Child Labor and Schooling Decisions in Ghana

Comments Welcome

by

Michael A. Boozer[1]

and

Tavneet K. Suri

First Draft: April, 2001
Revised: September, 2001
For submission to the September, 2001 NEUDC Conference

ABSTRACT

In this paper we investigate the choices involved in the tradeoff between a child's labor outside of household work and their schooling hours from a sample of data from Ghana in the late 1980's. While households are interviewed only once during the survey, the data were collected over an 11 month period for both the Northern and Southern regions of Ghana. Northern and Southern Ghana have rather distinct long-run rainfall patterns over the year. These two facts allow us to derive an identification strategy with good power for the child labor decisions made by households. Essentially, we use the month by region variation in child labor intensities as a source of exogenous variation in child labor, while controlling for secular month and region variation in child labor intensities. This approach uses the monthly span of the data to construct an identification strategy that uses the synthetic panel approach of Deaton (1985) to average out the unobserved preference and short-run income fluctuations across month by region cells. We also consider a refined version of this strategy, in using both the realized rainfall (including sets of month and region dummies), as well as realized rainfall deviated from its long-term month by region means, as direct rainfall shock measures. These differing ways of using the rainfall data extract the different behavioral responses to short run and long run rainfall, and

[1] Boozer is at the Department of Economics and the Economic Growth Center, Yale University. Email: michael.boozer@yale.edu. Suri is at the Department of Economics, Yale University. Email: tavneet.suri@yale.edu. We are thankful to the Ghana Statistical Service for permission to use their data and also to David Lister at the University of East Anglia, UK for access to the rainfall data. We thank Markus Goldstein, Chris Udry and the Development Lunch participants at Yale for helpful comments.

thus to a degree income patterns. We interpret the strategy that uses simply month by region variations as being indicative of an average of the long and short run effects and that using the rainfall shock as indicative of the pure short run effect. We find that both strategies give statistically similar impacts: an hour of child labor reduces school attendance by approximately 0.38 hours. Surprisingly, this effect is actually *larger* than the OLS estimate of the effect, -0.21. Also, these results imply that income, or poverty, may not be as important in determining child labor. A direct analysis of this effect indeed provides no evidence of heterogeneity of this substitution effect by our observed income measure.

1 Introduction

There has been an increased concern, both in academic and policy circles, about child labor and whether or not concerted attempts to eliminate it would enhance the overall well-being of the child and the economy in question. What actually constitutes child labor is often not very clear. The ILO Minimum Age Convention, 1973 (No. 138) states that ratifying members "shall raise progressively the minimum age for admission to employment or work to a level consistent with the fullest physical and mental development of young persons." The minimum age specified is 15 years, although there is some qualification to that: "... developing countries, whose economies, educational and administrative facilities are insufficiently developed, may initially permit children of 12-14 years of age to carry out light work of certain types and under certain conditions."

The main concern of the ILO is therefore not child work per se, but child work that is detrimental to a child's physical and or mental health. The ILO's Convention No. 182 (1999) has just come into force, which concerns the prohibition and elimination of the "worst forms" of child labor[2], citing poverty as the main cause of these forms of child labor. The ILO (1996) estimates that about 250 million children, ages five to fourteen, were engaged in economic activity in 1996 and of these, 120 million work full-time. The rest combine economic activities with schooling and other non-economic activities. Regionally, the ILO estimates that Asia has the largest number of child workers, but that incidence is highest in Africa (41% compared to 22% in Asia and 17% in Latin America).

The United Nations Convention on the Rights of the Child captures its version of these ideas in its Article 132, which urges states "... to recognize the right of the child to be protected from economic exploitation and from performing any work that is likely to be hazardous or *to interfere with the child's education*, or to be harmful to the child's health or physical, mental, spiritual, moral or social development". UNESCO estimates that about 20% of primary-school age children (i.e. about 128 million children) were not receiving any education in 1990, a figure often quoted by the ILO to emphasize the extent of child labor across the world. This stance has more of an emphasis on the education aspect of the situation, with the implicit assumption that a direct trade-off necessarily exists between education and child labor with regard to human capital accumulation, i.e. that children who are not in schools *must* be working.

This is where the motivation for this paper lies - does the fact that children are a part of the labor market *cause* them to get less education, in terms of the short run tradeoff between time spent in school versus time spent at work? It is important to note that the analysis of child labor here does not encompass

[2]What the ILO categorizes as the "worst forms" of child labor include slavery, prostitution, drug trafficking and any activities that harm the healthy, safety and morals of a child.

forms of child labor that fall into the categories of the worst forms of child labor, as referred to by the ILO.

We are interested in whether there is any evidence of a trade-off between hours worked and hours going to school. If so, it begs the question as to what policy is most appropriate to encourage these households to invest in their children's education. This analysis hopes to emphasize *where* this trade-off is coming from and how this should relate back to current policy issues. We make no claim as to the relative benefits of attending school versus non-household work with regards to human capital accumulation, as we have no way of empirically assessing the relative welfare impacts of lower school attendance vs. increased labor for these children. Similarly, we do not attempt to attempt to explain *why* households choose to make these decisions, i.e. whether it is because of poor school quality or lack of functioning (adult) labor and/or credit markets - we leave these questions to future research projects. What we are interested in is therefore just the substitution effect between child labor and schooling net of income, preference, and institutional factors such as school vacations. In addition, as earlier work has left unclear the empirical importance of income (or poverty status) in determining child labor intensity, we investigate this issue directly. The cross-section evidence is driven by both income and substitution effects, among other factors. We try to isolate the pure substitution effect and can thus infer the income effect by contrasting these two sets of estimates. Finally, we also gain some insight into the role income plays by examining the heterogeneity in the substitution effects by reported household income.

We use a cross-section of data from Ghana collected over October 1988 to August 1989 on 3374 households. The central identification problem we face in estimating the causal effect of child labor choices on schooling decisions is that these are simultaneously driven by household income, family preferences toward work and schooling decisions, and institutional factors such as school vacations. All of these are likely to induce a negative relationship between schooling and child labor in a single cross-section of data. To abstract from these confounding factors, we use the distinct rainfall patterns of Northern and Southern Ghana to provide variation in the demand for child labor on the production side of the household's decision process.

Empirically, we can utilize *any* variation that arises by month by region cells in the child labor and schooling decisions to identify the effect of child labor, while controlling for regional and monthly secular effects (which would account for cultural, crop, and vacation effects, etc.). Apart from how we treat the covariates, this IV strategy using cell dummies as instruments amounts to treating the month by region cells as synthetic panel data as in Deaton (1985). By allowing for secular month and region main effects, we argue that the residual variation in child labor and schooling is purged of systematic preference and income variation. But, it is identified by factors that induce variation in child labor decisions at the month by region level, which subsequently drive schooling decisions at the margin.

We also refine this IV strategy by utilizing a second strategy that relies solely on the rainfall patterns themselves, while controlling for secular month and region effects. These two strategies work well in tandem: the first strategy utilizes variation in the child labor driven *both* by long and short run characteristics that drive demand for child labor by month by region cell, while the second isolates the purely short run variations in realized rainfall that drive the demand for child labor. A combination of these two also enables us to distinguish what we interpret as the purely short-term from the purely long-term effects. It is important that we have the ability to look at both. One hypothesis that would destroy the IV strategy based on actual rainfall is if decisions made by the household are driven largely by current, not long-run, income as in a non-myopic life-cycle model of household decision making. Such a behavioral framework has strong *a priori* validity in Ghana, and so it is important we are able to address it with our dual identification strategies.

Both strategies have good power in describing the fluctuations in child labor patterns within month by region cells. They give similar results: a significantly negative effect of child labor on schooling, particularly for boys. We also analyze the extensive margin (i.e. children moving in and out of the labor market) vs. the intensive margin (i.e. the response of child labor, conditional on the child already being *in* the labor market). For boys, there is a negative significant impact at both of these margins, while for girls the relationship is strong at the extensive margin, but essentially non-existent at the intensive margin. This evidence seems to imply that for girls, being drawn into the non-household labor market by rainfall conditions favorable for working is more disruptive than for boys. However, conditional on working, the effect of the marginal hour of labor for girls has no detectable impact on the time spent in schooling, unlike for boys, where the impacts at both margins are more uniform.

The rest of the paper is structured as follows. In Section 2, we discuss the relevance of the question at hand to policy, relating our motivation behind this question to prior work on child labor. Section 3 lays out a static model in which to think about household decision making processes *vis a vis* child labor and schooling. In Section 4, we then discuss the empirical setting of the problem and the Ghanaian data. Section 5 deals with carefully describing the identification strategies and Section 6 discusses the results, looking at the gender heterogeneity at the extensive and intensive margins. Section 7 concludes.

2 Child Labor and Policy

Recent attempts have been made in the developed world to advocate strong policies to counteract the child labor phenomenon that seems to be characteristic of a large part of the developing world. This has led to policies such as straightforward bans on child labor or trading sanctions. One reason for the negative view of child labor, even apart from the welfare impact on the children themselves, is that it may result from behavior that is privately optimal, but socially sub-optimal. This may arise if the social return to schooling in a developing country is large (owing to externality effects), but the private return is lower than an individual's discount rate, and so they choose relatively low levels of schooling. Limiting the extent to which employers may use child labor may then lead to superior outcomes from a societal perspective owing to the externality effects, and so move the economy to a 'good equilibrium'.

The aim of this paper is to empirically determine the *causal* relationship between child labor and schooling attendance and in the process shed some light on what decisions households are faced with and how policy can be shaped to account for that. It is not immediately obvious that bans or trade sanctions are necessarily the correct way to approach child labor, if at all we want to. That is where a policy discussion based on our analysis becomes useful - to shed light on some of the driving forces that may or may not exist behind these decisions to send children to work vs. to school.

In terms of prior literature, the work on child labor has tended to be of two strands. On the theoretical side, for example, Basu and Van (1998) analyze a model in which child labor exists, only as one of multiple equilibria. They illustrate how, in this 'bad equilibrium', a ban would be benign. The ban simply shifts the economy to a 'good equilibrium' (i.e. one where there is no child labor), after which it plays no role (no individuals voluntarily *want* to shift away from this good equilibrium once they reach it). This aspect of the literature focuses on poverty being the main raison d'etre of child labor. Other theoretical literature, in a similar vein, analyzes the (possibly multiple) equilibrium scenarios that may be possible in economies where child labor is prevalent and looks at the implications of different equilibrium scenarios in terms of welfare and Pareto efficiency. For example, Baland and Robinson (2000) analyze a model where child labor is a facet of poverty and they show how an equilibrium in which child labor exists can be inefficient, even if we ignore the social return to human capital.

On the empirical front, there has been a large literature on child labor, mostly trying to analyze the determinants of this child labor phenomenon with the hope that such analyses can shed light on what direction policy should be aimed. For example, Moehling (1999) looks at the child labor laws in late nineteenth century in the US, and finds that they had little to do with the subsequent fall in child employment. The theoretical models often rely on fully functioning markets, like Basu and Van (1998) who rely on the existence of functioning labor markets. A number of empirical studies have tried to steer specifically away from this, often attempting to quantify the market imperfections that may exist. For example, Bhalotra and Heady (2000) emphasize the role that a lack of functioning labor markets can play and how asset ownership (in particular farm size) can play a role in the incidence of child labor. They concentrate their analysis on *only* households that own land. However, in our sample there is a high incidence of child labor among households who do not own any land. About 41% of households do not own any land and about 24% of children who participate in the labor market in our sample belong to households that do not own land.

On a slightly different note, Jacoby and Skoufias (1997) look at the interplay of credit market imperfections and human capital accumulation and how seasonal fluctuations in schooling are a form of self-insurance and do not result in substantial losses in the accumulation of human capital. We attempt to abstract from these sorts of studies by not quantifying any of the market imperfections that *may* (and probably do in Ghana) exist per se, but just focusing on the causal relationship between child labor and schooling without any welfare judgments.

