The Consequences of Child Labor in Rural Tanzania: Evidence from Longitudinal Data*

Kathleen Beegle
Development Research Group
The World Bank

Rajeev H. Dehejia
Department of Economics and the Fletcher School
Tufts University
and
NBER

Roberta Gatti
Development Research Group
The World Bank
and
CEPR

Sofya Krutikova
University of Oxford

February 2007

Abstract

This paper exploits a unique long-horizon longitudinal data set from Tanzania to examine the long-run consequences of child labor on education, employment choices, and marital status. Using crop shocks as instruments, our 2SLS estimates indicate that child labor is causally associated with reduced educational attainment (both as measured by the number of school years as well as by an indicator capturing completion of primary school). Interestingly, this result appears to be entirely driven by the sample of boys, for whom doubling labor hours from a mean prevalence (16 hours) would imply losing 80% of a school year. Boys who worked when young are more likely to be farming (as opposed to earning a wage), although we could not find evidence that child labor is associated with noticeable differences in the choices of crop (cash versus subsistence) or with subsequent migration. For girls, the main discernable effect is on early marriage: a higher level of child labor hours is associated with a substantially greater chance of being married 10 to 13 years later.

* We thank participants at the Fall 2006 Child Labor Conference at Indiana University, the Paris School of Economics, IDEI at the University of Toulouse, and the participants of the CSAE Conference at Oxford University for helpful comments. All errors are our own. Contact information: kbeegle@worldbank.org, rajeev@dehejia.net, rgatti@worldbank.org.

I. Introduction

This paper exploits a unique long-horizon longitudinal data set from Tanzania to examine the long-run consequences of child labor on education, employment choices, and marital status. The question we examine is important for many reasons. The assumption that child labor is harmful to children's development underpins both the theoretical literature and the policy debate. For example, from the policy perspective, there is a general perception that the worldwide returns to eliminating child labor are very large (see International Labour Organization [ILO], 2003). However, the evidence that rigorously quantifies the consequences of child labor is limited. Both theoretically and empirically, it is not clear whether child labor substantially displaces schooling. In rural settings in developing countries (and more than 70 percent of child labor in developing countries is rural; ILO, 2002), both school and child labor tend to be low-intensity activities, in contrast to the sweatshops and full-time work that characterize child labor in the popular imagination and which have existed historically in some urban settings in North America and Europe (see Basu, 1999). Furthermore, even if child labor does disrupt schooling, it presumably also provides the child with labor market experience that subsequently could lead to increased earnings. Which effect dominates is an empirical matter.

A growing empirical literature (reviewed in Section 2) analyzes the relationship between child labor and school attainment but, with few exceptions, this literature examines the correlation, not the causal relationship, between these variables. There are many reasons to doubt a causal interpretation of a simple correlation between child labor and education. Households that resort to child labor presumably differ along an array of dimensions, both observable (education, wealth, occupation) and unobservable (social networks, concern for children, etc.), from those that do not. Even within households, children's ability is not observed by the econometrician but is observable to parents. To the extent that parents send their least (most) motivated children to work, there is a negative (positive) correlation between child labor and school attainment simply based on selection.

Beegle, Dehejia, and Gatti (2005) estimate the causal impact of child labor on educational attainment, earnings, and health using two rounds of panel data from the Vietnam Standard Living Survey. In that work, we instrument for participation in child labor by using community shocks and rice prices, two variables that influence child labor but are plausibly exogenous with respect to household choices. With this strategy we estimate that, over the 5-year period spanned by our panel, the mean level of child labor reduces the probability of being in school by 30 percent and educational attainment by 6 percent. Indicators of health are not in general affected by child labor status. However, children who have experienced child labor are more likely to be working for wages five years later, and also have higher daily earnings (including both market wages and estimated farm wages). These estimates are significant at standard levels, and suggest that the returns to work experience are higher than the returns to schooling and that, overall, child labor might amount to a net benefit for children, at least until early adulthood. Using data on adults to simulate the future impact of child labor on current children, we find that returns to education increase with age, whereas returns to experience decline monotonically; the net present discounted value of child labor is positive for households with a discount rate of 11.5 percent or higher.

