If one district is experiencing an unusually dry summer, for example, the probability that the surrounding districts receive less rainfall than usual as well is higher. Table 2 shows that the mean of the rainfall shock variables tend to be relatively small. All means, however, are statistically significantly different from zero. Except for the two constructed leads of the rainfall shock variable, rainfall shocks are relatively symmetric around zero.

Table 2 also splits up the main rainfall shock variable into magnitude categories to give an overview of the distribution of the rainfall shock. As we can see, the distribution is skewed towards negative deviations of rainfall from long-term averages: About 27 percent of rainfall shocks fall into the category of 0.5 to 1.5 standard deviations less rainfall than normal, whereas only about 18 percent of rainfall shocks in my sample are 0.5 to 1.5 standard deviations higher rainfall than expected. About 40 percent of rainfall shocks are relatively small deviations from long-term averages within 0.5 positive or negative standard deviations. Extreme rainfall shocks that are larger than 2.5 standard deviations in absolute value are very rare.

Table 3 reports the break-down of children in my sample by age and gender. It also shows averages for household size, age of the household head, wealth index, and whether the household owns their house and any agricultural land. As we can see, the sample contains slightly more boys than girls, but the breakdown within gender by age is similar: 12 percent of girls are six years old, for example, whereas six-year old boys make up 13 percent of boys in my sample. Averages for the household variables are also similar between boys and girls, suggesting that the living conditions for girls and boys are close to identical. Both girls and boys live in households where the average age of the household head is 45 years. 95 percent of households own their house and two thirds own any agricultural land. The typical household size of a household with a boy is 7.6, which is only slightly smaller than the average household size for households with a girl, which is 7.8. Households with boys and girls also perform very similar for the wealth index. Overall, these summary statistics suggest that at least in my sample there is little room for sample selection concerns of households with boys and girls. Any results presented below are therefore not driven by differences in household characteristics.

5 Results

5.1 Main Results

The main results are presented in tables 4, 5, 6 and 7. Table 4 reports results with just rainfall shocks in the past 12 months as an explanatory variable, controlling for age and including month, year and district fixed effects, while Table 5 interacts the shock variable with age to look at non-linear effects of rainfall shocks with respect to age. Table 6 splits the rainfall shock variable up into indicator variables of different magnitude categories to analyze whether rainfall shocks of different sizes have non-linear effects on school enrollment. Table 7 presents the results of these magnitude categories interacted with age categories.

In Table 4, a rainfall shock of one standard deviation in the previous 12 months translates into an increase in the probability of being enrolled in school of 2.3 percentage points for girls and 1.3 percentage points for boys. As mean school enrollment in my sample is 0.66 for girls and 0.8 for boys, this translates into a 3.5 percent increase in school enrollment for girls and a 1.6 percent increase in male school enrollment. Both coefficients are highly statistically significant and the effect for girls is statistically significantly greater than the effect for boys at the 5 percent level. Results are similar for the subsample of children aged 6-10 years, where a shock of one standard deviation increases school enrollment by about 2.5 percentage points (3.7 percent) for girls and by 1.2 percentage points (1.5 percent) for boys. For 11-14 year olds, there is no longer a statistically different effect of rainfall shocks on school enrollment by gender. School enrollment for girls improves by 1.7 percentage points (2.9 percent) and that of boys increases by 1.3 percentage points (1.6 percent).

Table 5 looks at the non-linear effects of rainfall shocks with respect to age by interacting dummy variables for each age with the shock variable. The table shows that the effect of the rainfall shock is monotonically decreasing in age: 6 year old children experience the largest increases in school enrollment with close to five percentage points, and the effect decreases as children become older. The table also presents tests for the gender equality of the magnitude of the rainfall shock for every age. Significant gender differences occur for 8-10 year olds, where the rainfall shock has significantly larger effects on school enrollment for girls than for boys. The interaction effects for 8 and 10 year olds are also significantly different from each other for girls and boys, indicating that the impact of the rainfall shock declines more sharply for boys than girls at these ages. The rainfall shock remains highly statistically significant in this specification as well. This confirms that gender differences in the impact of rainfall shocks are driven by younger children.

Tables 6 and 7 focus on non-linearities in the magnitude of the rainfall shock by splitting up the rainfall shock variable into a number of dummy variables, depending on the magnitude of the shock. The reference category in these regressions are relatively small shocks of a magnitude between -0.5 and 0.5 standard deviations. As Table 6 shows, the impact of rainfall shocks on school enrollment is concentrated among negative shocks between -2.5 and -0.5 standard deviations: A rainfall shock between -2.5 and -1.5 standard deviations, for example, leads to a 6.2 percentage points decrease in school enrollment for 6-14 year old girls and to a decrease of 3.6 percentage points for boys of the same age. Except for 6-10 year old girls, where the category of shocks between 1.5 and 2.5 standard deviations is marginally significant, no positive rainfall shock category is statistically significantly different from zero. Significant gender differences between girls and boys are also concentrated on negative shocks and are again driven by 6-10 year old children, where girls experience a significantly greater decrease in school enrollment after a negative rainfall shock than boys. There is also little evidence that large absolute shocks of more than 2.5 standard deviations, which we might think of as floods or droughts, have particularly important effects on children's school enrollment. This may be partly due to the fact that, as Table 2 showed, such extreme rainfall shocks are very rare in my sample and their impact may therefore be imprecisely estimated. But these results are also consistent with the agricultural output graph from Figure 3 which showed that large absolute shocks may have no detrimental effects for some crops such as wheat so that the income shocks of large rainfall shocks may actually not be that different from income shocks caused by smaller rainfall deviations.

One big concern with using rainfall shocks is that rainfall shocks may not just act as income shocks, but may also change the opportunity costs for children. Children may for example need to help on the farm when the harvest is good. This would work in the opposite direction of the effect of alleviating credit constraints since more rainfall and thus a better harvest would now be associated with lower school enrollment. The same pattern would occur if roads are closed during heavy rainfall or schools inaccessible. Therefore, the positive relationship between rainfall shocks and school enrollment that we find even in this more detailed analysis here suggests that these factors cannot be dominant.

Table 7 interacts the magnitude categories of the rainfall shock variable with the age dummies. For an easier overview over the age and magnitude patterns, Table 7 reports the estimated overall effect for a given magnitude and a given age of a child, rather than the full regression table with main and interaction effects, which is given in appendix Table A1. Black borders show statistically significant gender differences at at least the 10 percent significance level. The table confirms the results from the previous tables by showing that most significant results as well as most significant gender differences are concentrated among 6-10 year olds. Consistent with Table 5, Table 7 demonstrates that rainfall shocks matter most for 6 and 7 year old children, but in a quite symmetric fashion: Negative rainfall shocks lead to significant decreases in school enrollment for both boys and girls, whereas positive rainfall shocks increase school enrollment, especially for girls who experience significantly bigger increases in school enrollment than boys. Significant gender differences are also present for 8-10 year olds, where girls are significantly more affected by rainfall shocks than boys. The effects are also more concentrated on the impacts of negative shocks than positive ones, suggesting that girls are more likely to be taken out of school after bad shocks rather than being enrolled after positive rainfall shocks at these ages. As in previous tables, children over 10 years old are relatively little affected by rainfall shocks.