A lot of the recent controversy regarding what policy approach to take when thinking of child labor centers around the role played by poverty, as dictated by the theoretical literature. A number of studies, however, for example Canagarajah and Coulombe (1997) find that, contrary to most beliefs, poverty is *not* the main culprit in determining child labor, though it may have indirect effects on child labor through the effects of poor quality schooling. Bhalotra (1999), too, tackles this poverty question head on. She defines her 'poverty hypothesis' to imply a case where the wage elasticity of child labor is negative and in such a case improving the returns to education is unlikely to have substantial impact on child labor. She finds that this wage elasticity is significantly negative for boys, but insignificantly different from zero for girls in rural Pakistan. Our paper is in a similar vein, in that we try to analyze a number of effects present in a cross section of households, focusing on household decision making and the gender heterogeneity that may exist. We try to distinguish *between* the different effects productivity shocks may have on household decision making and use these results to discuss what this may imply for policy stances. This paper also reminds us of how these issues are extremely economy specific and how general worldwide policies, like sanctions and bans are simply inappropriate in some parts of the world.

3 A Theoretical Framework to Consider the Influence of Rainfall on Child Labor and Schooling Choices

In this section we present a simplified theoretical framework for the household's decision process in choosing child labor and schooling. The model is dynamic to illustrate the short-run versus the long-run effects of rainfall fluctuations.

Our theoretical framework focuses on the *behavioral* factors, which affect the joint determination of child labor and schooling choices for a household. One key factor we consider in the econometric framework to abstract from this joint determination is a measure of a rainfall (or weather) shock that induces variations in child labor, and therefore, *subsequent* fluctuations in schooling hours. It would thus be ideal if the rainfall shock only entered the model on the production side of the household decision process, and there *only* impact the marginal product of child labor, thereby altering the implicit child shadow wage. So, in our modeling framework below, we assume that *current* rainfall - holding constant long run rainfall patterns (which we take to be isomorphic to log run income) - indeed affects only the current marginal productivity of child abor.[3]

It is easy to imagine a more general framework whereby rainfall also impacts the productivity of the household more generally and generates an *income* effect on the production side. We need to also consider the consumption side of the household and the possibility that rainfall leads to a decrease in the allocation of time to all activities including labor and schooling activities. These two possible confounding influences of the rainfall shock may lead to failure of the exclusion restrictions we consider where rainfall affects schooling choices *only* through its impact on child labor. And, these two reasons just given would imply that rainfall has direct and *distinct* influences on *both* child labor and schooling, which invalidates rainfall as a useful instrument. The behavioral model in this section delineates clearly the structure necessary to justify our empirical strategy.

In a dynamic setting, we assume the "assets" of the household accrue due to a household production function that is separable from the child labor production function, but which feeds into the child labor decision problem via the budget constraint. While this rules out any interaction between the marginal product of child and other household labor, it does not rule out the interesting case of child labor fulfilling an excess demand for labor role. If current rainfall is almost (mean) independent of rainfall in the next period, then it will induce largely a substitution effect of increasing the demand for child labor on the productivity side. However, if rainfall patterns are quite persistent, then current rainfall is a good proxy for long-run rainfall, and so a blip in current rainfall will induce both a positive substitution effect, but a negative income effect on child labor levels. More importantly, the necessary econometric exclusion restriction of rainfall versus the schooling decision is violated in that the income effect exerts a *direct* effect on the schooling decision. This would be true even if the purely *idiosyncratic* component of current rainfall induces a purely substitution effect on child labor, and thus only indirectly on child schooling via its effect on child labor intensities. The degree to which current rainfall (as opposed to a constructed 'shock' measure which tries to isolate the idiosyncratic or temporary

[3]We also assume the rainfall shock affects the marginal productivity of child labor homogeneously *regardless* the size of the rainfall shock. While as a global assumption this rules out the negative effects of flooding, for example, for the magnitudes of rainfall shocks observed in our data, this appears to be a good local assumption.

component) represents more of an income as opposed to a pure substitution effect depends on the persistence of rainfall across time for a given point in the landscape.

With this focus of our approach in mind, let us consider a generic dynamic household utility maximization problem over the arguments of interest to our question. We want to make clear we do not take the standard Beckerian approach of viewing child schooling as an investment good, but instead as an argument of the intertemporal utility function. Again, this only makes sense if our model is interpreted with poetic license, rather than as a literal statement of reality.[4] We denote the per- period schooling decision (in a binary sense, attend or not, and in the continuous sense of hours per week) as S_{it}. We denote the child labor decision (again, either as a discrete or continuous measure for the period under consideration) as L_{it}. We also make the necessary assumptions to allow the child labor schooling decision process to be separable from the household overall consumption decisions, and so take it to be completely ancillary from this process.[5]

We use an intertemporally separable utility function of the form:

$$\max_{\{L_{it}, S_{it}\}} \sum_{t=1}^{T} \beta^t U(F - L_{it}, S_{it}; a_{it}, b_{it}) \tag{1}$$

where $\beta \equiv (1+\rho)^{-1}$ is the household's subjective discount factor (and ρ is the subjective discount rate). The index i denotes a household and the argument F denotes the full time allocation for households, and so child labor enters the household utility function negatively. The finite time horizon for households is represented by T. The parameters a_{it} and b_{it} are taste parameters for child labor and schooling respectively. These will obviously need some additional assumptions made about them in order to allow this structure to be econometrically feasible, but for now just note that if they are taken to be household-specific preference parameters (but time invariant), then a fixed effects framework will account for such preference variation across households.

As we discussed above, we assume that the rainfall shock enters only the production side of the household's problem thus affecting only the household's budget constraint. Ignoring consumption, schooling is the only 'purchased' good. Since only the relative prices of child labor and schooling affect the household's choice, we normalize the price of schooling to unity. Conceptually, the price of

[4] In addition, the notion of schooling as a consumption is not contradicted by ethnographic evidence from Africa. There is a large amount of anecdotal evidence on attitudes towards education and how it well may be that for some individuals, schooling is at least partly a consumption good. The externality argument mentioned earlier also has this flavor.

[5] Clearly, this is just for clarity. Consumption could easily be added to the model, although it adds nothing to the empirical implications unless we know something about the complementarity or substitutability between the composite consumption good and the arguments of household utility we consider here. As it adds nothing to the empirical implications absent such assumptions, it is superfluous and so we ignore it.

schooling may be thought of as the amalgam of the opportunity cost of time for time spent in schooling, as well as the more direct shoe-leather cost of schooling such as travel to school and books, clothing, and direct school fees. We abstract from whether or not child laborers are paid actual wages, but instead rely on the weather shock, denoted as θ_{it}, to denote the shadow wage implicitly paid to the children.[6] As long as the benefits are recouped at the household level, the question of direct payment is irrelevant in a unitary household framework. With these considerations in mind, the evolution of the household's budget constraint may be written as:

$$A_{i,t+1}/p_{t+1} = (1 + r_t)(A_{it}/p_t + \theta_{it}L_{it} - S_{it}) \tag{2}$$

where A_{it} represents household assets which are brought into period t, and may include assets accumulated from household decisions not explicitly accounted for in our framework. Of prime interest would be if current rainfall were a substantial contributor to current household assets (i.e. by making *adult* labor more productive), and so rainfall generated both an income (via the aspects explicitly left out of our model) and a substitution effect (which is our focus here) on child labor and schooling for that reason. Here again, we will address this possibility by comparing effects driven by alternative sources of long and short-run rainfall patterns to assess the empirical importance of this conceptual issue.

We take the household and time-specific Lagrange multiplier on the per-period budget constraint to be λ_{it}. In a setting of perfect certainty (although the basic result generalizes to a setting of uncertainty with complete asset markets), the recursion relationship for the evolution of λ_{it} with a constant interest rate r implies:

$$\lambda_{it} = (\frac{1+r}{1+\rho})^t \lambda_{i0} \tag{3}$$

and using the approximation that $\log(1 + x) \approx x$ for small x this simplifies to:

$$\log(\lambda_{it}) = t(r - \rho) + \log(\lambda_{i0}) \tag{4}$$

where λ_{i0} is an household-specific Lagrange multiplier on the household's lifetime budget constraint. While this specific parameterization implies a linear

[6] Clearly, the marginal productivity of child labor could vary across, as well as within, households even apart from the weather shock. The thorny issue for our purposes would be if the cross-household marginal productivity differences also correlated with schooling decisions. In the U.S. literature, this would relate to the familiar 'ability bias' issues of those who get more education are also more productive in the workforce, generating a spurious correlation between the two. Here the sign of this correlation between the MP_L of child labor and the schooling decision is not so clear, even accepting it is present and used in the household decision problem. In addition, lacking any empirical measure of child labor 'quality', this issue is far beyond being empirically accounted for in our setting. While we acknowledge the issue of heterogeneity in child productivities is perhaps an important source of a (spurious) negative correlation between the child labor and schooling decision - more 'able' children are kept home to work and the less 'able' are sent to school - the careful investigation of this issue will have to wait for more suitable data.

trend in time should be included to control to the *aggregate* evolution of the Lagrange multiplier on the budget constraint, the result in the complete markets/uncertainty setting implies only that time- specific dummies (as opposed to a linear trend) be included to account for aggregate uncertainty. That is, we modify the last equation to:

$$\log(\lambda_{it}) = \lambda_t + \log(\lambda_{i0}) \tag{5}$$

With this setup for the dynamic household maximization problem and the law of motion for the relevant portion of the household budget constraint, we can write down the log-linear solution to the partial differential equations given by the first order conditions for the control variables S_{it} and L_{it} as:

$$\log(L_{it}) = \eta\log(\theta_{it}) + \delta\log(\lambda_{it}) + a_{it} + u_{it} \tag{6}$$

where we have appended an error term (u_{it}) to account for optimization errors, measurement errors in hours of work, and, in the case of uncertainty, forecast errors (we are not precise on this last part here for now) in updating the marginal utility of wealth. For the schooling equation, we have not specified if the within-period household utility function is separable in child labor and schooling. If it is (as in commonly assumed in many empirical settings on this and related topics such as adult labor supply), then of course the implicit 'wage' of current rainfall does not exert a direct effect on S_{it} *conditional* on the value of λ_{it}. In the separable case, the log-linear solution for the schooling equation - ignoring the generic stochastic error as in the child labor equation for simplicity - we have:

$$\log(S_{it}) = \pi\log(\lambda_{it}) \tag{7}$$

In the separable case, there is no direct *structural* dependence of the schooling choice on current rainfall *given* the value of the Lagrange multiplier λ_{it}, but if we substitute out for the (unobserved) λ_{it} using the two reduced forms, we get:

$$\log(L_{it}) = \eta\log(\theta_{it}) + \frac{\delta}{\pi}\log(S_{it}) + d_{it} + u_{it} \tag{8}$$

And just re-writing this equation to place the schooling choice on the LHS, we have:

$$\log(S_{it}) = -\frac{\eta\pi}{\delta}\log(\theta_{it}) + \frac{\pi}{\delta}\log(L_{it}) + \alpha_{it} + \epsilon_{it} \tag{9}$$

Finally, let us place one more restriction on the fundamentals on the behavioral equations by assuming the preference variation has the factor structure:

$$\alpha_{it} = \alpha_i + \gamma_t \tag{10}$$

Which implies that, in the cross-household dimension (represented by the factor α_i) households have varying attitudes towards the levels of child labor and schooling they wish their children to attain. And that across time (months of the year, represented by the factor γ_t) *all* households may feel secularly poorer or richer, perhaps owing to seasonal weather patterns and farming operations, leading to aggregate movements in child labor and schooling decisions. This leads to the structural schooling equation:

$$\log(S_{it}) = -\frac{\eta\pi}{\delta}\log(\theta_{it}) + \frac{\pi}{\delta}\log(L_{it}) + \alpha_i + \gamma_t + \epsilon_{it} \tag{11}$$

An item to note as regards this equation is that even conditional on the (endogenous) child labor decision L_{it}, the current rainfall enters the structural schooling equation. It is, however, downweighted relative to the effect of child labor by the parameter η, which is the direct effect of the rainfall shock on the child labor choice. In other words, while we expect current rainfall to exert an influence on the schooling decision not *purely* through the child labor decision, we expect the direct effect of rainfall to be substantially muted relative to the impact of the child labor decision. From an econometric perspective, this suggests we need to examine if the exclusion restriction of current rainfall is valid. If rainfall patterns are rather persistent and so largely absorbed by the combination of the cross- sectional dummies α_i and the time indicators γ_t, then the exclusion restriction of rainfall affecting the schooling decision *only* via its impact on the child labor decision may be valid. While the behavioral model is agnostic on this front, we leave it as an empirical matter to assess the validity of these exclusion restrictions as well as assessing the importance of the substitution versus income effects of the rainfall variables.