Although the empirical strategy employed in Beegle, Dehejia, and Gatti (2005) tries to overcome some of the limitations of the existing literature, the 5-year interval between the two rounds of data implies that long-run impacts must be extrapolated rather than estimated.

Conversely, with the Kagera data we can measure outcomes 10-13 years after child labor has taken place. In a previous paper we have documented the extent to which families use child labor in the Kagera region as a means to cope with agricultural shocks (Beegle, Dehejia, and Gatti, 2006). In particular, we find that in a sample of 7 to 15 year olds, crop shocks (as measured by crops accidentally lost to pests, insects, and fire) lead to a significant increase in the level of child labor and that households with assets are able to offset approximately 80 percent of these shocks. While educational enrolment decreases in response to shocks, households with a typical level of asset holdings are able fully to offset this effect. These results suggest that poorer households might be using assets as a buffer stock, drawing them down in times of need, whereas the behavior of wealthier households is consistent with an access to credit story. More importantly, given their characteristics, these shocks are plausible candidates to serve as instruments for participation in child labor: they are good predictors of child labor and, as they are not correlated over time within households, they appear to be exogenous at the household level (see detailed discussion in Section IV).

Using crop shocks as instruments, our 2SLS estimates indicate that child labor is causally associated with reduced educational attainment (both as measured by the number of school years as well as by an indicator capturing completion of primary school). Interestingly, this result appears to be entirely driven by the sample of boys, for whom child labor at the mean prevalence (16 hours) would imply losing 80% of a school year. Boys who worked when young are more likely to be farming (as opposed to earning a wage), although we could not find evidence that child labor is associated with discernible differences in the choices of crop (cash versus subsistence) or with subsequent migration. For girls, the main discernable effect is on marriage: a higher level of child labor hours is associated with a substantially greater chance of being married 10 to 13 years later.

The paper is organized as follows. Section II briefly reviews the existing literature. Section 3 introduces the data, and Section 4 describes the empirical methodology and discusses in detail the plausibility of our instrumental variable approach. Results are discussed in Section 5. Section 6 concludes.

II. Literature Review

There is a large literature that examines the tradeoff between child labor and schooling. In this section, we highlight a few of the existing results. Patrinos and Psacharopoulos (1995) show that factors that predict an increase in child labor also predict reduced attendance and an increased chance of grade repetition. Patrinos and Psacharopoulos (1997) estimate this relationship directly, and show that child work is a significant negative predictor of age-grade distortion. Akabayashi and Psacharopoulos (1999) show that in addition to school attainment, the reading competence of children (as assessed by parents) decreases with child labor hours. Finally, Heady (2003) uses objective measures of reading and mathematics ability and finds a negative relationship between child labor and educational attainment in Ghana.

The papers reviewed thus far examine the correlation between child labor and schooling, rather than the causal relationship. As we discuss in detail below, there are many reasons to doubt that these two coincide. A few recent papers address this issue. Using data from Ghana, Boozer, and Suri (2001) exploit regional variation in the pattern of rainfall as a source of exogenous variation in child labor. They find that a one hour increase in child labor leads to a 0.38 hour decrease in contemporaneous schooling. Cavalieri (2002) uses propensity score matching and finds a significantly negative effect of child labor on educational performance. Her results suggest that on average working in childhood is associated with a 10 percent reduction in the probability of being promoted to the next grade compared to children who do not work. Ray and Lancaster (2003) instrument child labor with household measures of income, assets, and infrastructure (water, telephone, and electricity) to analyze its impact on school outcome variables in seven countries. Their findings generally indicate a negative impact of child labor on school outcomes.[1] However, their two-stage strategy is questionable, as it relies on the strong assumption that household income, assets, and infrastructure satisfy the exclusion restriction in the schooling equations. Bezerra, Kassouf, and Arends-Kuenning (2007) use labor market indicators to instrument for child labor in their study of academic achievement in Brazil; these indicators include city population, state-level schooling and literary rates. While working seven hours or more per day results in a 10 percent decrease in achievement test scores relative to student who do not work, working up to 2 hours per day (14 hours per week) had minimal or no impact of school achievement. Finally, Ravallion and Wodon (2000) indirectly assess this relationship in their study of a food-for-school program in Bangladesh that exploits between-village variation in program participation. They find that the program led to a significant increase in schooling, but that only one eighth to one quarter of the increased schooling hours seem to have come from decreased child labor. This suggests that child labor does not offset schooling one for one.