Overall, the analysis shows that girls tend to be more affected by rainfall shocks than boys, with results being driven by the primary school aged children rather than children that are older. If we interpret rainfall shocks as income shocks for these rural households, then the results suggest that younger children are more vulnerable to income shocks than older children and that gender differences of income shocks are driven by children in the 8-10 year age group. Furthermore, younger children in general are more vulnerable to rainfall shocks than older children, especially at ages 6 and 7.

5.2 Robustness Checks

While the results in the previous section already suggest a consistent pattern where the impact of rainfall shocks is decreasing in age and shocks have larger effects on girls than boys in the 8-10 year range, a number of checks can be performed to ensure that they are robust to different empirical specifications.

One potentially important concern is the chosen interval of the shock variable. The constructed rainfall shock takes into account rainfall in the 12 months prior to the interview month. If there are significant lags in the translation of agricultural output into household income, however, it may be that what matters more for school enrollment is not rainfall in the previous 12 months, but rainfall that occurred prior to this. In order to check whether this is an important concern, Table 8 extends the analysis of Table 4 by also including two lags in addition to the main rainfall shock variable. Lag 1 is constructed by taking into account deviations of rainfall from long-term averages between 13 and 24 months prior to the interview months, and lag 3 analoguously focuses on rainfall in the interval of 25 to 36 months prior to the interview date. If these lags have important effects on school enrollment of children, we should see significant estimated coefficients in Table 8 or at least changes in the estimated coefficient of our main rainfall shock variable. Table 8 shows that neither is the case: The estimated coefficients for the lag variables are in general small and are never statistically significant; furthermore, the estimated coefficients for the main rainfall shock variable are very robust to the inclusion of lags and hardly change from the estimates reported in Table 4. This justifies just focusing on rainfall in the previous 12 months in this paper.

Another potentially important concern is measurement error. Matching rainfall data to the closest districts necessarily introduces measurement error since local conditions will be different from those in the district capital. In their paper, Maccini and Yang (2009) instrument for rainfall shocks from the closest source (in their case weather stations) with rainfall from the next nearest weather stations since rainfall should be correlated, but measurement error uncorrelated between rainfall data from different weather stations. I use a similar technique in Table 9 where I instrument for the closest grid point of rainfall of a district with the 5 next closest points. As the table shows, measurement error is not much of a concern in my case, since IV results are basically identical to the main results presented in Table 4 of this paper[21]. The first stage IV

[21] This is porbably more driven by the rainfall data that I use, which has already been smoothed by

results are documented in appendix table A2.

A third problem that might be driving the results is selection into my sample, for example through death or migration correlated with rainfall shocks. If selection is not an issue, then regressing cohort sizes of girls and boys on the rainfall shocks should not produce any effects: Whether rainfall in the previous 12 months was particularly high or low should not affect how many children of a particular age and gender interviewers were able to locate. The results of this analysis are shown in Table 10 and demonstrate that selection indeed is not an issue since none of the estimated coefficients is close to being statistically significant.

The hypothesis that the effects of rainfall on school enrollment work through the household income channel, rather than through any other variables correlated with rainfall, also lends itself to being tested empirically by focusing on a falsification test. If rainfall shocks lead to household income shocks through the channel of agricultural output and then influence school enrollment, we should find that the impact of rainfall shocks on school enrollment is significantly lower in urban areas, where agricultural production is less important for household income. Table 11 presents the results for urban areas, focusing on households living in large cities[22]. The results show that the impact of rainfall shocks on school enrollment is different in urban areas: A one standard deviation of additional rainfall here has no statistically significant effect on school enrollment for girls and older boys, and is even associated with a large negative effect of 5.5 percentage points for 6-10 year old boys. Table 11 also reports the p-value of testing the equality of coefficients in urban areas with the rural coefficients estimated in Table 4; the urban coefficients are significantly lower than those in rural areas for boys of all ages and for the overall sample of girls, although not for the subgroups of 6-10 and 11-14 year old girls. Overall, however, Table 11 supports the idea that rainfall shocks matter more in rural areas and are thus working through household income shocks.

5.3 Discussion

The previous analysis has demonstrated that rainfall shocks have a robust impact on school enrollment for 6-14 year old children in India. The effect is concentrated among young children, with the magnitude of the effect decreasing in age. 6 and 7 year old children experience large increases in school enrollment after positive rainfall shocks, and important decreases in the probability of attending school after a negative rainfall

using data availlable from various rainfall stations rather than due to measurement error not being an issue in my case.

[22]The NFHS datatsets distinguish between the countryside, towns, small cities and large cities. Only the countryside gets coded as a rural area, whereas the rest is coded as urban even though towns may lie in major agricultural areas and have close economic connections to the surrounding rural areas. Agricultural production may therefore still play an important role in towns and small cities, but should have much less of an impact on household income in large cities. Consistent with this idea, in results not shown the impact of rainfall on school enrollment is higher the more rural an area becomes: School enrollment for both boys and girls are lowest in large cities and then increase for small cities and towns.

shock. Girls in this age category benefit significantly more from a positive rainfall shock than boys, although there are no gender differences with respect to negative rainfall shocks. Rainfall shocks also affect girls significantly more than boys in the 8-10 year age range, where girls seem to be especially vulnerable to being taken out of school after adverse rainfall shocks. There are no significant gender differences in the impact of the rainfall shocks at older ages, and 11-14 year old children in general are affected relatively little by these shocks.

This pattern of results is consistent with the conceptual framework set up earlier, which suggested that young children are more likely to be out of school because of credit constraints, whereas at older ages more and more children are not attending school because the net benefits do not exceed the opportunity costs of school enrollment. This could be due to opportunity costs increasing with age, with older children being more productive in doing household chores, working on the farm or working for pay; or it could be due to declines in the returns to education, broadly defined to include marriage market concerns as well as expected lifetime earnings. Regardless of the main underlying reason, these considerations predict that young children should benefit more from a positive income shock than older children because their opportunity costs are likely to be low, and returns to education are likely to be perceived as high by parents. The empirical results support this hypothesis by showing decreasing impacts of rainfall shocks in age.

The idea that young children are mostly out of school because of credit constraints is also supported by the non-linearity results in Tables 6 and 7. As was shown, the school enrollment of 6-7 year old boys and girls responds to positive and negative income shocks in a symmetric fashion: Positive rainfall shocks improve school enrollment, whereas negative rainfall shocks lead to fewer children being enrolled in school, suggesting that credit constraints play an important role for school enrollment among these young children. Since children usually start school at these ages, the results can be interpreted as parents sending their children to school if they can afford to do so, but delaying their school entry if the household experiences negative income shocks. The fact that this pattern holds for both boys and girls at these ages demonstrates that parents are not opposed to the idea of educating their daughters in general, but may suffer from credit constraints. The results also suggest, however, that parents already prioritize male school enrollment at these ages: School enrollment for girls at these ages is already significantly lower than that for boys, but girls benefit significantly more from positive rainfall shocks than boys. Consistent with the conceptual framework set up earlier, this suggests that boys are sent to school first and additional resources can then only be allocated to girls, who are disproportionately out of school because of credit constraints. Intra-household gender discrimination for these young children therefore takes the form of parents allocating resources to their sons first; daughters are only sent to school if the family is not credit constrained.

This reasoning also helps explain the results for 8-10 year old children. As Table 1 shows, the gender gap in school enrollment increases with age, so a higher and higher percentage of out-of-school children will be girls at older ages. For 8-10 year olds, the empirical results are consistent with most boys who can ever be affected by income shocks already attending school, probably not least because their school enrollment was prioritized from early on, so that the only children that can be affected by further household income increases are girls. Therefore, the finding that girls benefit significantly more from a positive rainfall shock than boys is consistent with the idea that an important percentage of out-of-school girls are involuntary 0s who are not enrolled in school because of credit constraints. Girls in this age group are therefore more vulnerable to rainfall shocks as they are sent to school last after a good shock, and especially taken out of school first when a bad shock occurs.