4 Institutional Setting and Data

The household data used is the Ghana Living Standards Survey for 1988/1989 from the World Bank. The GLSS Surveys are a set of repeated cross sectional nationwide surveys that gather information at the individual and household level over 1987/88, 1988/89, 1991 and 1998. We make use of the GLSS2 survey conducted over an 11 month period from October 1988 to August 1989. As we discuss below, we take advantage of the fact that the survey was *not* conducted at a narrow interval in time in devising our identification strategy. We also take advantage of the ability to match the GLSS data to geographically proximate rainfall stations, which allows us to construct both long run (i.e. approximately 50 year averages over a 100 year time span) and contemporaneously observed monthly rainfall for our sample of households. The surveys cover about 180 clusters and 3374 households across Ghana. The schooling and labor data in the survey are based on weekly time logs of all the children present in each of the households interviewed above the age of 7. The extract we draw contains 4718 children in the age group 7-18. These children are involved in three main work activities, categorized as main job, secondary job and house work. It is the first two of these that are the most relevant.[7]

Table 1 shows the participation rates of children in various activities for this cross section of 2043 households[8]. About 31% of the children are involved in a main and/or secondary job, 52% attend school (i.e. report non zero hours), 86% are involved in some house work and 12% report both positive non-house labor and schooling hours (i.e. are involved in both schooling and labor). We also break down these into participation rates by gender and then by climatic region (i.e. north and south), to illustrate the variation in participation across genders and across regions.

INSERT TABLE 1 HERE

Table 2 shows the average hours spent per child in each of these activities, across genders and regions, *conditional* on participation. We can see that boys do less housework than girls, and that children already in the labor market in Northern Ghana work, on average, more hours than those in Southern Ghana. This already gives us an indication of some possible heterogeneity by region. We come back to this in more detail later.

INSERT TABLE 2 HERE

When we look at the simple cross-sectional association between child labor and schooling hours using the data from the survey week time logs, we observe a statistically strong negative relationship between the two measures. In Column A of Table 3 we present the OLS estimates of the relationship between schooling hours and child labor hours net of a few control variates which help control for within and between household heterogeneity as regards this relationship. The point estimate indicates that an additional hour spent in child labor is associated with a little more than a 0.2 reduction in hours spent in school during the survey week. The relationship is highly significant with a t- statistic of just over 20. We emphasize here that conclusion from Table 1 that just under half the sample of children report zero hours of schooling during the survey week, and almost 70 percent report no labor hours. The OLS relationship simply pools the zero

[7] We did look at house work as a separate category of time use as such, but our work with these data indicates that, across months for a given region, house work does not vary much. Obviously, as noted by others, girls tend to do more house work than boys. We can also see this in Table 1. However, since the emphasis of this paper is to look at substitution effects, and no such effects are evident in in house work, we restrict our analysis to look at just main and secondary jobs. On average, in our sample, children who engage in house work, work about 13 hours a week in chores and other household duties. And, children who are engaged in the labor market, work an average of about 20 hours a week.

[8] The reason the number of households we have in our sample is less than the full number of households in the GLSS data (3374) is because we restrict our sample to children aged between 7 and 18.

and non-zero observations on both schooling and child labor to present the total association between the two variables in Column A. Later in our empirical work, as we refine the identification strategy, we will consider the separate effects at these two margins, particularly when we look at the results broken out by gender.

INSERT TABLE 3 HERE

Of course, there are many reasons as to why the OLS estimate in Column A may tell us little about the tradeoffs between these two choices faced by a *given* household. For example, one reason why we may expect to see a negative relationship between these two variables has little to do with household choices: i.e. school vacations. In that case schooling hours would necessarily be no larger than when school is in session, and work hours may be larger owing to the greater net (of schooling) time endowment. Similarly, when looking *across* households as OLS does, we may expect to see a negative relationship between child labor and schooling if both child leisure and schooling are normal goods and incomes vary sufficiently across households. We do try to account for this in our OLS regression in Table 3 by including the measure of income available in the GLSS as a covariate, but this measure is of notoriously poor quality, and even so, may not appropriately control for long or short run income differences across households. Indeed, our income measure is generally not significant at conventional levels in explaining the variation in schooling across households. Finally, a third major reason why we may see a negative relationship between child labor and schooling across households, even though for any *particular* household a different relationship may hold, is that households may vary in their preferences as regards the schooling and labor intensities of the children in the household. If some household prefer low child labor and high schooling relative to other households which prefer high labor and low schooling, then again we would expect to see a negative relationship estimated between the two variables in the cross-section.

5 Identification

Many of these problems might be surmounted if we had data on the same households over time, and so via a standard household fixed analysis look at the fluctuations for a *given* household in the choices of child labor and schooling. Differencing out the purely cross-sectional variation in this tradeoff might go a long way in netting out differences due to taste and permanent income across households. However, a household fixed effects analysis might be a bad idea if, for example, it is largely current, as opposed to permanent or long run differences in income that drive this relationship. In addition, the rather fluid notion of a household containing extended family members and the like may also undermine the purely fixed effects analysis. Thus, as is often the case in other empirical applications (see e.g. Boozer (2000)) the fixed effects analysis, even if we had access to repeated observations on these variables for a given household over time, may actually be a rather bad idea.

We instead consider alternative comparisons based on the cross-household data whereby variations in child labor are plausibly driven by variations in the marginal productivity in child labor. Ideally, child labor contributes little to the overall income of the household so that we can argue these variations in the marginal productivities induce primarily only a substitution effect between child labor and schooling, and not an income effect. As we shall see, this exclusion restriction whereby the 'price' of child labor affects schooling only through its effect on child labor will be somewhat testable as this implicit wage will generate relatively more of an income versus a substitution effect via its permanent versus transitory components in the weather shock.

To set the stage for these two identification strategies, capturing different relative shares of the permanent and transitory aspects of the weather shocks between them, we discuss the data we have on the rainfall patterns across space and time in Ghana. The data come from historical reports from 29 rainfall stations across Ghana. We have about a 50 year time series available for each station, with data covering a time span as far back as 1888 for some stations. Figure 1 shows the monthly variations in rainfall for Northern and Southern Ghana, and it shows the distinct differences in these patterns by the North and South regions. In particular, we see that for the South there are two rainfall seasons during the calendar year, whereas in the North there is just one. In addition, the South tends to get relatively more rainfall on average than the North, and all of these climatic differences lead to different agricultural practices and crops in the North and in the South. By our regional definitions the South of Ghana consists of places like Accra, Cape Coast, Saltpond, and there are two rainfall seasons, the first between about February and July and the second starting from around August through till November. Northern Ghana consists of places like Navorongo, Wa, Gambaga, and here there is just one rainfall season, between April and October.

INSERT FIGURE 1 HERE

The implication of Figure 1 is that to the extent that children are employed in farming labor activities, we can use the variation in child labor intensities across months within a given region as a source of variation in child labor to exploit in estimating the causal relationship between child labor and schooling. The advantage of using purely month by regional variation is that we can then control for idiosyncratic differences that occur in this relationship by netting out a set of pure monthly and pure region dummies. Net of these month and region main effects, we then argue the groups of households sampled across different month by region cells are sufficiently homogeneous as regards our omitted variables of primary concern (i.e. institutions such as vacations, preferences, and household income) to identify a causal impact of child labor on schooling. Furthermore, by investigating the sensitivity of this relationship to the permanent (or long-run) and transitory (or, more precisely the short run - as we make no claim that weather shocks are independent across time) weather shocks, we address the question of how this relationship varies with long and short run income differences for the household.

Recall the behavioral model of the relationship between schooling (S_{it}) and child labor (L_{it}) that we derived in section 3 above:

$$\log(S_{it}) = -\frac{\eta\pi}{\delta}\log(\theta_{it}) + \frac{\pi}{\delta}\log(L_{it}) + \alpha_i + \gamma_t + \epsilon_{it} \tag{12}$$

While we have only a single cross-section of data, the sample has the advantage (for our purposes) of sampling households at different points of time and with the ability to match to a number of rainfall stations across space. So while this behavioral equation cannot be directly estimated given only a single cross-section of data collected at a point in time, we turn next to how our data can allow this model to be estimable if synthetic groupings to time-by-cluster cells can allow us to control for secular time and cluster effects. If the behavioral model aggregates to the cluster level (Deaton (1985)), then this will be a way to control for secular preference/permanent income differences across space as well as institutional factors such as vacations, which may be described by the secular time effects.

To be more specific on the empirical strategy, first note that each household, owing to the sample design, belongs to a month by region cell, indexed as m, r whereby $m \in$ (Jan., Feb., March, April, May, June, July, August, Oct., Nov., Dec.) and $r \in$ (North, South). Using Figure 1, discussed at the beginning of this section as motivation, note that we can define the set of 9 dummies (we use April as the omitted base month throughout, the survey was not conducted in September, and December lacks sufficient observations to be useful) with the notation δ_m as the monthly dummies. Similarly, we can define the dummy ξ_r as the regional dummy equaling 1 if the household is in the North. Finally, let $\lambda_{m,r}$ denote the set of 9 month by region interactions, or cell indicators. As Figure 1 suggests, even if we include the month and region main effects, δ_m and ξ_r in the regression, we can use the month by region indicators $\lambda_{m,r}$ as excluded instruments to pick out the variation in child labor that occurs within month by region cells, net of secular variation purely across months or across regions (and of course, net of variation due to the regressors). This strategy will have power since, as Figure 1 indicates, the rainfall patterns do vary differentially across the months in the North and South, and so even conditional on these main effects, there will exist variation in things like rainfall patterns at the purely month by region level, and so we may expect to see overall variation in child labor intensities at the month by region level.

Using the month and region subscripts to denote the clusters (or cells) of households, consider the individual level regression above, augmented to include secular month (which will go partially towards eradicating the problem of school vacations) and region dummies as main effects. As excluded instruments, we use the month by region indicators $\lambda_{m,r}$, to pick out the variation in child labor across households that is due to the purely month by regional variation (which is plausibly associated with weather variation):

$$S_i = \beta L_i + x_i'\tau + \delta_m + \xi_r + \nu_i \tag{13}$$

(we continue to use the index i to emphasize this is a household level regression, even though each i is associated with an (m, r) cell) with the associated first stage relationship given by:

$$L_i = \pi\lambda_{m,r} + x_i'\kappa + \delta_m + \xi_r + \zeta_i \tag{14}$$

Call the two-stage least squares estimator for β from this regression $\hat{\beta}^{IV}$.

Notice that the two-stage procedure just given in equations (10) and (11) has an interpretation in the form of repeated cross-section or synthetic panel models of the type considered by Deaton (1985). To see this, first ignore the presence of the control variables x_i'. Then the first stage regression is simply the projection onto a set of dummy variables. In particular, the projection of the child labor measures on the month by region dummies, $\lambda_{m,r}$, simply returns the cell means. Thus, purely in terms of estimating coefficients (and not standard errors, etc.), estimating equations (10) and (11) by two-stage least squares is numerically equivalent to grouping the data on schooling and child labor to the 20 month by region cells, and running OLS on that, in the manner of a synthetic panel analysis:

$$S_{m,r} = \beta L_{m,r} + \delta_m + \gamma_r + \omega_{m,r} \tag{15}$$

(Where the subscripts denote the level to which we have taken the sample means of the schooling and labor variables.) Call the estimator derived from this (purely heuristic - we do not actually do this in the paper) synthetic panel procedure $\hat{\beta}^{SP}$.[9] Then, in the absence of covariates it is straightforward to show that $\hat{\beta}^{IV} = \hat{\beta}^{SP}$.

This alternative interpretation of the IV procedure we use in the paper is quite useful since we can plausibly argue that the households within month-by-region cells are plausibly homogeneous across cells *once* we condition on the secular month and region main effects, as regards the unobserved fixed error

[9] Note that this grouped model has only 20 observations on the cell means. Thus, actually proceeding empirically with the synthetic panel approach is infeasible here, even before we start to consider the addition of covariates.

components α_i and γ_t in the behavioral model above.[10] Similarly, to the extent that school vacations vary *purely* at the monthly or regional level, this problem will also be eliminated by the inclusion of the month and region main effects.