The link between child labor and subsequent labor market outcomes is examined by Emerson and Souza (2006). They examine the effect of child labor and child schooling on adult outcomes. Using an instrumental variable strategy, they argue that, even controlling for

[1] For instance, they find that in Belize the initial hour of child labor leads to a reduction in years of schooling by 2.6 years. Note that in some cases they find the marginal impact of child labor to be positive. In particular, for Sri Lanka, the impact is positive for all school outcomes.

completed schooling, child labor has a negative effect on adult earnings. Their instruments for child labor and child schooling include the number of schools per child in the state, the number of teachers per school, and GDP per capita at age 12. Their paper has a number of potential strengths (a large, nationally representative data set), but also potential weaknesses: the key child labor question is retrospective and is asked only to those individuals who are working as adults.

Krutikova (2006) uses the same data as our analysis. She focuses on education 10-13 years after the initial child labor as the outcome, and builds on the design suggested by Beegle, Dehejia, and Gatti (2006) by using shocks in the early rounds of the Kagera data as instruments for child labor. She finds significant, negative effects of shocks on education. The present paper builds on and extends this in a number of important directions. First, we examine several outcomes associated with child labor: wages, occupational choice, schooling, and marital status. We also use information on farming to estimate a simple agricultural production function (this is work in progress). Second, we address some potentially important sources of bias in Krutikova's results. We provide evidence for the plausibility of the exclusion restriction by estimating the reduced form effect of the instruments (crop shocks) on a range of the long-run outcome variables. We also use a more refined specification that includes a single endogenous variable (hours worked) and excludes from the sample children that, as of the baseline, had little or no chance to enter schooling (i.e. older children who had never attended school at the baseline; see the discussion below). We also correct some potentially significant econometric limitations of her design (in particular, the use of household level instruments in a second-stage regression that includes household fixed effects). Finally, we investigate whether child labor has different effects on boys and girls.

The literature investigating other medium and long run consequences of child labor is sparse, mostly because of data constraints. As discussed in the introduction, Beegle, Dehejia, and Gatti (2005) investigates the causal impact of child labor on education, labor market outcomes, and health in Vietnam. In a sample of children from rural areas, increasing child labor by the average number of hours is associated, 5 years later, with a half–year reduction in schooling, with no significant impact on the prevalence of illness, and with a substantial increase in wages. Using the same data and an instrumental variable technique, O'Donnell, Rosati and Van Doorslaer (2005) investigate the impact of child labor on health outcomes. Their results differ in part from ours, as they find some evidence that work during childhood has a negative impact on health outcomes five years later.[2]

III. Data description

III.1 Data set

The Kagera Region of Tanzania is located on the western shore of Lake Victoria, bordering Uganda to the north and Rwanda and Burundi to the west. The population (1.3 million in 1988, about 2 million in 2004) is overwhelmingly rural and primarily engaged in producing bananas and coffee in the north and rain-fed annual crops (maize, sorghum, cotton) in the south. This study uses baseline data from the Kagera Health and Development Survey (KHDS), a longitudinal socioeconomic survey conducted from September 1991 to January 1994 covering the entire Kagera region (World Bank, 2004). Because adult mortality of the working age population (15-50) is a relatively rare event and HIV/AIDS was unevenly distributed in Kagera, the KHDS household sample was stratified based on the agro-climatic features of the region, levels of adult mortality from the 1988 Census (including both high and low mortality areas), and

[2] See Beegle, Dehejia, and Gatti (2005) for a detailed discussion of this discrepancy.

household-level indicators thought to be predictive of elevated adult illness or mortality, in order to capture a higher percentage of households with a death while retaining a control group of households without a death.