For 11-14 year olds, on the other hand, the relatively small magnitudes of the impact of rainfall shocks on school enrollment suggest that for these children income shocks matter relatively little for their education, indicating that most of out-of-school children in this age category are true 0s rather than credit constrained. Despite a further widening in the gender gap in school enrollment at these ages there are no significant gender differences in the impact of rainfall shocks. This suggests that more 11-14 year old girls are true 0s: Boys are sent to school because doing so is seen as worthwhile (although the effect for boys also decreases with age), whereas more and more girls are consciously taken out of school and can therefore no longer be affected by an income shock.

6 Conclusion

Overall, the empirical results in this paper are consistent with parents systematically discriminating against girls in the household in education decisions at all ages: Among children in the 6-10 year age category, discrimination takes the form of parents sending their sons to school first and only enrolling their daughters in school if they can still afford to do so. So at these ages, girls tend to be out of school because of credit constraints. In the 11-14 year age group, on the other hand, discrimination against girls takes the form of girls being out of school because the net benefits from school enrollment are perceived to be low for them. Girls' school enrollment at this age is therefore hardly affected by household income shocks. These results also imply that even transitory income shocks like deviations in rainfall can have important long-run impacts on socio-economic outcomes especially on young children since they influence the acquisition of human capital: A child that experienced positive income shocks early in the school career may be able to accumulate more human capital than a similar child that did not benefit from higher than average rainfall.

My empirical analysis has potentially important policy implications, as policies may need to target girls and their families age-specifically in order to close the gender gap in school enrollment. While younger girls are most vulnerable to unanticipated transitory income shocks, the results suggest that school enrollment for girls in this age group may improve as household income rises permanently. Parents do seem to see the benefits of sending young girls to school for at least some time.

For older children in general, and girls especially, however, just alleviating credit constraints has only limited beneficial effects on children's school enrollment. In order to increase enrollment rates in this area, it seems like parents' perceptions about the benefits and costs of school enrollment need to be changed. This may be very challenging, especially when requiring changes in traditions or kinship practices.

While the validity of the conceptual framework proposed in this paper cannot be verified directly, alternative explanations would need to be able to explain the various empirical patterns presented in this paper in a compelling way. In any case, the fact that the empirical results are consistent with the conceptual framework in this paper suggests that empirical results, especially when they are supposed to shed light on intra-household discrimination, need to be interpreted carefully. There is also room for important future research in looking more closely at why parents discriminate against girls in education decisions, since both opportunity costs of school enrollment and returns to education, broadly defined, can potentially play an important role in understanding parents' behavior. Only if parents' motives are better understood may it be possible to target policies to successfully improve school enrollment for older children.

References

[1] Agnihotri, S. Survival of the Girl Child: Tunnelling Out of the Chkravyuha. *Economic and Political Weekly*, 38(41): 4351-60, 2003.

[2] Ahmad and Murdoch, Identifying Sex Bias in the Allocation of Household Resources: Evidence from Limited Household Surveys from Bangladesh, 1993.

[3] Asfaw, A., Klasen, S. and Lamanna, F. Intra-Household Disparities in Children's Medical Care before Death in India. IZA DP No. 2586, 2007.

[4] Aslam, M. and Kingdon, G.G. Gender and Household Education Expenditure in Pakistan. *Applied Economics*, 40(20):2573 - 2591, 2008.

[5] Barcellos, S.H., Carvalho, L. and A. Lleras-Muney. Child Gender and Parental Investments in India - Are Boys and Girls Treated Differently? RAND Working Paper, March 2010.

[6] Bardhan, P.K. On Life and Death Questions. *Economic and Political Weekly*, Special Number, 1293-1304, 1974.

[7] Bhalotra, S. and C. Attfield. Intrahousehold Resource Allocation in Rural Pakistan: A Semiparametric Analysis. *Journal of Applied Econometrics*, 13(5):463-480, 1998.

[8] Bhatty, K. Educational Deprivation in India, A Survey of Field Investigations. *Economic and Political Weekly*, 33(27): 1731-1740, 1998.

[9] Burra, N. Cultural Stereotypes and Household Behaviour: Girl Child Labour in India. *Economic and Political Weekly*, 36(5): 481-488, 2001.

[10] Caldwell, J.C., Reddy, P.H. and P. Caldwell. Educational Transition in Rural South India. *Population and Development Review*, 11(1): 29-51, 1985.

[11] Chanana, K. Structures and Ideologies: Socialisation and Education of the Girl Child in South Asia. *The Indian Journal of Social Science*, 3(1), 1990.

[12] d'Haultfoeuille, X. and A. Maurel. Inference on a Generalized Roy Model, with an Application to Schooling Decisions in France, IZA Working Paper No 4606, December 2009.

[13] Das Gupta, M., Zhenghua, J., Bobua, L., Zbenming, X., Chung, W. and B. Hwa-Ok. Why is Son Preference So Persistent in East and South Asia? World Bank Development Research Group, Policy Research Working Paper 2942, 2002.

[14] Das Gupta, M., Chung, W. and L. Shuzhuo. Is There an Incipient Turnaround in Asia's 'Missing Girls' Phenomenon?, The World Bank Development Research Group, Policy Research Working Paper 4846, 2009.

[15] De, U.S., Dube, R.K. and G.S. Prakasa Rao. Extreme Weather Events over India in the Last 100 Years. *Journal of Indian Geophysical Union*, 9(3):173-187, 2005.

[16] Deaton, A. Looking for Boy-Girl Discrimination in Household Expenditure Data. *The World Bank Economic Review*, 3(1):1-15, 1989.

[17] Department of Agriculture and Cooperation, Government of India. Annual Report 06-07, 2007.

[18] Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Government of India. Agricultural Statistics at a Glance. 2010.

[19] Dreze, J. Patterns of Literacy and Their Social Context in V. Das et al. (eds) The Oxford India Companion of Sociology and Social Anthropology. Vol II. Oxford University ress, New Delhi, 2003.

[20] Dreze, J. and M. Saran. Primary Education and Economic Development in China and India: Overview and Two Case Studies in Basu, K., Pattanaik, P., and K. Suzumura (eds), Choice, Welfare, and Development: Essays in Honour of Amartya Sen, Oxford, Clarendon Press, 1995.

[21] Dreze, J. And G. Kingdon. School Participation in Rural India. *Review of Development Economics*, 5(1), 2001.

[22] Duflo, E. Grandmothers and Granddaughters: Old Age Pension and Intrahousehold Allocation in South Africa. *World Bank Economic Review*, 17(1): 1-25, 2003.

[23] Duflo, E. and R. Hanna. Monitoring Works: Getting Teachers to Come to School. MIT and NYU Poverty Action Lab Working Paper, 2006.

[24] Filmer, D. and L. Pritchett. Educational Enrollment and Attainment in India - Household, Wealth, Gender, Village and State Effects. *Journal of Educational Planning and Administration*, 3(2), 1999.

[25] Fuwa, N., Ito, S., Kubo, K., Kurosaki, T., and Y. Sawada. Gender Discrimination, Intrahousehold Resource Allocation, and Importance of Spouses' Fathers: Evidence on Household Expenditure from Rural India. The Developing Economies, XLIV(4):398-439, 2006.