We are still left with some alternative sources of bias, however, and to help deal with those, we next consider a more refined version of the instrumental variables strategy based only on the month by region dummies. The IV strategy just discussed utilizes *all* of the variation in child labor that varies across month by region cells. However, we may want to utilize only this variation in child labor correlated with current (i.e. contemporaneous) rainfall patterns. Note that by including the set of month and region main effects, then from a Frisch-Waugh perspective, we can interpret the use of observed rainfall deviated from a set of month and region means, as analogous to a monthly rainfall shock measure.[11] Furthermore, in work we consider below, we also use the current rainfall data together with the month by region interactions $\lambda_{m,r}$ to try to isolate the long and short run variations in the behavioral responses to rainfall, as neither procedure alone completely isolates one or the other of these sources of variation. We also make use of the long-run historical averages in rainfall as a source of variation, with the idea that this may tell us about the relative importance of permanent versus current income (as driven by rainfall patterns) in the household choices between child labor and schooling.

5.1 Identifying the Response to Long and Short Run Household Income Fluctuations

So far, we have succeeded well in eliminating confounding factors such as taste differences across households and institutional factors such as vacations via one

[10]It is entirely possible, for example, households always feel poorer in the early part of the year before the rains come and it is planting season, for example, thus inducing higher child labor and lower schooling for the households interviewed in those months. Similarly, as the North receives less total rainfall and has only one growing season, households from there may also feel secularly poorer. For these and other reasons, this is why it is of key importance that we not exclude the region and month main effects from the second stage, but instead include them to purge the observed data on schooling and child labor of spurious correlations as regards the choices made for a *given* household.

[11]In fact, in an earlier draft of this work, we used a rainfall shock measure that averaged the current rainfall for the current *season* deviated from the long-run seasonal mean observed in our historical rainfall data. The results were largely the same as the ones we present here, but were slightly more imprecise. One major problem with using a constructed shock measure is that we need insights as to how households form expectations as to rainfall patterns given the historical and contemporaneous rainfall data. Clearly if we had better priors on the expectations process, we could construct a more efficient procedure than the one described in the text, which simply uses the monthly variation deviated of its long run mean. However, in the absence of better information as regards this process, the procedure described in the text is perhaps the most robust, as it places little structure on the expectations process.

or both of our identification strategies.[12] However, while we have hinted that our instrumental variables strategies get at the problem of imperfectly observed short and long run income fluctuations for a household, we have until now been purposefully vague as to how we precisely combine our strategies so as to more accurately extract the importance of these two types of income variations for the household. We devote this subsection to this discussing issue so as to give it the focus it needs.

A central question regarding household behavior in Ghana is what role current, as opposed to permanent, income plays a role in decision making. In a typical lifecycle decision framework used to model consumption and labor supply choices in developed economies, access to well-functioning credit markets and overall wealth help justify such smoothing behavior that such models describe. Were this the relevant framework, then even if current income shocks for the household could not be measured accurately, as long as permanent income is appropriately measured, the econometric exercise can proceed. The flip side of this notion may be more appropriate for Ghana, where current income and not permanent income plays a large role in household choices. This becomes a more severe problem for our framework as the weather shock affects not only the marginal productivity of child labor, but the overall productivity of the farm, and thus household income. The critical problem is that it may well be that the weather shock no longer satisfies the exclusion restriction of affecting the schooling decision only by altering the marginal productivity of child labor, but has a *direct* effect as well by changing the level of current income.

We address this point by measuring the difference in the way long run versus short run rainfall patterns affect the estimated impact of child labor on schooling. Simply put, if the permanent income model is the relevant context for Ghana, then cross-household differences in current rainfall should only have explanatory power to the extent they contain signal as to the long run averages. But the problem of omitted short run income in the second stage schooling equa-

[12] Of course, it is not *entirely* possible to exclude the possibility that more exotic versions of these counfounding factors still bias our results. But we can rule out the stylized versions of them. At the end of the day *anything* that varies at a rainfall shock by month-region cell that is correlated with the choice between child labor and schooling and which is not controlled for in our framework we will mistakenly be attributing to weather shocks. For example, it is possible to imagine that even within month by region cells, school vacations are called if the rains or the harvest are unusually good so that children may work in the fields. We would be attributing the resulting negative correlation between schooling hours and child labor as due to a household *choice* as opposed to an institutional feature. However, to the extent the vacation was called as so few households would *choose* to send their children to school in such a situation, then this may be an effect we would want to include rather than throw away through an even more refined identification scheme. A rather strong case can be made that while we may still be guilty of misappropriating *how* these decisions induced by weather shocks are made, the total impact that we measure with our strategies is likely not misleading even with such concerns present. Indeed, were such spontaneous vacations an important facet of what we observe, it would be useful if future research could isolate who makes such vacation decisions and what their objectives and decision factors are.

tion should be limited if long run averages are used as the instrument. Current rainfall deviated from its long run mean, however, should have essentially no explanatory power for child labor or schooling under this null however, analogous to the Euler equation tests of the permanent income model of consumption proposed by Hall (1978). Similarly, if only current income matters, we can use the month by region variation in child labor and *include* the rainfall shock in the first and second stage regressions to examine to what effect failure to adequately control for short run income fluctuations in the schooling equation has on our results. If the responses are much the same with or without including the current rainfall as an additional covariate when we use the month by region indicators as instruments, this would indicate that while current income may play a role in the quantity of schooling directly, it does not appear to impact the *tradeoff* between child labor and schooling that is the focus of our paper. Such results go more fundamentally towards the question of the role of income as regards child labor. One interpretation of the finding of lack of sensitivity of the measured tradeoff between child labor and schooling decisions to alternative income definitions and concepts may be that while income alters the total quantity consumed of either of these two goods, it does not appear to affect the tradeoff. Such would be the case, for example, if the children in poorer households simply worked harder, as defined by total hours spent in the labor market and in school, than children from richer families.

With these concepts in mind, especially as regards the treatment of the effects of alternative income measures, we now briefly discuss the operational aspects of our instrumental variables strategies. Our first strategy, which we denote as Strategy A, uses the month by region interactions as the excluded instruments. All strategies will include, in addition to the covariates, a set of month and region main effects. Thus, Strategy A will use *any* variation across month by region cells (net of the main effects) in child labor, and not that which is solely due to weather shocks. In terms of the long run and short run differences in behavioral responses just discussed, this strategy will utilize both the planned (owing to secular long run differences in rainfall patterns across the month by region cells evident in Figure 1) and the unanticipated response to the short run weather fluctuations.[13] Our second instrumental variables strategy (henceforth Strategy B) refines strategy A by not using *all* of the variation in child labor patterns at the month by region level, but only that variation due to

[13] Clearly this reasoning depends implicitly on the nature of rainfall shocks within cells. For example, if each cell consisted of only one observation on current rainfall, then as far as we can tell, the responses to long run trends and short run shocks are isomorphic. In fact, in Strategy C, which we discuss below, we control for current rainfall while using the Strategy A instruments of month by region interactions to isolate the effect of long run patterns on child labor. In the one observation of rainfall per cell scenario, this model would not be estimable, as the excluded instruments would be collinear with the rainfall variable and so we would have exact collinearity: the different effects of long and short run fluctuations in rainfall would not be estimable. In work near the end of the paper we use the long run rainfall patterns directly to refine this idea further.

current rainfall is used. In fact, owing to the different farming seasons depicted in the long run average rainfall patterns in the North and South of Ghana, we allow the effect of current rainfall to vary by region, and so in Strategy B, the excluded instruments consist of a current rainfall main effect as well as the rainfall by North interaction.

Strategy B, by virtue of including month and region main effects in the first and second stages, *almost* has the interpretation of picking up the responses in child labor to the short run weather fluctuations. This is because the main effects play the role of deviating the current rainfall observations from their month and region means, which proxy for the long run averages. In fact, as we indicated in footnote 9, in an earlier version of this work we used the current rainfall deviated from its long run average as a direct measure of the shock and we obtained the same sets of results as in this version of the paper, the only difference being that the earlier results were slightly less precise. As this version manipulates the observed data even less, but retains the same basic interpretation, we opted to present this cleaner method. In addition, this relates well to our Strategy C, which in some sense combines Strategies A and B, but uses the month by region dummies as the excluded instruments while *including* the current rainfall and its interaction with region in the first and second stages as additional controls.

The idea motivating Strategy C is the concern that current income (or a driving factor such as current rainfall) cannot be credibly excluded from the schooling equation apart from its influence on child labor intensities. If, as a developing economy, credit markets and intertemporal substitution in general are not as possible in Ghana, then in stark contrast to the permanent income model, current income may play a large role in household consumption of all goods including schooling. By using current rainfall as an excluded instrument, we are assuming that either the impact of rainfall on income fluctuations is small or that the permanent income model is the relevant model if the rainfall impacts not only the marginal productivity of the child laborers, but the productivity of the entire household. Strategy C makes use of the (roughly) distinct roles Strategies A and B play to *include* the current rainfall in the schooling equation while continuing to use the cross month by region cell variation in child labor to identify the labor-schooling tradeoff.

Finally, another specification check we perform which pertains to the maintained hypotheses of the model outlined in Section 3 concerns the impact of the rainfall shock on the consumption side of the household decision process. Strategies A, B, and C deal with the precise way in which the rainfall shock enters the production side of the model, and concerns investigating the sensitivity of the results to the possible failure of the exclusion restriction regarding the rainfall shock. The model in Section 3 presumed the rainfall shock would produce a pure substitution effect and a relatively small income effect if the share of household (implicit) income due to child labor was small. However, rainfall might also play a role on the consumption side if rainfall (and its accompanying hot weather) makes travel to school, and schooling itself, too difficult. This again would produce a spurious negative correlation between child labor and schooling that is not of the type we are trying to measure. In an effort to gauge the magnitude of this confounding effect, we simply look at the correlation of the total time devoted to both activities with the rainfall shock.

6 Empirical Results

6.1 The Marginal Effect of Child Labor on Schooling

Given the discussion in the previous section which outlines our identification strategies, as well as several robustness checks, we now turn to the estimation of the marginal effect of child labor induced via our instruments.

We start first by presenting the cross-sectional OLS results which will capture not only the substitution effect we are interested in isolating, but also the institutional and preference factors which may induce a spurious negative correlation between these two variables, as well as the plethora of confounding income effects. Column A in Table 3 shows a simple OLS regression of schooling on child labor and a number of controls. We control for age of the child, sex, demographic composition of the household, religion, language, whether the child is the oldest in the family, household income and a dummy for whether the household is in a rural or an urban area. There is a strong negative coefficient on child labor, as expected - at present, we make no claims on causality. Note that income has been included as a control, even though the income measure is extremely noisy. Ideally, we would like to control for *permanent* income, but are unable to measure this. Column B of Table 3 then looks at the question of an income effect and shows the OLS results if we exclude income as a control, with no significant changes in the results. We come back to this issue of the income effect in more detail later.

There is, at a first level, heterogeneity in the monthly effects, by region, as well as possible heterogeneity by gender. We come back to the issue of heterogeneity by gender later and first look at the heterogeneity in monthly effects by *region*. If we fail to account for this heterogeneity by region by month, and simply construct an instrumentation strategy, using the rainfall measures we have[14], then, as Column C in Table 3 shows, we do not find any significant

[14]The excluded instrument in this case is the monthly realized rainfall, deviated from its month by region effects. The results here do not change much when the constructed seasonal shock is used as the excluded instrument. This seasonal shock was constructed by look at the rainy season in each region (north and south), i.e. two such seasons for Southern Ghana and one for Northern. For each household, then, given the date it was surveyed in, we calculated the total seasonal rainfall it had realized so far. As a measure of the shock we then deviated this from the long term mean of what this particular household would have expected for the season so far.

effects of child labor on schooling. This instrumented effect of child labor on schooling, when not accounting for heterogeneity that may be present, is not significantly different from zero.

If we then think more carefully about identification, giving due course to the heterogeneity in month by region effects, we can implement the identification strategies outlined in Section 5. The first of these, Strategy A, was to use just the *month by region* interactions to capture the effects of different climate driving the need for varying child labor intensities. Column A in Table 4 gives the first stage results, i.e. using month by region interactions to explain the variation in child labor, using the same controls as in the regressions in Table 3. The first stage has immense power, with a large number of the month by region interactions being very significant, especially towards later in the year. The F test of joint significance of the excluded instruments supports this and is also reported in Table 4. Table 5 then shows the IV (second stage) results: we find a negative significant effect of child labor on schooling. It is worth noting that this negative effect of child labor on schooling that we find here is extremely similar to the OLS results from Table 3. This, together with the other results in Table 3 and in light of the theoretical model, seem to point to the income effect not being of large importance.