In 2004, another round of data collection was completed (Beegle, De Weerdt, and Dercon, 2006a). The goal of the KHDS 2004 was to re-interview the sample of 6,210 respondents from the 1991-1994 survey; this excludes 169 individuals who died over the course of the baseline rounds. In addition to the household survey, the KHDS 2004 included additional community-level surveys consistent with those carried out in the 1991-1994 rounds. A community questionnaire was administered to collect data on the physical, economic and social infrastructure of the baseline communities, as well as shocks experienced at the community-level. Over the course of 10-13 years, it was anticipated that a considerable number of individuals would have migrated from the dwelling occupied in 1991-1994. Considerable effort was made to track surviving respondents to their current location, be it in the same village, a nearby village, within the region, or even outside the region.

Because of the long time frame of the KHDS panel, we are able to study behaviors of children in conjunction with outcomes for these children as young adults. Among children ages 7-15 studied in Beegle, Dehejia, and Gatti (2006), 75% were re-interviewed in 2004, 21% were not located, and 4% were deceased. Among the children we study here (see details on the sample restriction in Section III), 76% were re-interviewed in 2004. Of these, 18% had moved far from their original village but still resided in Kagera, 11% resided outside Kagera but in Tanzania, and 2% were residing in Uganda. These children were, on average, 11 years old in their last interview from the baseline rounds. By 2004, they were almost 23 years old (Table 1).

III.2 Descriptive statistics

Our definition of child labor is the total hours spent working in economic activities and chores in the previous week (including fetching water and firewood, preparing meals, and cleaning the house). Economic activities for children consist predominately of farming, including tending crops in the field, processing crops, and tending livestock. We include chores as well as economic activities for two reasons. First, the concept of child labor (by ILO standards) is not restricted to only economic activities.[3] Second, in the largely rural sample of households in this study, it may be difficult to distinguish time in household chore activities and time spent preparing subsistence food crops. Children in the sample work on average a total of 17 hours per week, of which 10 are chores (Table 1). Girls spend on average 2.5 hours more time working on household chores than boys, and this difference is more pronounced among older girls.

Our education outcome variables are years of schooling and an indicator variable for having completed seven or more years of education (primary level). Individuals in the sample have an average of 6.4 years of schooling and 78% of them have completed primary school.

We can measure labor market outcomes with an array of different variables. As the economy in the Kagera region is based mainly on extensive farming, an important indicator of success is whether the individual earns a salary or if he or she is involved in cash cropping (mainly tobacco and coffee) rather than subsistence farming. Moreover, the literature indicates that child labor might help a child acquire plot-specific experience, which could be particularly important in rural economies (Rosenzweig and Wolpin, 1985). If this is the case, we should expect child labor to be associated with a lower level of individual mobility, in a setting where

[3]It should also be mentioned that the concept of child labor does not necessarily refer to simply any work done by a child, but, rather, work that stunts or limits the child's development or puts the child at risk. However, in survey data it is difficult (perhaps impossible) to appropriately isolate the portion of time spent working on the farm that qualifies under this very nuanced definition.

plot-specific experience cannot easily be exported to other agricultural contexts. We therefore investigate if child labor has an impact on the probability that individuals moved from their villages. This is possible because, unlike most of other surveys, the Kagera survey tracks individuals. While, in wave 5, 70% of the re-interviewed individuals in the sample were still living in the same or in neighboring villages, mobility is associated with significantly higher income gains for panel respondents (Beegle, De Weerdt, and Dercon, 2006b). So, while there may be advantages to experience on specific farm plots, lower mobility may hinder economic growth for these children.

Finally, we explore whether child labor significantly affects marital status. This is particularly interesting for our sample of girls, who tend to work more hours than boys, especially on household chores.[4] Since marriage is universal in Tanzania, we are effectively examining the influence of child labor on the likelihood of earlier marriage. Although we do not assume that marriage yields positive outcomes for those who marry, the work is motivated by the perception that the age at marriage can have significant effects on the future lives of women and their children.[5] Younger marriages increase health risks for women as well as potentially result in "worse" marriage matches.[6]

[4] For example, girls between 10 and 15 work 22 hours per week (15 of which are for house chores), against 18 hours for boys (11 of which are for house chores).