[26] Gibson, J. Testing for Boy-Girl Discrimination with Household Expenditure Data: Results for Papua New Guinea. *Applied Economics Letters*, 4(10):643-46, 1997.

[27] Gibson, J. and S. Rozelle. Is it Better to be a Boy? A Disaggregated Outlay Equivalent Analysis of Gender Bias in Papua New Guinea. *Journal of Development Studies*, 40(4):115-136, 2004.

[28] Gong, X., van Soest, A. and P. Zhang. Sexual Bias and Household Consumption: A Semiparametric Analysis of Engel Curves in Rural China. *Journal of Applied Econometrics*, 20: 509-527, 2005.

[29] Government of India, Department of Elementary Education and Literacy, Department of Secondary and Higher Education, Ministry of Human Resource Development. Annual Report 2005-06. 2006.

[30] Haddad, L. and T. Reardon. Gender Bias in the Allocation of Resources within Households in Burkina Faso: A Disaggregated Outlay Equivalent Analysis. *Journal of Development Studies*, 29(2):260-276, 1993.

[31] Heckman, J. and E. Vytlacil. Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica*, 73, 669-738, 2005.

[32] Himaz, R. Intrahousehold Allocation of Education Expenditure and Returns to Education: The Case of Sri Lanka. University of Oxford Department of Economics Discussion Paper Series No.393, 2008.

[33] Jayachandran, U. Socio-Economic Determinants of School Attendance in India. Delhi School of Economics Working Paper, No. 103, 2002.

[34] Jayachandran, S. Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy*, 114(3):538-575, 2006.

[35] Jayachandran, S. Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires. Mimeo, Stanford University, 2005.

[36] Jayachandran, S. and I. Kuziemko. Why Do Mothers Breastfeed Girls Less Than Boys? Evidence and Implications for Child Health in India, Working Paper, May 2009.

[37] Jejeebhoy, S.J. Family Size, Outcomes for Children, and Gender Disparities. *Economic and Political Weekly*, 28(35): 1811-1821, 1993.

[38] Jensen, R. Equal Treatment, Unequal Outcomes? Generating Gender Inequality through Fertility Behavior, unpublished manuscript, John F. Kennedy School of Government, Harvard University, 2002

[39] Jethwani-Keyser, M. "'When Teachers Treat Me Well I Think I Belong"': School Belonging and the Psychological and Academic Well being of Adolescent Girls in India, Ph.D. Dissertation, 2008.

[40] Jha, J. and D. Jhingran. Elementary Education for the Poorest and Other Deprived Groups. Centre for Policy Research. New Delhi, 2002.

[41] Kaur, R. Dispensable Daughters and Bachelor Sons: Sex Discrimination in North India. *Economic and Political Weekly*, 109-114, 2008.

[42] Kingdon, G.G. The Quality and Efficiency of Private and Public Education: A Case-Study of Urban India. *Oxford Bulletin of Economics and Statistics*, 58(1): 55-80, 1996.

[43] Kingdon, G.G. Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure. *Economic Development and Cultural Change*, 53:409 - 451, 2005.

[44] Kingdon, G. The Progress of School Education in India, Global Poverty Research Group Working Paper 071, March 2007

[45] Kishor, S. 'May God Give Sons to All': Gender and Child Mortality in India. *American Sociological Review*, 58(2): 247-265, 1993.

[46] Klasen, Stephan. Nutrition, Health and Mortality in Sub-Saharan Africa: Is There a Gender Bias? *Journal of Development Studies*, 32(6):913-932, 1996.

[47] Kochar, A. 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. 1999.

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

[49] Kumar, H. and R. Somanathan. Mapping Indian Districts Across Census Years, 1971-2001. Centre for Development Economics, Delhi School of Economics Working Paper Series, No. 176, April 2009.

[50] Lancaster, G., Maitra, P. and R. Ray. Endogenous Power, Household Expenditure Patterns, and New Tests of Gender Bias: Evidence from India. Technical Report Discussion Paper no. 20/03, Department of Economics, Monash University, 2003.

[51] Lee, Y. D. Do Families Spend More on Boys than on Girls? Empirical Evidence from Rural China. *China Economic Review*, 19:80-100, 2008.

[52] Levine, D.I. and D. Yang. A Note on the Impact of Local Rainfall on Rice Output in Indonesian Districts. Working Paper, January 2006.

[53] Liu, J. and C. Hsu. Economies of Scale, Gender Discrimination, and Cost of Children. *Applied Economics Letters*, 11(6):377-382, 2004.

[54] Maccini, S. and D. Yang. Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall. *American Economic Review*, 99(3):1006-1026, 2009.

[55] Messer, Ellen. Intra-Household Allocation of Food and Health Care: Current Findings and Understandings - Introduction. *Social Science and Medicine*, 44(11): 1675-1684, 1997.

[56] Miller, Barbara D. Social Class, Gender and Intra-Household Food Allocations to Children in South Asia. *Social Science and Medicine*, 44(11): 1685-1695, 1997.

[57] Murdoch, J. J. and H. S. Stern. Using mixture models to detect sex bias in health outcomes in Bangladesh. *Journal of Econometrics*, 77:259-276, 1997.

[58] Nambissan, G. The Social Context of Schooling Girl Children in India. textitPerspectives in Education, 11(3-4), 1995.

[59] Nambissan, G. Equity in Education - The Schooling of Dalit Children in India. *Economic and Political Weekly*, 1996.

[60] National Portal of India, $http://india.gov.in/citizen/agriculture/index.php$, accessed on 02/21/2011.

[61] NFHS Supplemental Documentation: Data Set Version 52, September 2010.

[62] Pande, R. and N. Astone. Explaining Son Preference in Rural India: The Independent Role of Structural Versus Individual Factors. *Population Research and Policy Review*, 26(1): 1-29, 2007.

[63] Qian, N. Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance. *Quarterly Journal of Economics*, 123(3): 1251-1285, 2008.

[64] Ramachandran, V. Gender and Social Equity. Hierarchies of Access. The European Commission, 2002.

[65] Ribot, J. C., Najam, A. and G. Watson. Climate Variation, Vulnerability and Sustainable Development in the Semi-Arid Tropics. In: Ribot, J. C., Magalhaes, A. R., Panagides, S. S. (Eds.), Climate Variability, Climate Change and Social Vulnerability in the Semi-Arid Tropics. Cambridge: Cambridge University Press, 1-10, 1996.

[66] Rose, E. Consumption Smoothing and Excess Female Mortality in Rural India, Review of Economics and Statistics, 81(1):41-49, 1999.

[67] Rose, E. Gender Bias, Credit Constraints and Time Allocation in Rural India. The Economic Journal, 110(465):738-758, 2000.

[68] Rosenzweig, M.R. and T.P. Schultz. Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India. *American Economic Review*. 72(4): 803-815, 1982.

[69] Sauerborn, R., Berman, P. and Nougtara, A.. Age bias, but no gender bias, in the intra-household resource allocation for health care in rural Burkina Faso. *Health Transition Review*, 6:131-145, 1996.

[70] Seetharamu, A.S. and U. Devi. Education in Rural Areas - Constraints and Prospects. A.P.H. Publishing Corporation, New Delhi, 2007.

[71] Sinha, A. Disaster Management - Lessons Drawn and Strategies for the Future. 2002.