We now look at Strategy B, where the observed rainfall, by region are used as the excluded instruments, i.e. we use the realized rainfall (and not a measure of the shock the households face) and include month by region interactions *and* main effects in both stages of the IV estimation. Column B in Table 4 shows the first stage results, which again are highly significant, and Table 5 the IV results. We find very similar results to the first strategy, strongly negative coefficients on child labor, but this time with a slightly larger coefficient.

INSERT TABLES 4 & 5 HERE

The distinction between what is used in the two different instrumentation strategies is that the first strategy allows for *permanent* differences in child labor intensities, and therefore has more of a long-term interpretation. It is important to remember that this instrumentation strategy in fact captures both the long-term as well as short-term effects and anything else that may vary by month by region. However, when using the rainfall by region to instrument, we use only the short-term by controlling for the month by region effects (i.e. any anticipated effects). The first strategy therefore speaks to a more long- term effect and is comparable to the second as long as there are no secular differences in permanent income by month by region cells. Within this interpretation, we can then in essence isolate a long-term impact, by using the same excluded instruments as Strategy A, but also control for the short-term effects, i.e. rainfall by region, in both stages for the procedure. We report these results as Strategy C, with the first stage shown in Table 4[15] and the second stage in Table 5.

[15] Note that this first stage is the same as the first stage for Strategy B, except that Strategy

Again, we see a highly significant negative impact of child labor on schooling, and an effect that is comparable in magnitude to the other two strategies.

6.2 Extensive vs. Intensive Margins and Gender Heterogeneity

There are two relevant *margins* of effects here - an extensive margin, i.e. the effect of a child switching in and out of the labor market and an intensive margin, i.e. an effect of hours worked, *conditional* on a child already being in the labor market. This is where the interplay of gender effects becomes more evident. The gender breakdowns are interesting partly because boys are more likely to work in the field, and, more importantly, when both boys and girls work in the field, they tend to different types of crops[16]. We may, therefore, well expect to see differential effects by gender of the rainfall and even differential effects on these two margins.

We analyze these margins separately, looking at the extensive margin by considering the child labor variable as a 0-1 decision, i.e. a simple description of whether the child is in the labor market or not and looking at the intensive margin by looking only at those children *already* in the labor market. Table 6 shows the results of analyzing child labor at these two margins, using all three instrumentation strategies. We can see, in all three cases, large effects at the extensive margin in both the long and short term, i.e. children switching in and out of the labor market has big impacts on the schooling of these children. At the intensive margin, there are also effects, mostly of a long-term nature, given the interpretation of our strategies.

INSERT TABLE 6 HERE

We then go on to look at gender heterogeneity, first across the full sample of children and leaving aside the extensive and intensive margins and then we specifically look at the gender heterogeneity at each of these margins. Table 7 shows these results, again for each of our instrumentation strategies. The results support the hypothesis that there may indeed be differential effects by gender, and perhaps more so in the short-term, as illustrated by Strategy B, where the effect we are interested in for girls is not significantly different from zero.

INSERT TABLE 7 HERE

B uses *only rainfall by region* effects as the excluded instruments and Strategy C uses *only month by region effects* as the excluded instruments

[16]This raises the fixed-cost and/or minimum scale issues of boy girl-labor, which we will, at present, abstract from.

We now look at whether there is any evidence of differential effects by gender across the extensive and intensive margins. Table 8 shows the results at the extensive margin across genders. We can see how all three strategies show strong effects on boys - there are big impacts of child labor on schooling in both the short-term as well as the long-term. These effects are also existent for girls, much stronger in the long-term and not as evident in the short-term. Again, both the short-term and long-term coefficients are very similar.

INSERT TABLE 8 HERE

At the intensive margin, things begin to look even more different, as seen in Table 9. Once again, strong negative effects are evident for boys at this margin - working extra hours for boys has big impacts on their schooling. However, for girls, there is no evidence of a significant effect, let alone a negative impact. The evidence seems to point to the impact of child labor on schooling, once children are already in the labor market, being extremely small for girls but quite large for boys.

INSERT TABLE 9 HERE

6.3 The Schooling Consumption Effect

In light of the theoretical model presented in Section 3, we have so far shown strong evidence of an impact that child labor has on schooling. Any effect of child labor on schooling, as we mentioned earlier, could be one of three effects: a substitution effect, an income effect and a direct consumption of schooling effect. Up to now, we have shown how this effect we identify in each of our strategies does not seem to be an income effect. We can also show how we manage to abstract from a simple consumption of schooling explanation of our results. We do this by looking at the effects on the total time allocation of children across our sample. By total time allocation, we mean the total time children have accounted for in terms of labor and schooling hours and so, what is left out of this measure of total time allocated is house work and/or leisure. Table 10 shows the effects of rainfall by region on total time allocation defined in this way. Table 10 also breaks out these effects by gender. As we can see, the impacts of rainfall and rainfall by region on total time allocation are not significantly different from zero for both the entire sample as well as by gender, supporting the hypothesis that what we pick in our instrumentation is *indeed* a substitution effect between schooling and child labor.

INSERT TABLE 10 HERE

6.4 The Income Effects and the Relevance of Poverty

We now try to tackle the question of whether poverty plays a role in explaining differential child labor effects. We want to try and analyze whether child labor is indeed a phenomenon associated with purely *poorer* households and whether this tradeoff between child labor and schooling is stronger for households with lower income. We have seen so far how there is no evidence that the effect we find is indeed largely an income effect.

Intuitively, already, our results speak to this idea of a poverty hypothesis. As we can see from Tables 6 on, our long run and short run instrumentation strategies look extremely similar in terms of coefficients on the tradeoff between child labor and schooling. Girls seem to be insulated from any short run effects while boys are affected at both the short run as well as the long run. If households are indeed constrained, then we would expect the short run to be a lot more important, relative to the long run, contrary to our results.

However, we can actually analyze this relevance of poverty or income directly. We can check whether there is any evidence of heterogeneity in these effects we find by income group. We have an extremely noisy measure of income and so, we split our sample into three 'income groups' - low, medium and high and that way we can compare the two ends of the distribution, as such. This broad grouping should be more believable, given that income is noisy and we try and abstract from that by just looking across these low and high income groups. We define the low income group as consisting of households with annual incomes less than 135,000 cedis and the high income group as those with annual incomes greater than 420,000 cedis. Those with incomes between these two figures are in the middle income group. We only compare the high and low income groups. Each of these income groups comprises about a third of our sample in that about a third of the children in our sample come from households that belong to each of these income groups. Note that the mean household income in our sample is about 564,000 cedis per annum (with a standard deviation of about 994,000 cedis).

Table 11 shows the results - it looks across the full sample in each income group, ignoring the extensive vs. intensive margins and any gender heterogeneity that may be present - it is therefore comparable to Table 5. Similar to previous analyses, we use all three instrumentation strategies. We find very similar results, in terms of coefficients across the two income groups, also similar to the pooled results in Table 5. The standard errors are noticeably larger, hence the effects may seem insignificant, but the coefficients are very across income groups are similar to each other and to what we found earlier. There is no evidence, therefore, of heterogeneity across these two income groups, and no evidence that the effects are larger if the child belongs to a household in the low income group.

INSERT TABLE 11 HERE

7 Conclusion

This paper aims to tackle the question of whether a tradeoff between child labor and schooling does exist, in the context of an exogenous shock to the productivity of the children in question. We draw on a theoretical framework to describe the households decisions *vis a vis* child labor and schooling, in which there exists a preference heterogeneity across groups of households that would bias any simple reduced form analysis. This necessitates our empirical approach of attempting to identify variation in child labor across groups of households that will speak to the effects we are interested in. For the purposes of this paper, we are interested in only the marginal effect of child labor on schooling, and in analyzing this show how the empirical methodology speaks to just this effect and not other confounding influences.

We find that, accounting for the heterogeneity in effects by region (given the different climatic patterns in Northern and Southern Ghana) by month, that there is a significant tradeoff between child labor and schooling. This tradeoff is more obvious for boys than for girls, though it does exist for girls in the long run as well as at the extensive margin. We also show how this effect should be mostly a substitution effect by looking, in turn, at the other two confounding hypotheses. We find no evidence of an income effect and by looking at the effects on total time allocation for the children, we find no evidence that these are direct impacts on schooling consumption as such.

This then begs the question of what, if any, is the policy relevance of these findings. The gender heterogeneity in child labor effects or impacts has been widely documented, for example see Canagarajah and Coulombe (1997) and Bhalotra (1999). At present, a lot of governments, NGO's and other external agencies are interested in policy measures to help alleviate poverty, believing this to be the main reason for the existence of child labor. In doing so, they hope to discourage households from engaging in child labor, for example by giving them subsidies to go to school. An example of this is the Food for Education Program in Bangladesh, the program discussed in Ravallion and Wodon (1999). The other direction policy could take would be direct public spending on schools to improve their quality and relevance in households' decision making processes, for example the Back to School Program in Indonesia, as discussed in Sayed (2000). We find little evidence of poverty playing a role in the effects of child labor and so, for Ghana, policy aimed at alleviating poverty may not be appropriate. However, it is also important to remember how context and economy specific these analyses are and so translating them into any worldwide policy is not appropriate.

Note that we have abstracted from making any specific welfare judgments for the children themselves. We make no claims as to whether the children are worse off by being in the labor market. The aim here is to analyze household decision making processes *vis a vis* child labor and schooling, and until we fully do so, it seems impossible to make any welfare judgments. Similarly, we aim to identify *what* drives this tradeoff as opposed to *why* households make these tradeoffs. We leave that to future research projects, trying to look at whether the existence of these tradeoffs are a result of a lack of functioning adult labor markets or due to a lack of functioning credit markets or simply a lack of high quality schooling in these areas.

References

[1] Akabayashi, Hideo, and George Psacharopoulos, (1999), 'The Trade- off Between Child Labor and Human Capital Formation: A Tanzanian Case Study', *Journal of Development Studies*, 35, June 1999, pp. 120-140.

[2] Altonji, J. G., (1983), 'Intertemporal Substitution in Labor Supply: Evidence from Micro-Data', *Journal of Political Economy.*

[3] Angrist, Joshua D., (2000), 'Estimation of Limited-Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice', *NBER Technical Working Paper*, No 248.

[4] Baland and Robinson, (2000), 'Is Child Labor Inefficient?', *Journal of Political Economy*, August 2000, pp. 663-679.

[5] Bardhan, P., and C. Udry, (1999), *Development Microeconomics*, Oxford University Press.

[6] Basu, Kaushik, (1999), 'Child Labor: Cause, Consequence and Cure, with Remarks on International Labor Standards', *Journal of Economic Literature*, 37, September 1999, pp. 1083-1119.

[7] Basu, Kaushik, and Pham Hoang Van, (1998), 'The Economics of Child Labor', *American Economic Review*, 88, 1998, pp. 412-427.

[8] Bhalotra, S., (2000), 'Is Child Work Necessary?', *STICERD Working Paper*, DEDPS No. 26.

[9] Bhalotra, S., and C. Heady, (1999), 'Child Farm Labor: Theory and Evidence', *STICERD Working Paper*, DEDPS No. 24.

[10] Boozer, Michael A., (2000), 'The Relative Value of Within and Between Variation Using Data With a Group Structure', working paper, Department of Economics, Yale University.

[11] Canagarajah and Coulombe, (1997), 'Child Labor and Schooling in Ghana', Part of the series undertaken as part of a World Bank Economic and Sector Work (ESW) on *Ghana: Labor Markets and Poverty.*

[12] Deaton, Angus, (1985), 'Panel Data From a Time-Series of Cross- Sections', *Journal of Econometrics*, 30, pp. 109-126.

[13] Glewwe, Paul, and Hanan Jacoby, (1995), 'An Economic Analysis of Delayed Primary School Enrollment in a Low Income Country: The Role of Early Childhood Nutrition', *Review of Economics and Statistics*, 1995, pp. 156- 169.