[5] Behrman et al. (2006) establish a casual nexus between education and age at marriage.

[6] Younger mothers are more likely to suffer from micronutrient deficiencies and be unaware of the health risks associated with pregnancy; they are also more likely to have children soon after marriage with increased risk of maternal and infant mortality (World Bank, 2007). Younger ages at marriage may result in curtailed education for girls, although it is difficult to ascertain the causality. In any case, a younger bride may be less able to assert power and authority in her marriage especially given that women marry men who are on average several years older.

IV. Empirical methodology

IV.1 Specification

We are interested in the relationship between outcomes in wave 5 (including education, occupation, and marital status) and the level of child labor intensity (which we measure through mean child labor hours in waves 1 to 4). An OLS regression of the form

$$Y_{i,\, t+10} = \alpha + \beta T_{i,t} + \gamma X_{i,t+10} + \varepsilon_{i,t+10}, \qquad (2)$$

where $Y_{i,\, t+10}$ are outcomes in wave 5, $T_{i,t}$ is mean child labor hours in waves 1 to 4, and X_i are household and community-level controls, is unlikely to estimate a causal relationship. The principal concern is omitted variable bias. The child labor decision is likely to be correlated with both household- and child-level covariates, not all of which will be observable to the researcher. For example, though we can control for parents' education we cannot control for their discount rates. At the child level, we have few covariates other than age, and thus, for example, cannot control for ability. Reverse causality is less of a concern because the outcome is measured 10 years after child labor intensity.

We address concerns with the OLS specification using an instrumental variables strategy. Our instrument, S_{it}, is an indicator for whether agricultural shocks (crop accidentally lost to pests and fire) occurred in waves 1 to 4. Thus our basic specification is a two-stage least squares procedure of the form:

$$T_{i,t} = \mathrm{a} + \mathrm{b}S_{i,t} + \mathrm{c}X_{i,t} + \nu_{i,t} \qquad (2)$$

$$Y_{i,\, t+10} = \alpha + \beta \hat{T}_{i,t} + \gamma X_{i,t+10} + \varepsilon_{i,t+10}. \qquad (3)$$

The instrument is motivated by previous work (Beegle, Dehejia, and Gatti 2006) in which we show that the extent to which households use child labor in response to shocks varies according to whether families have sufficient assets to buffer the impact of the shock. Furthermore, families tend to increase labor to greater extent for older children than younger children.

Thus, our instrument (an indicator variable for whether a shock occurred in a family in a given wave) is interacted both with the (log) level of assets and with the age of the child[7], providing variation within the household over time and also between children in a given round. In particular, our first-stage specification is

$$T_{ijt} = \alpha_j + \delta_t + \gamma_w + \beta_1 X_{ijt} + \beta_2 \cdot S_{jt} + \beta_3 (S_{jt} \cdot A_{jt}) \\ + \beta_4 \cdot A_{ijt} + \beta_5 (S_{jt} \cdot Age_{ijt}) + \beta_6 \cdot Age_{ijt} + \varepsilon_{ijt}, \quad (4)$$

where α_j, δ_t, and γ_w are household, time (season), and survey-round fixed effects respectively, and predicted hours are estimated for individual *i* in household *j* at time *t*.

. We impose several restrictions on the sample we examine. Following our previous work, we consider children between the ages of 7 and 15 in the baseline survey. Note that the prevalence of labor among younger children is low. Likewise, by some definitions, labor at age 14 and above would not be viewed as a particularly serious form of child labor. We also have information on whether children have ever been to school by wave 4. Tabulation of this variable shows that only 32% of 7-8 year olds had attended school at some point in time, which is consistent with widespread tendency to delay enrollment, while among children age 13 and above only 12% had never been school. It is unlikely that these older children will enroll in the future. At the same time, the data suggest that, in response to a shock, households are more likely to employ the labor of the older, more productive children. Because of this, if we included these children in our sample, we would be likely to find a strong negative correlation between years of schooling and child labor. As a result, our sample includes all 7-15 year olds who were in school

[7] More precisely, our asset measure captures all physical stock, including durables but excluding land.

at the relevant wave and those children who have not yet entered school but are still young enough to have a chance to enroll (7-9 year olds).