[72] Sinha, A. Basic Education for All. The Elusive Quest for Social Justice in a Dynamic Democracy. The Case of India, 2003.

[73] Sipahimalani, V. Education in the Rural Indian Household: A Gender Based Perspective, Mimeo, NCAER, 1999.

[74] Sreenivasan, P.S. and J.R. Banerjee. The Influence of Rainfall on the Yield of Rain-Fed Rice at Karjat (Colaba District). *Agricultural Meteorology*, 11:285-292, 1973.

[75] Srinivasan, S. and A.S. Bedi. Tamil Nadu and the Diagonal Divide in Sex Ratios. *Economic and Political Weekly*, 56-63, 2009.

[76] Subramaniam, R. Gender-Bias in India: The Importance of Household Fixed-Effects. *Oxford Economic Papers*, 48:280-299, 1996.

[77] Subramanian, S. and A. Deaton. Gender Effects in Indian Consumption Patterns. Discussion Paper no.147, Research Program in Development Studies, Woodrow Wilson School, Princeton University, 1990.

[78] Tilak, J.B.J. How Free is Free Primary Education in India? *Economic and Political Weekly*, 31(5): 275-282, 1996.

[79] Tilak, J.B.G. Determinants of Household Expenditure on Education in Rural India. Working Paper Series 88. NCAER, New Delhi. 2002.

[80] Urquiola, M. and V. Caldern. Apples and Oranges: Educational Enrollment and Attainment Across Countries in Latin America and the Caribbean. *International Journal of Educational Development*, 26, 572-590, 2006.

[81] Walker, T. and J. Ryan. Village and Household Economies in India's Semi-Arid Tropics. Baltimore MD: Johns Hopkins University Press, 1990.

[82] Webster, P.J. and C. Hoyos. Prediction of Monsoon Rainfall and River Discharge on 15-30-Day Time Scales, 2004.

[83] Zimmermann, Laura. Reconsidering Gender Bias In Intra-Household Allocation in India. Working Paper, March 2010.

Table 1: School Enrollment in Rural Areas by Age and Gender

	1992/1993		1998/1999		2005/2006	
	girls	boys	girls	boys	girls	boys
age						
6	0.47	0.57	0.71	0.76	0.51	0.53
7	0.57	0.71	0.78	0.86	0.72	0.75
8	0.56	0.76	0.78	0.85	0.77	0.82
9	0.61	0.78	0.80	0.89	0.85	0.89
10	0.56	0.76	0.72	0.85	0.79	0.87
11	0.58	0.81	0.74	0.87	0.84	0.91
12	0.50	0.73	0.63	0.79	0.74	0.84
13	0.46	0.74	0.60	0.78	0.70	0.81
14	0.39	0.68	0.52	0.72	0.66	0.74
6 - 10	0.55	0.72	0.76	0.84	0.73	0.77
11 - 14	0.48	0.74	0.62	0.79	0.74	0.82

Notes: Author's calculation using NFHS 1992/93, 1998/99 and 2005/06. Observations are weighted.

Table 2: Summary Statistics of Rainfall Shocks

	mean	std.dev	min	max
shock	-0.2396	0.9919	-3.0739	2.8243
lag 1	-0.0182	0.7901	-2.3098	2.1448
lag 2	0.0901	0.9483	-3.2978	3.7539
lead 1	-0.2054	0.8823	-2.8513	4.0243
lead 2	0.1059	1.1071	-2.7740	4.1786
distribution of shock variable by magnitude (percent of shocks):				
shock<-2.5	2.1			
-2.5 < shock < -1.5	7.7			
-1.5 <shock < -0.5	27.4			
-0.5 < shock < 0.5	40.8			
0.5 < shock < 1.5	17.9			
1.5 < shock < 2.5	3.7			
shock > 2.5	0.5			
N	158532			

Table 3: Summary Statistics of Children by Age and Gender

	girl	boy
age 6	0.12	0.13
age 7	0.12	0.11
age 8	0.13	0.13
age 9	0.1	0.1
age 10	0.13	0.14
age 11	0.08	0.08
age 12	0.13	0.13
age 13	0.09	0.09
age 14	0.1	0.1
household size	7.78	7.57
age of household head	45.28	45.19
own house	0.95	0.95
own agricultural land	0.66	0.66
wealth index	0.38	0.35
N	76541	81965

Note: Wealth index was created by principal component analysis of asset and living condition questions for rural households. Own house, own agricultural land and wealth index are available only for a subsample.

Table 4: The Impact of Rainfall Shocks on School Enrollment in Rural Areas

	(1)	(2)	(3)	(4)	(5)	(6)
	girls	boys	girls	boys	girls	boys
	6 - 14	6 - 14	6 - 10	6 - 10	11 - 14	11 - 14
shock	0.0230***	0.0131**	0.0252***	0.0123*	0.0174**	0.0129*
	(0.0070)	(0.0058)	(0.0074)	(0.0064)	(0.0085)	(0.0068)
age dummies	Yes	Yes	Yes	Yes	Yes	Yes
fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
p-value of gender equality test	0.047		0.025		0.549	
Observations	76228	81549	45700	48920	30528	32629
R-squared	0.191	0.1063	0.1917	0.1299	0.2076	0.1001
mean of dependent variable	0.6556	0.7949	0.6881	0.7985	0.6070	0.7895
percent increase	3.5	1.6	3.7	1.5	2.9	1.6

Robust standard errors for clustering at district level in parenthesis

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Fixed effects here include month, year and district fixed effects

Table 5: Age Non-linearities in the Impact of Rainfall Shocks

	(1)	(2)
	girls 6 - 14	boys 6 - 14
shock	0.0492***	0.0474***
	(0.0088)	(0.0077)
shock-age interactions		
shock*age 7	-0.0064	-0.0093
	(0.0072)	(0.0067)
shock*age 8	-0.0169***	-0.0317***
	(0.0059)	(0.0065)
shock*age 9	-0.0190**	-0.0309***
	(0.0073)	(0.0067)
shock*age 10	-0.0244***	-0.0432***
	(0.0066)	(0.0071)
shock*age 11	-0.0370***	-0.0443***
	(0.0080)	(0.0074)
shock*age 12	-0.0445***	-0.0476***
	(0.0077)	(0.0080)
shock*age 13	-0.0473***	-0.0588***
	(0.0082)	(0.0081)
shock*age 14	-0.0526***	-0.0552***
	(0.0088)	(0.0076)
age dummies	Yes	Yes
fixed effects	Yes	Yes
p-value of F test of joint significance	0	0
p-value of gender equality test		
shock	0.831	
shock at age 7	0.5453	
shock at age 8	0.0244	
shock at age 9	0.0726	
shock at age 10	0.0024	
shock at age 11	0.2825	
shock at age 12	0.5238	
shock at age 13	0.1303	
shock at age 14	0.6009	
Observations	76228	81549
R-squared	0.1923	0.1084

Robust standard errors for clustering at district level in parenthesis
*** p<0.01, ** p<0.05, * p<0.1
Fixed effects here include month, year and district fixed effects

Table 6: The Non-linear Impact of Rainfall Shocks on School Enrollment in Rural Areas