[14] Grootaert, Christian, and Ravi Kanbur, (1995), 'Child Labor: An Economic Perspective', *International Labor Review*, Vol. 134, 1995, pp. 187-203.

[15] Grosh, M. and Paul Glewwe, (1995), 'A Guide to Living Standards Measurement Surveys and Their Data Sets', *LSMS Working Paper*, No 120, The World Bank.

[16] Hall, Robert E., (1978), 'Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence', *Journal of Political Economy*, 96, pp. 971-987.

[17] Heady, Christopher, (2000), 'What is the Effect of Child Labor on Learning Achievement? Evidence from Ghana', UNICEF, Innocenti Working Paper No. 79.

[18] International Labor Organization, (1996), 'Child Labor: What is to be Done', ILO Homepage.

[19] Jacoby, Hanan and Emmanuel Skoufias, (1997), 'Risk, Financial Markets, and Human Capital in a Developing Country', *Review of Economic Studies*, 64, 1997, pp. 311-335.

[20] Jensen, Peter, and Helena Skyt Nielsen, (1997), 'Child Labor or School Attendance? Evidence from Zambia', *Journal of Population Economics*, 10, 1997, pp. 407-424.

[21] Moehling, Carolyn, (1997), 'The Impact of Children's Income on Intrahousehold Resource Allocations', Yale University.

[22] Moehling, Carolyn, (1999), 'State Child Labor Laws and the Decline of Child Labor', *Explorations in Economic History*, 36, 1999, pp. 72-106.

[23] Patrinos, Harry Anthony, and George Psacharopoulos, (1997), 'Family Size, Schooling and Child Labor in Peru: An Empirical Analysis', *Journal of Population Economics*, 10, 1997, pp. 387-405.

[24] Patrinos, Harry Anthony, and George Psacharopoulos, (1995), 'Educational Performance and Child Labor in Paraguay', *International Journal of Educational Development*, 15(1), pp. 47-60.

[25] Psacharopoulos, George, (1997), 'Child Labor versus Educational Attainment: Some Evidence from Latin America', *Journal of Population Economics*, 10, 1997, pp. 377-386.

[26] Ravallion, Martin, and Quentin Wodon, (1999), 'Does Child Labor Displace Schooling? Evidence on Behavioral Responses to an Enrollment Subsidy', World Bank Working Paper.

[27] Ray, Ranjan, (2000), 'The Determinants of Child Labor and Child Schooling in Ghana', University of Tasmania.

[28] Ray, Ranjan, (2000), 'How Child Labor and Child Schooling Interact with Adult Labor', University of Tasmania.

[29] Sayed, Haneen, (2000), 'Indonesia: Back to School Program', Paper presented at Conference on Child Labor and Child Protection, The World Bank.

[30] United Nations Educational, Scientific and Cultural Organization.

[31] Vijverberg, Wim P. M. (****), 'Educational Investments and Returns for Women and Men in Cote d"Ivoire', *Journal of Human Resources*, 1993, 28, pp. 933-974.

Figure 1
Geographical Rainfall Patterns, Northern and Southern Ghana

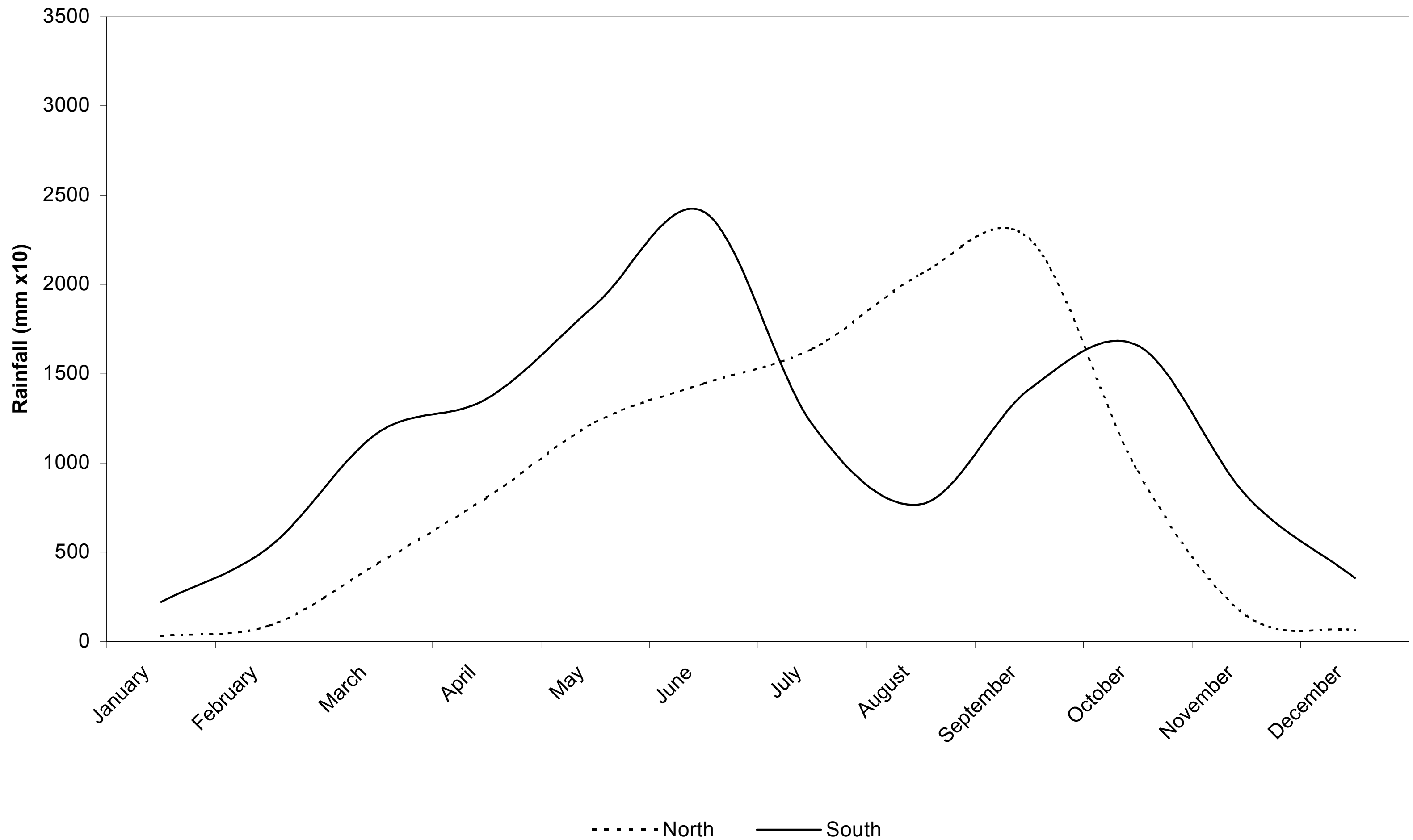

Figure 2
Current Year Rainfall, Northern and Southern Ghana

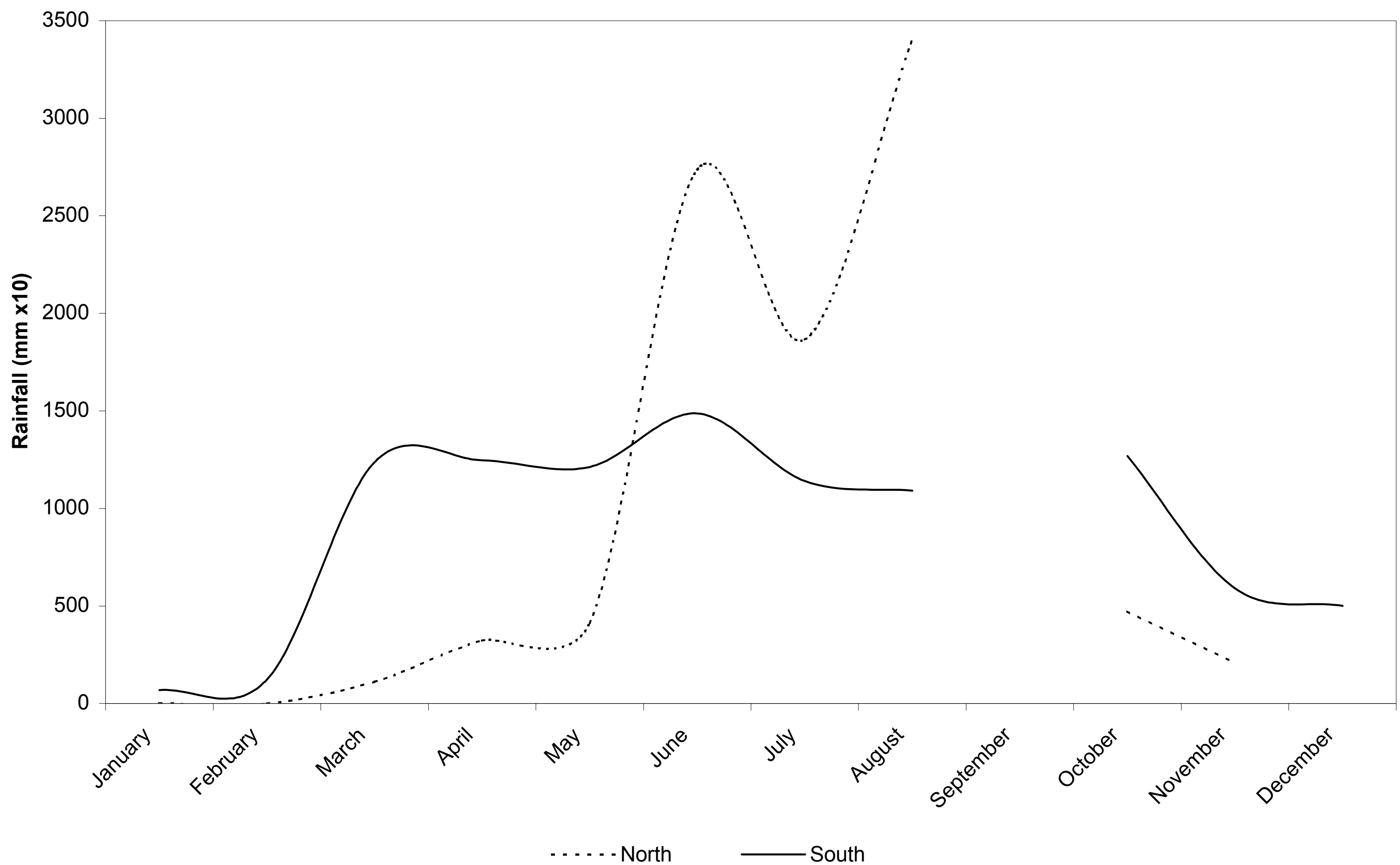

Figure 3
Long-term Rainfall, Northern and Southern Ghana
(50 Year Average)