IV.2 Discussion: the validity of the instruments

In this section, we discuss whether crop shocks plausibly satisfy the requirements for a valid instrumental variable.

Both our previous work and the estimates based on the sample used in this paper confirm that crop shocks are significant predictors of child labor. Table 2 reports estimates from a first-stage regression, where total child labor hours (column 1) and chore hours (column 2) are regressed on the instruments and other regressors (such as gender, region, age squared, and log per capita expenditure). The occurrence of a shock is associated with an average increase of about 4 working hours for a 10 year old child. The instruments are jointly significant at the one percent level (with an F-statistic of 4), but are borderline with respect to the critical values suggested for weak instrument tests. We are currently working to expand the instrument set to include rainfall variation, which will hopefully increase our first stage power.

Our previous work has documented that crop shocks are transitory and uncorrelated with household characteristics. In particular, we show that the occurrence of a shock is neither significantly predicted by past shocks nor by household characteristics (with the exception of an indicator for female-headed households which we control for in our subsequent specifications). Thus, the evidence supports the view that crop shocks are plausibly exogenous with respect to household characteristics.

The remaining concern is whether crop shocks satisfy the exclusion restriction, i.e., that these shocks affect education and labor outcomes only through child labor. The relevance of this concern is supported by an influential strand of literature suggesting that transitory shocks can have long term consequences for households (see, for example, Ravallion and Lokshin, 2005). In our previous work we explored the contemporaneous effects of the crop shocks on household wealth, and found no significant effect on cash per capita and physical assets per capita. We did, however, find a negative and significant effect of the shocks on durable assets, which is stronger among poorer households.

Even if the contemporaneous effects of shocks are limited to child labor and durable assets, shocks could nonetheless affect long-run outcomes through channels other than child labor. We investigate this concern in a number of ways. First, we examine the reduced form effect of agricultural shocks in waves 1 to 4 on a range of outcomes 10+ years later.[8] In particular, we regress wave 5 measures of household wealth, including (log) values of physical and business assets, durables, farm equipment, land, and occupied dwellings on shocks, while controlling linearly and interactively for initial wealth. The effect of shocks and their interaction with initial wealth values are not typically significant (Table 3). The exception is durable assets for which the shock-asset interaction is positive. Note however that the direct effect of shocks on low-wealth households, although negative, is not significant. The positive interaction implies that the effect of shocks on higher wealth households is even smaller in absolute terms. These results suggest, albeit indirectly, that given initial conditions, shocks in waves 1 to 4 did not have permanent effects on a number of important variables in wave 5. In other words, they support the hypothesis that while shocks account for a significant variation in child labor, their effects on outcomes of interest are likely to be only short term.

---

[8] This could be extended by verifying that shocks are not correlated with such causes of attrition in the sample as mortality and destitution; this is work in progress.

Second, we exploit cross-sectional variation in the size of shocks to test whether our second stage results are driven only by large shocks or whether we obtain similar results for smaller shocks. Smaller shocks are less likely a priori to have long-lasting impacts on households, except through their impact on contemporaneous variables (child labor). We present these results below (in Section V.2 along with other robustness checks).

Finally, we use the adults in our data as a comparison group for children. In particular, using the same specification that we use for children, we test whether for adult labor hours in waves 1 to 4 have a significant impact on outcomes in wave 5 when instrumented with agricultural shocks. Since the outcomes we examine (schooling, farming, working for a wage, marital status) are typically predetermined for adults in waves 1 to 4, we would find a significant effect under one of two circumstances: (1) if the shocks in wave 1 are sufficiently large to affect adult outcomes such as farming, working for a wage, or marital status, or (2) if the instruments are correlated with household-level unobservables (e.g., if households that experience shocks are for unobservable reasons more likely to have lower levels of education). We also present these results in Section V.2 below.

V. Results

V.1 Baseline OLS estimation

We first present the results of OLS regressions of our main outcomes, as measured in wave 5, on average child labor over waves 1 to 4 (Table 4). Although we are aware that in this context OLS estimates are likely biased due to unobservables, we use them as baseline estimates.