	(1)	(2)	(3)	(4)	(5)	(6)
	girls	boys	girls	boys	girls	boys
	6 - 14	6 - 14	6 - 10	6 - 10	11 - 14	11 - 14
shock<-2.5	0.0195	-0.0269	0.0390	-0.0101	-0.0033	-0.0396
	(0.0285)	(0.0467)	(0.0277)	(0.0457)	(0.0373)	(0.0633)
-2.5<shock<-1.5	-0.0623***	-0.0363**	-0.0603**	-0.0235	-0.0581*	-0.0542***
	(0.0237)	(0.0183)	(0.0248)	(0.0199)	(0.0297)	(0.0207)
-1.5<shock<-0.5	-0.0352**	-0.0207*	-0.0353**	-0.0160	-0.0327*	-0.0285**
	(0.0151)	(0.0118)	(0.0164)	(0.0138)	(0.0180)	(0.0128)
0.5<shock<1.5	0.0044	-0.0016	0.0109	0.0085	-0.0080	-0.0151
	(0.0175)	(0.0132)	(0.0185)	(0.0151)	(0.0216)	(0.0157)
1.5<shock<2.5	0.0423	0.0083	0.0644*	0.0289	0.0128	-0.0260
	(0.0295)	(0.0238)	(0.0350)	(0.0273)	(0.0361)	(0.0243)
shock>2.5	0.0516	0.0212	0.0974	0.0100	-0.0256	0.0230
	(0.0731)	(0.0554)	(0.0671)	(0.0680)	(0.0819)	(0.0469)
age dummies	Yes	Yes	Yes	Yes	Yes	Yes
fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
p-value of F test of joint significance	0.0104	0.2992	0.003	0.545	0.355	0.0577
p-value of gender equality test						
shock<-2.5	0.227		0.2520		0.46	
-2.5<shock<-1.5	0.076		0.037		0.866	
-1.5<shock<-0.5	0.148		0.098		0.789	
0.5<shock<1.5	0.571		0.848		0.681	
1.5<shock<2.5	0.106		0.146		0.213	
shock>2.5	0.257		0.003		0.315	
Observations	76228	81549	45700	48920	30528	32629
R-squared	0.1913	0.1063	0.1922	0.1299	0.2078	0.1005

Robust standard errors for clustering at district level in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Table 7: The Non-linear Impact of Rainfall Shocks and Age

	shock<-2.5		-2.5<shock<-1.5		-1.5<shock<-0.5	
	girl	boy	girl	boy	girl	boy
age						
6	-0.0139	-0.0922*	-0.0858***	-0.0950***	-0.0810***	-0.0720***
7	-0.0389	-0.1026**	-0.0975***	-0.0878***	-0.0556***	-0.0723***
8	0.0339	-0.0378	-0.101***	-0.0179	-0.0441**	-0.0104
9	0.0090	-0.0367	-0.0703**	-0.0371	-0.0241	-0.0182
10	-0.0197	-0.0118	-0.0675**	-0.0091	-0.0305*	-0.0019
11	-0.0006	0.0151	-0.0162	-0.0388*	-0.0254	0.0040
12	0.0646	0.0156	-0.0151	-0.0101	0.0021	0.0035
13	0.0650	-0.0007	-0.0290	0.0124	-0.0253	0.0094
14	0.12**	0.0695	-0.0419	-0.0313	-0.0241	-0.0166
	0.5<shock<1.5		1.5<shock<2.5		shock>2.5	
	girl	boy	girl	boy	girl	boy
age						
6	0.0358*	0.0252	0.1088**	0.0651**	-0.0182	0.0095
7	0.0010	0.0038	0.1033***	0.0191	0.2008***	0.0826***
8	0.0071	0.0183	0.0552	0.0192	0.0708	0.0663
9	0.0098	0.0060	0.0907**	0.0169	0.1021	0.0760
10	0.0140	0.0011	0.0103	0.0114	0.0074	-0.0449
11	-0.0031	-0.0206	0.0457	0.0013	-0.0075	0.0891*
12	0.0022	0.0013	0.0384	-0.0174	0.0072	-0.0281
13	-0.0097	-0.0274	-0.0320	-0.0221	0.1003	0.0234
14	-0.0242	-0.0371**	-0.0368	-0.0259	0.0036	-0.0664

Robust standard errors for clustering at district level in parenthesis
*** p<0.01, ** p<0.05, * p<0.1
Fixed effects here include month, year and district fixed effects

Table 8: The Impact of Lagged Rainfall Shocks on School Enrollment

	(1)	(2)	(3)	(4)	(5)	(6)
	girls	boys	girls	boys	girls	boys
	6 - 14	6 - 14	6 - 10	6 - 10	11 - 14	11 - 14
shock	0.0233***	0.0133**	0.0253***	0.0120*	0.0178**	0.0136**
	(0.0069)	(0.0058)	(0.0073)	(0.0065)	(0.0084)	(0.0068)
lag 1	0.0109	0.0064	0.0103	0.0043	0.0140	0.0104
	(0.0081)	(0.0063)	(0.0086)	(0.0072)	(0.0096)	(0.0070)
lag 2	-0.0021	-0.0017	-0.0011	0.0019	-0.0034	-0.0057
	(0.0072)	(0.0061)	(0.0080)	(0.0074)	(0.0084)	(0.0061)
age dummies	Yes	Yes	Yes	Yes	Yes	Yes
fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
p-value of F test of joint significance	0.004	0.116	0.0037	0.26	0.0847	0.134
p-value of gender equality test						
shock	0.048		0.022		0.582	
lag 1	0.423		0.334		0.657	
lag 2	0.927		0.556		0.722	
Observations	76228	81549	45700	48920	30528	32629
R-squared	0.1911	0.1063	0.1918	0.13	0.2078	0.1003

Robust standard errors for clustering at district level in parenthesis

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Fixed effects here include month, year and district fixed effects

Table 9: IV Results of the Impact of Rainfall Shocks on School Enrollment

	(1)	(2)	(3)	(4)	(5)	(6)
	girls	boys	girls	boys	girls	boys
	6 - 14	6 - 14	6 - 10	6 - 10	11 - 14	11 - 14
shock	0.0239***	0.0124**	0.0252***	0.0118*	0.0190**	0.0120*
	(0.0075)	(0.0062)	(0.0080)	(0.0071)	(0.0086)	(0.1004)
age dummies	Yes	Yes	Yes	Yes	Yes	Yes
fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76043	81365	45583	48803	30460	32562
R-squared	0.1915	0.1066	0.1922	0.1303	0.2081	0.1004

Robust standard errors for clustering at district level in parenthesis

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Fixed effects here include month, year and district fixed effects

shock was instrumented using rainfall shocks in the 5 closest districts as instruments

Table 10: Test for Selection into the Sample

	(1)	(2)
	girls	boys
	6 - 14	6 - 14
shock	0.1025	0.1432
	(0.3393)	(0.3455)
age dummies	Yes	Yes
fixed effects	Yes	Yes
Observations	8835	8961
R-squared	0.036	0.0368

Robust standard errors for clustering at district level in parenthesis
*** p<0.01, ** p<0.05, * p<0.1
Fixed effects here include month, year and district fixed effects

Table 11: The Impact of Rainfall Shocks on School Enrollment in Urban Areas

	(1)	(2)	(3)	(4)	(5)	(6)
	girls	boys	girls	boys	girls	boys
	6 - 14	6 - 14	6 - 10	6 - 10	11 - 14	11 - 14
shock	-0.0046	-0.0545**	0.0190	-0.0479*	-0.0340	-0.0637
	(0.0144)	(0.0266)	(0.0231)	(0.0253)	(0.0364)	(0.0397)
age dummies	Yes	Yes	Yes	Yes	Yes	Yes
fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
p-value of gender equality test	0.072		0.002		0.659	
p-value of equality with rural	0.093	0.013	0.795	0.02	0.172	0.057
Observations	9095	9841	5106	5574	3989	4267
R-squared	0.0595	0.0528	0.0561	0.0622	0.0735	0.0545