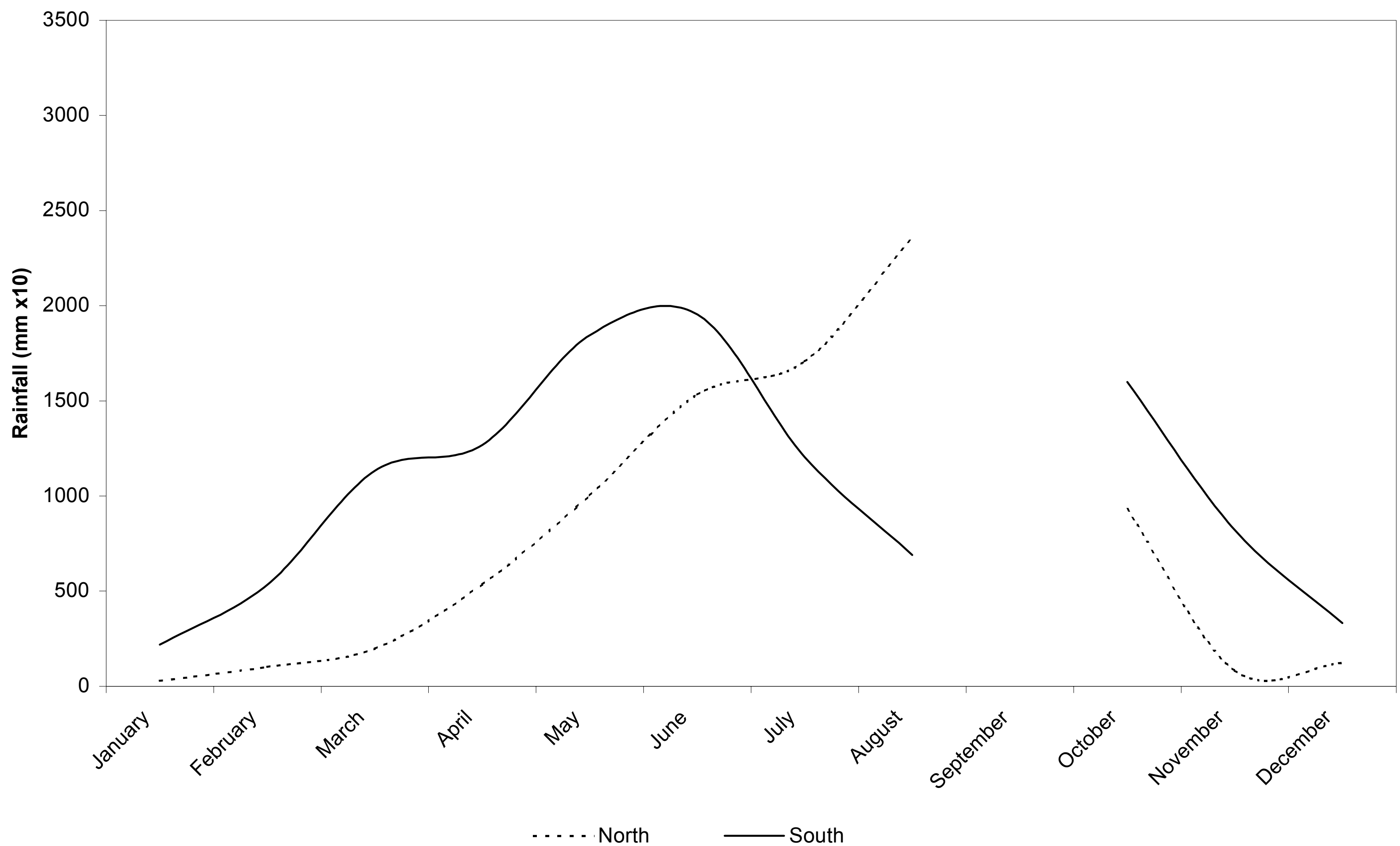

Figure 4
Rainfall Shock (Monthly), Northern and Southern Ghana

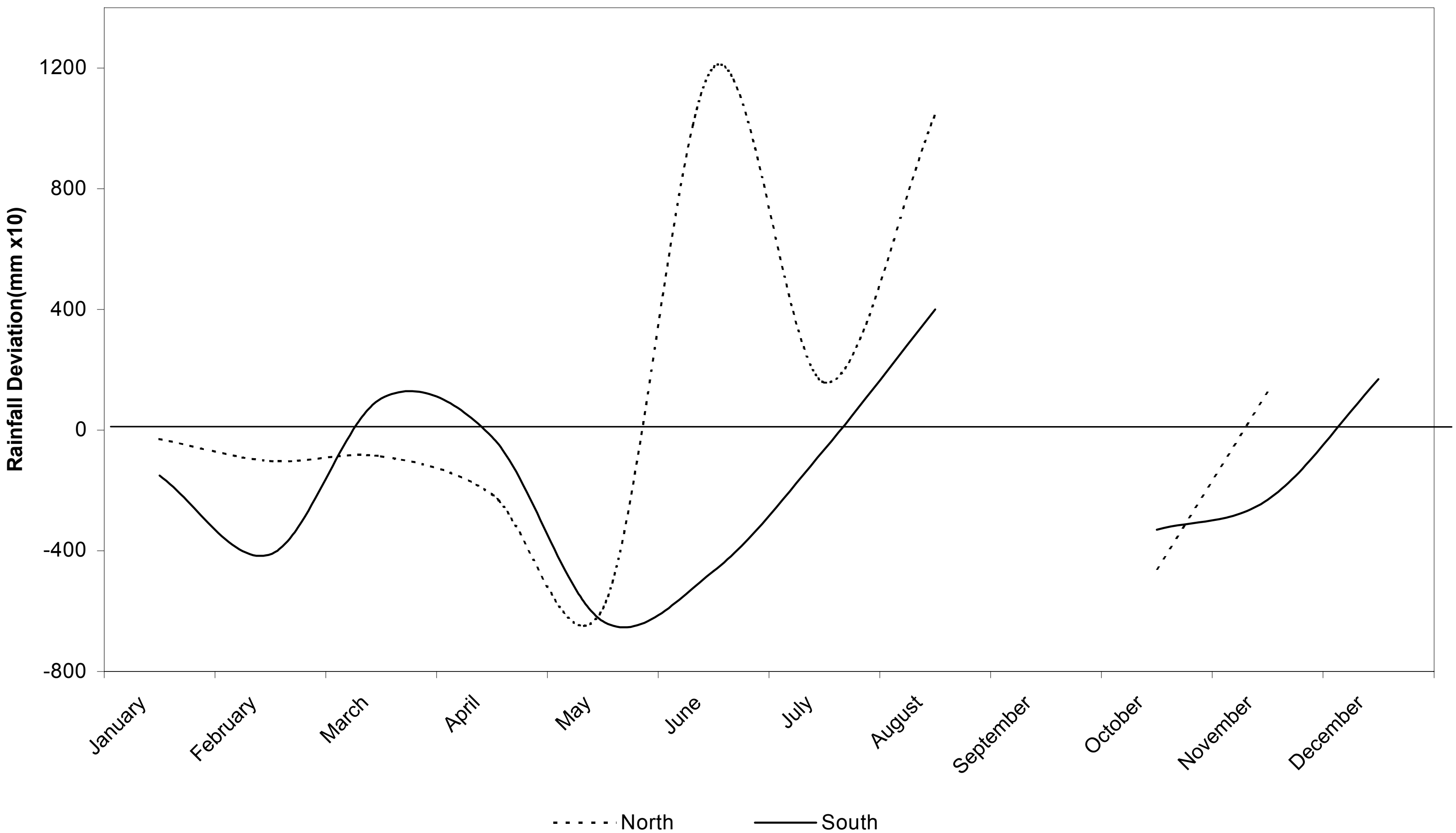

Table 1
Participation Rates of Children, Aged 7-18, in Various Activities

Sample	Schooling	Labor (main & secondary jobs)	House Work
Full sample	52%	31%	86%
Girls	46%	30%	89%
Boys	57%	33%	80%
North	28%	38%	66%
South	57%	29%	88%

Table 2
Average Hours Spent in Each Activity, Conditional on Participation

Sample	Schooling	Labor (main & secondary jobs)	House Work
Full sample	21.9	19.6	12.9
	(6.2)	(17.1)	(8.5)
Girls	22.0	19.8	15.2
	(6.1)	(16.5)	(9.4)
Boys	21.8	19.5	10.5
	(6.3)	(17.6)	(6.8)
North	24.3	30.9	11.0
	(5.3)	(18.2)	(9.2)
South	21.6	16.3	13.2
	(6.3)	(15.3)	(8.4)

Notes: Standard deviations are in parentheses.

Table 3
The Effect of Child Labor on Schooling

Explanatory Variable	A OLS	B OLS	C IV
Child labor (hours)	-0.212 (0.010)	-0.213 (0.010)	-0.174 (0.270)
Income ($x10^6$)	-0.272 (0.193)		-0.327 (0.216)
Age	-0.217 (0.058)	-0.221 (0.058)	-0.286 (0.337)
Sex (female)	-2.852 (0.379)	-2.859 (0.379)	-2.622 (0.441)
Urban	0.204 (0.429)	0.266 (0.428)	-0.048 (0.866)
Region (north)	-1.386 (0.641)	-1.341 (0.641)	-2.033 (1.250)
Oldest	-0.722 (0.384)	-0.659 (0.382)	-0.671 (0.394)
Fraction of household that is male children	-0.943 (1.395)	-0.747 (1.390)	-1.240 (1.434)
Fraction of household that is female children	0.145 (1.421)	0.323 (1.419)	-0.294 (1.464)
R-squared No. of observations	0.160 4499	0.159 4499	0.160 4313
F test P value	-	-	$F_{1,\,4284} = 7.20$ 0.007

Notes: Robust standard errors are in parentheses.
Excluded instrument (see F test) in Column C: realized rainfall deviated from month and region effects. Results are similar when using the constructed seasonal shock.
All regressions henceforth also control for occupation of the household head, religion of the household and language.

Table 4
IV Regressions: First Stage Results: Effects on Child Labor

Explanatory Variable	Strategy A Month by Region		Strategy B Rainfall by Region		Strategy C	
January * North	***5.099***	***(2.413)***	2.255	(2.631)	***2.255***	***(2.631)***
February * North	***1.035***	***(2.422)***	-2.422	(2.696)	***-2.422***	***(2.696)***
March * North	***2.771***	***(2.705)***	1.019	(2.774)	***1.019***	***(2.774)***
May * North	***8.376***	***(2.969)***	8.957	(2.974)	***8.957***	***(2.974)***
June * North	***2.099***	***(2.807)***	22.46	(8.053)	***22.46***	***(8.053)***
July * North	***8.492***	***(2.596)***	21.11	(5.465)	***21.11***	***(5.465)***
August * North	***19.09***	***(2.489)***	44.94	(10.11)	***44.94***	***(10.11)***
October * North	***13.63***	***(2.309)***	19.03	(2.470)	***19.03***	***(2.470)***
November * North	***12.23***	***(2.374)***	10.76	(2.431)	***10.76***	***(2.431)***
December * North	***12.18***	***(2.706)***	-	-	-	-
North	-4.248	(2.069)	-0.551	(2.420)	-0.551	(2.420)
Rain ($x10^3$)	-	-	***0.537***	***(0.480)***	0.537	(0.480)
Rain * North ($x10^3$)	-	-	***-8.853***	***(3.172)***	-8.853	(3.172)
R-squared	0.210		0.210		0.210	
No. of observations	4499		4313		4313	
F test on excluded instruments	$F_{10,\ 4451} = 10.32$		$F_{2,\ 4264} = 2.00$		$F_{9,\ 4264} = 9.75$	
P value	0.000		0.135		0.000	

Notes: Robust standard errors are in parentheses.
Excluded instruments: (Emphasized above)
Strategy A: only month by region interactions.
Strategy B: rainfall by region & rainfall main effect.
Strategy C: only month by region interactions & rainfall effects in both first and second stages.

Table 5
IV Regressions: Second Stage Results
Effects of Child Labor on Schooling

Explanatory Variable	Strategy A Month by Region		Strategy B Rainfall by Region		Strategy C	
Child labor	-0.317	(0.071)	-0.843	(0.471)	-0.380	(0.073)
Rainfall ($x10^3$)	-	-	-	-	-0.325	(0.381)
Rain * North ($x10^3$)	-	-	-	-	1.194	(0.494)
Income ($x10^6$)	0.042	(0.164)	0.270	(0.289)	0.055	(0.167)
Age	-0.052	(0.104)	0.578	(0.584)	0.010	(0.104)
Sex (female)	-2.938	(0.349)	-3.179	(0.591)	-2.775	(0.359)
Urban	-0.496	(0.449)	0.579	(1.114)	-0.311	(0.458)
Region (north)	-1.896	(0.705)	1.547	(3.484)	-2.948	(0.807)
Oldest	-0.881	(0.344)	-0.775	(0.436)	-0.865	(0.354)
Fraction of household that is male children	-0.604	(1.305)	-0.970	(1.844)	-0.910	(1.347)
Fraction of household that is female children	0.664	(1.330)	1.089	(1.693)	0.252	(1.366)
R-squared	0.302		-		0.287	
No. of observations	4499		4313		4313	

Notes: Robust standard errors are in parentheses.
Excluded instruments: Strategy A: only month by region interactions.
Strategy B: rainfall by region & rainfall main effect.
Strategy C: only month by region interactions & rainfall effects in both first and second stages.