Child labor (as measured by the average number of hours worked by each child in the 4 baseline waves) is associated with reduced schooling as measured by number of years of schooling attained and the probability of completing primary education. The negative association also holds for the probability of staying in or near the same village in wave 5 as in the last round of the baseline and the probability of being a wage worker in the last 12 months. The results further suggest that child labor is positively correlated with the probability of being a farmer in the last 12 month, growing cash crops, and being married. Although these results are not statistically significant, the magnitudes are not trivial. For instance, the mean prevalence of child labor is associated with losing one quarter of a school year. As we include (log) per capita expenditure in specification, the regression controls for relative economic status cross-sectionally; thus, it is unlikely that in this context child labor hours are simply picking up an effect of household poverty. Splitting the sample by gender shows that among girls child labor is associated with a statistically significant increased probability of marriage (which, given the sample characteristics, typically implies marrying at a younger age).

However, because of the potential sources of bias in the OLS specification, as discussed in Section IV.1, we are reluctant to interpret these coefficients causally.

V.2 Instrumental variable estimation

We now discuss the results of estimating the impact of child labor hours on outcomes using a two stage least squares procedure. In the first stage, labor hours are predicted from a regression of child labor hours on shocks and their interactions (equation (4) above). We average the 4 predicted values obtained for each child (one per wave) into a summary measure of average predicted child labor per child. These values of child labor are then used in the second stage of the regression, where standard errors are bootstrapped to correct for the use of predicted hours.[9]

[9] Note that as we have on average 2.3 children in the relevant age range per household, the second stage could be estimated with household fixed effects. However, the variation of education outcomes across children within a

Table 5, columns (1) and (2), shows that the 2SLS estimates of the effect of child labor on education are negative, statistically significant, and triple in magnitude to those estimated with OLS. The mean level of child labor is associated with a decrease in half a year of schooling and a 7 percentage point reduction in the chance of completing primary school. These results are in line with those obtained for Vietnam in Beegle, Dehejia, and Gatti (2005), where we find that doubling average work hours (7 extra hours in that sample) is associated with a similar decline in school attainment. In both papers we find 2SLS effects are greater than OLS effects.

To the extent that families send the least gifted children to work and skills in the classroom and in the field are positively correlated, OLS would overestimate the impact of child labor on schooling relative to the causal effect (as estimated by IV). However, our results instead lend support to one of two alternative views: either that classroom and agricultural skills are positively correlated and families send their most gifted children to work, or that academic and agricultural skills are negatively correlated and families send the children most suited to work in the field to work. Both views are plausible. The first is justified to the extent that child work in response to an agricultural shock is very valuable to the household, so that parents' may decide to use their most talented children. The second is justified by parents trying to minimize the harm they cause their children by increasing their labor activities.[10]

In column (3), child labor does not appear to be significantly associated with migration. However, column (4) shows that individuals who worked more when young are more likely to be farming in young adulthood (wave 5). Child labor has no significant impact either on the choice

---

single family is low and a second stage fixed effect specification is unable to estimate any of the seven coefficients of interest with any degree precision.

[10] This validates one of the key predictions of the model presented in Horowitz and Wang (2004). Another possible interpretation is attenuation bias due to measurement error in child labor hours in the baseline.

between farming cash or subsistence crops (column (5)),[11] or on whether the individual had a wage or salary job in the past 12 months (column (6)). Overall, these results suggest that the primary impact of increased child labor in waves 1 to 4 is on the farming / not-farming margin, rather than migration or choice of crop. This is consistent with the Rosenzweig and Wolpin (1985) framework in which child labor can impart plot specific experience that is difficult to transfer to other activities.

Finally, in column (7), we find that child labor is associated with a significant increase in the probability of marriage. At the mean level of child labor intensity, marriage by round 5 is 10 percent more likely. As noted in the discussion above, since marriage is almost universal in Kagera, this result suggests that child labor is associated with earlier marriage.