Robust standard errors for clustering at district level in parenthesis

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Fixed effects here include month, year and district fixed effects

Figure 1: Adaptation of the Roy model to illustrate the three possible states of children's school enrollment problem

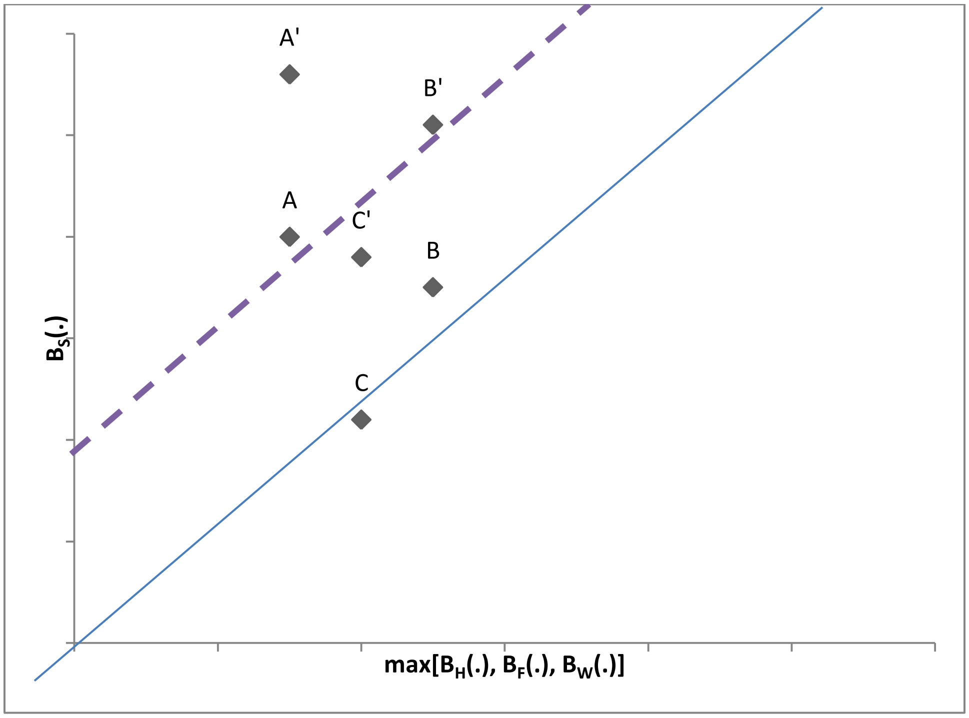


Figure 2

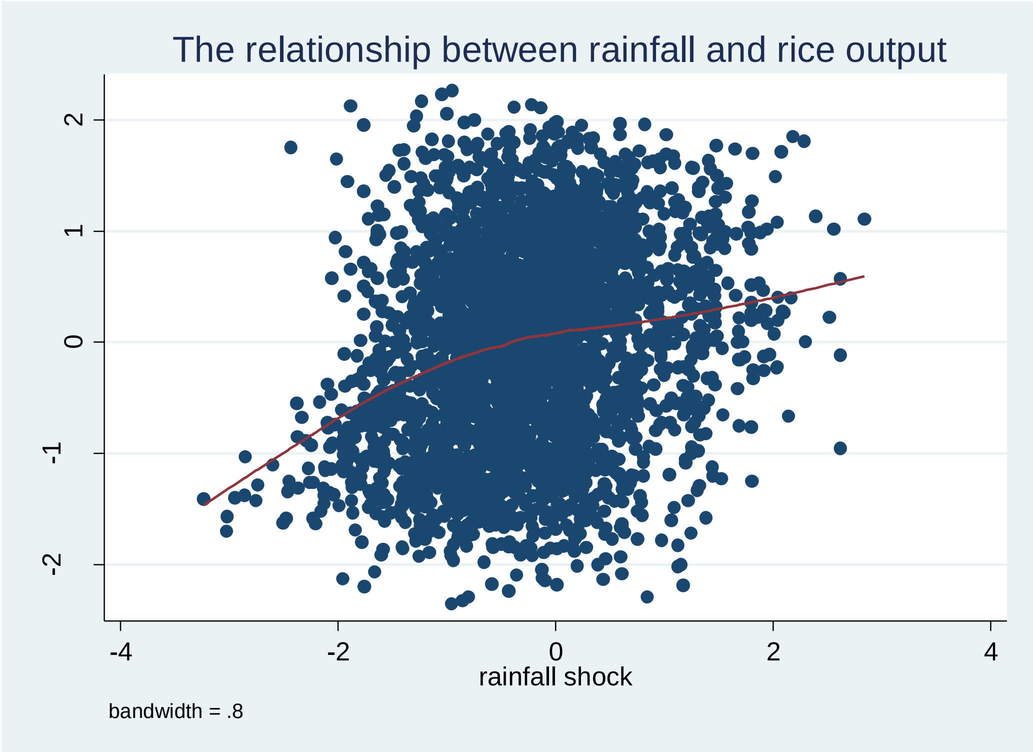
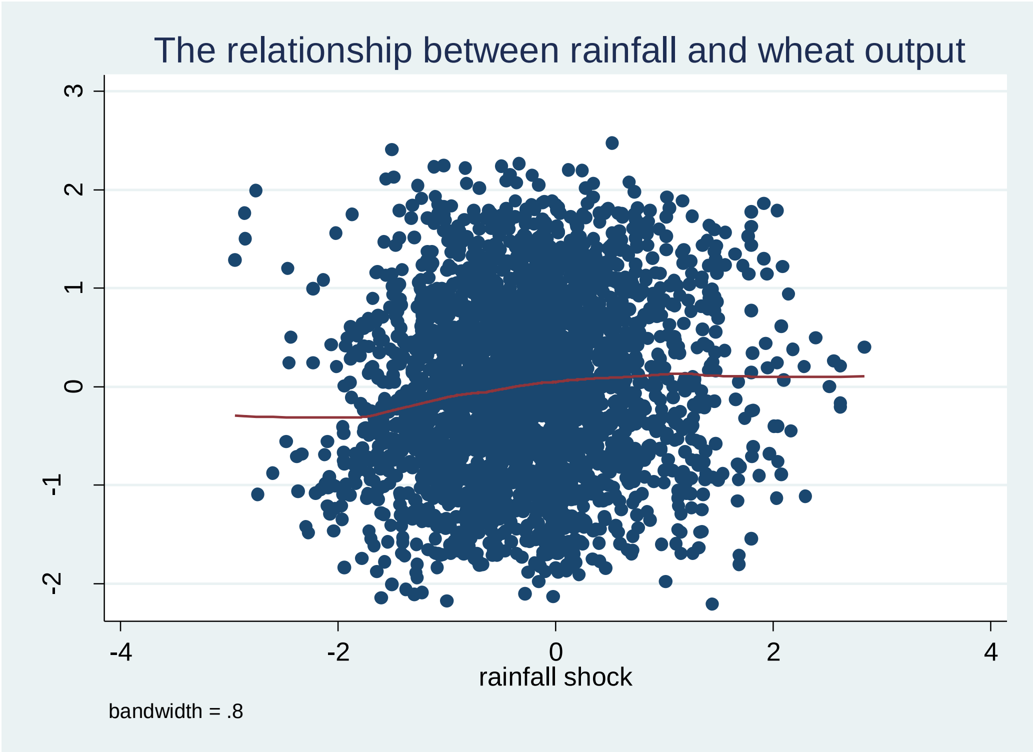