Table 6
Effects of Child Labor on Schooling: Extensive vs. Intensive Margins

Explanatory Variable	Strategy A		Strategy B		Strategy C	
	Extensive Margin	Intensive Margin	Extensive Margin	Intensive Margin	Extensive Margin	Intensive Margin
Child labor	-9.820	-0.147	-9.942	-0.243	-18.74	-0.164
	(2.373)	(0.079)	(5.250)	(0.361)	(3.357)	(0.074)
Rainfall	-	-	-	-	0.677	-0.041
($x10^3$)					(0.489)	(0.630)
Rainfall * North	-	-	-	-	1.53	0.168
($x10^3$)					(0.584)	(0.707)
Income	-0.054	0.216	-0.034	0.743	-0.009	0.235
($x10^6$)	(0.161)	(0.400)	(0.166)	(0.993)	(0.179)	(0.398)
Age	0.024	0.415	-0.002	-0.320	0.410	-0.401
	(0.129)	(1.438)	(0.250)	(0.536)	(0.169)	(0.142)
Sex (female)	-2.959	-4.003	-2.700	-3.583	-2.925	-3.644
	(0.370)	(0.570)	(0.390)	(0.760)	(0.438)	(0.591)
Urban	0.311	-0.602	0.297	-0.770	1.540	-0.717
	(0.578)	(0.814)	(0.883)	(1.481)	(0.705)	(0.794)
Region (north)	-3.325	-2.787	1.264	9.882	-5.228	-3.219
	(0.630)	(1.560)	(2.389)	(8.213)	(0.882)	(1.692)
Oldest	-1.016	-0.481	-0.896	-0.276	-0.925	-0.388
	(0.365)	(0.552)	(0.371)	(0.609)	(0.428)	(0.561)
Fraction of household	-0.799	-0.005	-0.559	1.317	-1.708	0.583
that is male children	(1.376)	(2.032)	(1.480)	(2.082)	(1.606)	(2.037)
Fraction of household	-0.266	2.951	-0.250	4.156	-1.408	3.265
that is female children	(1.400)	(2.142)	(1.503)	(2.602)	(1.639)	(2.163)
R-squared	0.222	0.401	0.230	0.421	-	0.399
No. of observations	4499	1375	4313	1294	4313	1294
F test (first stage)	11.63	4.62	11.03	1.07	7.73	5.54
P value	0.000	0.000	0.000	0.344	0.000	0.000

Table 7
Effects of Child Labor on Schooling: Gender Heterogeneity

Explanatory Variable	Strategy A		Strategy B		Strategy C	
	Boys	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**
Child labor	-0.184	-0.338	-0.724	0.133	-0.229	-0.476
	(0.069)	(0.094)	(0.364)	(1.498)	(0.061)	(0.119)
Rainfall	-	-	-	-	-0.432	-0.194
($x10^3$)					(0.517)	(0.585)
Rainfall * North	-	-	-	-	1.341	0.562
($x10^3$)					(0.657)	(0.726)
Income	-0.152	0.207	0.164	-0.067	-0.138	0.252
($x10^6$)	(0.227)	(0.241)	(0.355)	(0.692)	(0.228)	(0.258)
Age	-0.091	-0.160	0.604	-0.744	-0.042	-0.021
	(0.117)	(0.137)	(0.486)	(1.771)	(0.109)	(0.158)
Urban	-1.177	-0.280	0.103	-0.651	-1.060	0.0003
	(0.591)	(0.615)	(1.170)	(2.026)	(0.588)	(0.649)
Region (north)	-2.307	-2.187	8.702	6.104	-3.535	-2.743
	(0.969)	(0.898)	(5.198)	(16.16)	(1.078)	(1.164)
Oldest	-0.571	-1.374	-0.621	-1.294	-0.567	-1.320
	(0.473)	(0.503)	(0.560)	(0.550)	(0.478)	(0.532)
Fraction of household	-2.147	0.313	-0.777	3.097	-2.495	-0.029
that is male children	(1.763)	(1.911)	(2.358)	(7.081)	(1.790)	(2.037)
Fraction of household	-2.442	4.103	-2.654	3.109	-2.981	4.350
that is female children	(1.745)	(2.054)	(2.067)	(4.428)	(1.771)	(2.173)
R-squared	0.340	0.279	0.107	0.224	0.342	0.231
No. of observations	2366	2133	2265	2048	2265	2048
F test on excluded	10.11	6.06	2.46	0.19	9.88	5.09
instruments (first stage)						
P value	0.000	0.000	0.085	0.827	0.000	0.000

Table 8
Effects of Child Labor on Schooling
Gender Heterogeneity & Extensive Margin Effects

Explanatory Variable	**Strategy A**		**Strategy B**		**Strategy C**	
	Boys	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**
Child labor	-5.164	-12.38	-9.180	-12.25	-13.27	-16.88
	(2.821)	(3.184)	(5.784)	(12.51)	(3.375)	(4.453)
Rainfall	-	-	-	-	0.537	0.235
($x10^3$)					(0.613)	(0.673)
Rainfall * North	-	-	-	-	1.735	0.645
($x10^3$)					(0.769)	(0.785)
Income	-0.235	0.129	-0.163	0.039	-0.209	0.136
($x10^6$)	(0.232)	(0.233)	(0.247)	(0.238)	(0.246)	(0.246)
Age	-0.085	0.025	0.077	-0.020	0.271	0.198
	(0.158)	(0.177)	(0.277)	(0.587)	(0.191)	(0.224)
Urban	-0.852	0.886	-0.329	1.064	0.383	1.453
	(0.768)	(0.757)	(1.115)	(1.684)	(0.895)	(0.866)
Region (north)	-3.292	-3.384	2.842	-0.713	-5.253	-4.506
	(0.885)	(0.908)	(3.006)	(5.600)	(1.121)	(1.282)
Oldest	-0.616	-1.483	-0.550	-1.275	-0.652	-1.309
	(0.491)	(0.551)	(0.510)	(0.567)	(0.532)	(0.615)
Fraction of household	-2.642	0.918	-2.748	1.316	-3.409	0.608
that is male children	(1.829)	(2.043)	(1.937)	(2.461)	(2.012)	(2.261)
Fraction of household	-2.980	3.288	-3.369	3.414	-4.547	3.125
that is female children	(1.829)	(2.166)	(2.054)	(2.224)	(2.035)	(2.369)
R-squared	0.293	0.168	0.260	0.183	0.170	0.047
No. of observations	2366	2133	2265	2048	2265	2048
F test on excluded instruments (first stage)	11.09	4.52	10.25	1.95	6.86	3.43
P value	0.000	0.000	0.000	0.142	0.000	0.000

Table 9
Effects of Child Labor on Schooling
Gender Heterogeneity & Intensive Margin Effects

Explanatory Variable	Strategy A		Strategy B		Strategy C	
	Boys	Girls	Boys	Girls	Boys	Girls
Child labor	-0.287 (0.078)	0.125 (0.079)	-0.222 (0.251)	0.123 (0.419)	-0.311 (0.074)	0.105 (0.090)
Rainfall ($x10^3$)	-	-	-	-	-1.405 (0.989)	0.460 (1.011)
Rainfall * North ($x10^3$)	-	-	-	-	2.131 (1.048)	-1.556 (1.144)
Income ($x10^6$)	0.501 (0.521)	-0.279 (0.622)	0.794 (0.653)	-0.080 (1.70)	0.566 (0.531)	-0.187 (0.647)
Age	-0.111 (0.190)	-0.871 (0.190)	-0.202 (0.457)	-0.971 (0.552)	-0.015 (0.186)	-0.928 (0.203)
Urban	-1.912 (1.171)	1.398 (1.271)	-1.584 (1.379)	1.429 (1.868)	-2.016 (1.173)	1.324 (1.289)
Region (north)	-0.844 (2.105)	-4.348 (1.638)	9.218 (7.350)	1.627 (3.221)	-3.284 (2.277)	-2.937 (2.380)
Oldest	0.212 (0.745)	-1.710 (0.956)	0.475 (0.940)	-1.458 (1.043)	0.154 (0.771)	-1.419 (0.977)
Fraction of household that is male children	0.519 (2.825)	3.6 (3.450)	1.906 (4.001)	4.097 (4.973)	0.648 (2.900)	4.321 (3.478)
Fraction of household that is female children	0.351 (2.778)	5.867 (3.827)	2.361 (2.927)	6.030 (4.768)	0.662 (2.890)	6.163 (3.855)
R-squared No. of observations	0.456 772	0.186 603	0.473 724	0.195 570	0.442 724	0.211 570
F test on excluded instruments (first stage)	4.99	29.99	2.51	0.74	5.83	30.88
P value	0.000	0.000	0.082	0.477	0.000	0.000

Table 10a
Effects of Rainfall on Total Time Allocation

Explanatory Variable	**With Only Month and Region Effects**	**Also Including Month *by* Region Effects**
Rainfall	-0.226	0.339
($x10^3$)	(0.527)	(0.548)
Rainfall * North	2.464	-0.950
($x10^3$)	(0.670)	(5.111)
Income	0.381	0.335
($x10^6$)	(0.236)	(0.235)
Age	0.770	0.772
	(0.072)	(0.072)
Sex (female)	-3.321	-3.310
	(0.467)	(0.465)
Urban	1.218	0.868
	(0.563)	(0.557)
Region (north)	-1.473	1.586
	(1.071)	(3.866)
R-squared	0.152	0.163
No. of observations	4313	4313

Notes: Robust standard errors are in parentheses.

Table 10b
Effects of Rainfall on Total Time Allocation by Income Group

Explanatory Variable	With Only Month and Region Effects		Also Including Month *by* Region Effects	
	Low Income Group	**High Income Group**	**Low Income Group**	**High Income Group**
Rainfall ($x10^3$)	-1.341 (0.893)	-0.244 (0.995)	0.258 (0.942)	-0.153 (1.032)
Rainfall * North ($x10^3$)	3.375 (1.075)	1.438 (1.330)	-3.873 (7.485)	9.685 (19.57)
Income ($x10^6$)	9.29 (9.50)	-0.280 (0.269)	3.26 (0.933)	-0.320 (0.274)
Age	0.750 (0.121)	0.887 (0.135)	0.750 (0.122)	0.868 (0.135)
Sex (female)	-4.496 (0.820)	-1.411 (0.792)	-4.567 (0.825)	-1.437 (0.796)
Urban	1.004 (1.072)	2.018 (0.968)	0.465 (1.095)	1.893 (0.987)
Region (north)	-3.709 (1.988)	-1.479 (1.958)	4.688 (5.218)	-23.91 (69.19)
R-squared	0.199	0.177	0.218	0.183
No. of observations	1396	1459	1396	1459

Notes: Robust standard errors are in parentheses.

Table 11
Effects of Child Labor on Schooling by Income Group

Explanatory Variable	Strategy A		Strategy B		Strategy C	
	Low Income	High Income	Low Income	High Income	Low Income	High Income
Child labor	-0.403	-0.159	-0.641	-0.273	-0.405	-0.191
	(0.099)	(0.180)	(0.487)	(1.300)	(0.096)	(0.261)
Rainfall ($x10^3$)	-	-	-	-	-0.952	-0.071
					(0.714)	(0.665)
Rainfall * North ($x10^3$)	-	-	-	-	1.307	2.169
					(0.839)	(1.078)
Income ($x10^6$)	5.28	-0.291	3.91	-0.239	4.27	-0.271
	(7.46)	(0.188)	(8.13)	(0.233)	(7.60)	(0.188)
Age	0.087	-0.302	0.337	-0.156	0.067	-0.263
	(0.149)	(0.273)	(0.567)	(1.833)	(0.146)	(0.378)
Sex (female)	-3.401	-2.796	-3.769	-2.213	-3.241	-2.321
	(0.662)	(0.618)	(1.279)	(1.530)	(0.673)	(0.641)
Urban	-1.099	-0.651	0.480	-0.102	-0.889	-0.500
	(0.919)	(0.840)	(1.611)	(3.572)	(0.910)	(1.071)
Region (north)	-2.342	-1.747	3.681	11.24	-4.362	-2.764
	(1.358)	(1.037)	(3.215)	(2.805)	(1.548)	(1.509)
R-squared	0.282	0.372	0.173	0.378	0.286	0.380
No. of observations	1475	1537	1396	1459	1396	1459
F test on excluded instruments (first stage)	5.86	2.14	1.16	0.22	4.95	1.25
P value	0.000	0.024	0.314	0.803	0.000	0.267

Notes: Robust standard errors are in parentheses.
Income groups are defined as follows. Low income groups are households whose annual income is less than 135,000 cedis and high income groups are households whose annual income is greater than 420,000 cedis. About a third of children in the sample belong to low income groups and about a third to high income groups.