Figure 3



Appendix Table A1: Complete Regression Results for Non-Linearities in Rainfall Shocks and Age

	(1)	(2)
	girls 6-14	boys 6-14
shock<-2.5	-0.0139	-0.0922*
	0.0419	0.054795
shock-age interactions		
shock*age 7	-0.0250	-0.0104
	0.0557	0.046459
shock*age 8	0.0478	0.054433
	0.0344	0.037869
shock*age 9	0.0229	0.055516
	0.0526	0.034058
shock*age 10	-0.0058	0.0804**
	0.0305	0.037983
shock*age 11	0.0133	0.1073**
	0.0418	0.047385
shock*age 12	0.0785*	0.1078***
	0.0403	0.033353
shock*age 13	0.0789	0.0915*
	0.0528	0.055166
shock*age 14	0.1339***	0.1617***
	0.0460	0.055537
		-
-2.5<shock<-1.5	-0.0858***	0.0950***
	0.0282	0.028905
shock-age interactions		
shock*age 7	-0.0117	0.007152
	0.0250	0.029136
shock*age 8	-0.0152	0.0771***
	0.0228	0.028223
shock*age 9	0.0155	0.0579**
	0.0270	0.028174
shock*age 10	0.0183	0.0859**
	0.0231	0.034488
shock*age 11	0.0696**	0.0562*
	0.0310	0.028686
shock*age 12	0.0707**	0.0849**
	0.0278	0.038549
shock*age 13	0.0568*	0.1074***
	0.0301	0.032483
shock*age 14	0.0439	0.0637*
	0.0298	0.033193
		-
-1.5<shock<-0.5	-0.0810***	0.0720***

	0.0195	0.019155
shock-age interactions		
shock*age 7	0.0254	-0.00027
	0.0180	0.016115
shock*age 8	0.0369**	0.0616***
	0.0153	0.015925
shock*age 9	0.0569***	0.0538***
	0.0190	0.019056
shock*age 10	0.0505***	0.0701***
	0.0184	0.017574
shock*age 11	0.0556***	0.0760***
	0.0209	0.019192
shock*age 12	0.0831***	0.0755***
	0.0212	0.019666
shock*age 13	0.0557**	0.0814***
	0.0232	0.020691
shock*age 14	0.0569**	0.0554**
	0.0231	0.022174
0.5<shock<1.5	0.0358*	0.0252
	0.0216	0.0187
shock-age interactions		
shock*age 7	-0.0348**	-0.0214
	0.0164	0.0149
shock*age 8	-0.0287*	-0.0069
	0.0167	0.0161
shock*age 9	-0.0260	-0.0192
	0.0181	0.0167
shock*age 10	-0.0218	-0.0241
	0.0172	0.0154
shock*age 11	-0.0389*	-0.0458**
	0.0221	0.0184
shock*age 12	-0.0336	-0.0239
	0.0212	0.0189
shock*age 13	-0.0455*	-0.0526***
	0.0231	0.0201
shock*age 14	-0.0600**	-0.0623***
	0.0250	0.0200
1.5<shock<2.5	0.1088**	0.0651**
	0.0431	0.0294
shock-age interactions		
shock*age 7	-0.0055	-0.0460*
	0.0332	0.0255
shock*age 8	-0.0536*	-0.0459
	0.0278	0.0307
shock*age 9	-0.0181	-0.0482*
	0.0367	0.0279
shock*age 10	-0.0985***	-0.0537*

	0.0375	0.0309
shock*age 11	-0.0631	-0.0638**
	0.0423	0.0318
shock*age 12	-0.0704	-0.0825**
	0.0433	0.0395
shock*age 13	-0.1408***	-0.0872**
	0.0434	0.0386
shock*age 14	-0.1456***	-0.0910**
	0.0516	0.0371
shock>2.5	-0.0182	0.0095
	0.0469	0.0366
shock-age interactions		
shock*age 7	0.2190**	0.0731
	0.0931	0.0585
shock*age 8	0.0890	0.0568*
	0.1117	0.0303
shock*age 9	0.1203	0.0666
	0.0975	0.0405
shock*age 10	0.0256	-0.0544
	0.1315	0.0665
shock*age 11	0.0107	0.0796
	0.1031	0.0527
shock*age 12	0.0253	-0.0376
	0.0692	0.0630
shock*age 13	0.1185	0.0139
	0.0923	0.0309
shock*age 14	0.0218	-0.0759
	0.0482	0.0764
age dummies	Yes	Yes
fixed effects	Yes	Yes
p-value of F test of joint significance	0	0
p-value of gender equality test		
shock<-2.5	0.136	
shock at age 7	0.1971	
shock at age 8	0.1682	
shock at age 9	0.4554	
shock at age 10	0.8377	
shock at age 11	0.8231	
shock at age 12	0.3894	
shock at age 13	0.4229	
shock at age 14	0.4085	
-2.5<shock<-1.5	0.777	
shock at age 7	0.7411	
shock at age 8	0.0043	
shock at age 9	0.2487	
shock at age 10	0.0345	
shock at age 11	0.4657	
shock at age 12	0.8445	
shock at age 13	0.2094	
shock at age 14	0.7533	

-1.5<shock<-0.5	0.629	
shock at age 7	0.3738	
shock at age 8	0.0254	
shock at age 9	0.7416	
shock at age 10	0.099	
shock at age 11	0.1245	
shock at age 12	0.9378	
shock at age 13	0.0782	
shock at age 14	0.7424	
0.5<shock<1.5	0.592	
shock at age 7	0.8791	
shock at age 8	0.5351	
shock at age 9	0.8336	
shock at age 10	0.4317	
shock at age 11	0.3739	
shock at age 12	0.9655	
shock at age 13	0.4008	
shock at age 14	0.5531	
1.5<shock<2.5	0.313	
shock at age 7	0.0027	
shock at age 8	0.2754	
shock at age 9	0.0809	
shock at age 10	0.9752	
shock at age 11	0.2378	
shock at age 12	0.1351	
shock at age 13	0.7785	
shock at age 14	0.7852	
shock>2.5	0.575	
shock at age 7	0.0376	
shock at age 8	0.9441	
shock at age 9	0.6398	
shock at age 10	0.2267	
shock at age 11	0.2248	
shock at age 12	0.5601	
shock at age 13	0.0646	
shock at age 14	0.4719	
Observations	76228	81549
R-squared	0.193	0.1031

Robust standard errors for clustering at district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Fixed effects here include month, year and district fixed effects

Appendix Table A2: First Stage IV Results

	(1) girls 6-14	(2) boys 6-14
IV1	0.5611***	0.5689***
	(0.0700)	(0.0701)
IV2	0.2323***	0.2345***
	(0.0681)	(0.0664)
IV3	0.0586	0.0506
	(0.0537)	(0.0535)
IV4	0.0885	0.0916
	(0.0579)	(0.0567)
IV5	0.0332	0.0308
	(0.0351)	(0.0360)
age dummies	Yes	Yes
fixed effects	Yes	Yes
Observations	76043	81365
F statistic	393.361	386.335
R-squared	0.9767	0.9763

Robust standard errors for clustering at district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Fixed effects here include month, year and district fixed effects