One of the most striking features of Table 5 is that when we split the results by gender, we find that the education and farming result is driven by boys and that the marriage result is driven by girls. One rationalization of this finding is that child labor does not affect female education because girls' education is a lower priority to begin with. (The flipside of this argument is that to the extent that female education is a lower priority those girls who are adding to their education between rounds 4 and 5 are committed to it, despite the increased demands of child labor). It is also worth noting that girls participate to a greater extent in chores than agricultural work, and it is possible that chores are less harmful to education than agricultural work. The marriage result suggests the possibility that child labor plays a role in marriage markets (for example that additional experience in housework increases the value of a girl in the marriage market). We plan to explore this in greater detail in future work.

[11] Note that farming and working for a wage are not mutually exclusive. Of 1,318 people: 180 do neither; 142 working for wage/salary, not farming; 653 farming, not working for wage/salary; 343 both farming and working for wage/salary.

V.3 Robustness checks

In this section, we subject our results to a series of robustness checks.

First, we check the sensitivity of our results to the choice of sample. As discussed in Section IV.1, the regressions in Tables 2-5 are estimated on the sample of 7-15 year olds who were either in school as of each wave or who were not in school but still young enough to start school at some later point (age 9). This restriction is particularly important for education outcomes because while child labor and education are simultaneous decisions, we are interested in identifying the impact of child labor (rather than, say, delayed enrollment) on educational attainment. Including children who are unlikely ever to attend school (i.e., those older than 9 and not at school in wave 4) in the sample would naturally bias our education results towards finding a stronger negative impact of child labor on education (since we would be including in the sample children who are working and not in schools in waves 1 to 4 and who are unlikely to obtain additional schooling between waves 4 and 5). When we run our instrumented regression on the full sample of children between ages 7 and 15 (Table 6), we find, as expected, larger coefficients for school years and primary school completion, but no discernable difference in the effect on the other outcomes (Table 6).

Second, one could argue that our shocks might proxy for underlying risk faced by the households in our sample. An extensive literature highlights that there might be significant costs associated with households' response to risk – including crop and labor choices – that are not necessarily captured by the measured response to shocks (see Morduch 1995 and Pörtner 2006). We address this concern by introducing a proxy for underlying agricultural risk into the specification, namely the standard deviation of rainfall in the village (Table 6). This has no effect on our coefficient of interest.

Finally, we exploit cross-sectional variation in the size of shocks to test whether our second stage results are driven by large shocks or whether we obtain similar results for smaller shocks. As discussed in Section IV.2, we are more confident that the exclusion restriction required for a valid instrumental variable (namely that the instrument affects the outcome only through the endogenous variable) will be satisfied for small shocks. In Table 7, we re-estimate our results, using three alternative definitions of the shock (an agricultural shock that results in a loss of respectively at most 5,, 10 and 20 percent of the crop) and excluding those individuals who experienced shocks larger than the respective thresholds (for example at the 20 percent threshold we exclude 71 individuals from our original sample who experienced a crop loss of greater than 20 percent). We find that our results are similar in magnitude to our baseline specification in Table 5 for each threshold and statistically significant for shocks of 10 percent of crop loss or greater.

VI. Discussion and future research

In this paper we investigate the impact of child labor on education and labor market outcomes using panel data from the Kagera region of Tanzania. Building on our previous work, we use the occurrence of crop shocks and their interaction with relevant household and individual characteristics (assets and age) as instruments for child labor. In the two-stage least squares regressions, we find a negative and significant effect of child labor on school years and the probability of completing primary school 10 to 13 years later. Moreover, child labor is significantly positively associated with the probability of being a farmer. This suggests that the extra farm labor associated with a shock might induce plot-specific experience that ties children more closely to the land, an interpretation which is consistent with Rosenzweig and Wolpin

(1985). These results are mainly driven by the sample of boys. The only significant effect for girls is on the probability of marriage, which suggests the possibility that child labor increases girls' value on the marriage market. This is consistent with the findings of Behrman et al. (2006).

In future work we intend to use data on bride prices to examine whether child labor in fact increases girls' value on the marriage market. We also intend to explore further avenues for instrumenting child labor, in particular historical satellite data on rainfall.

References

Akabayashi, H., and Psacharopoulos, George. (1999) “The Trade-off between Child Labor and Human Capital: A Tanzanian Case.” *Journal of Development Studies* 35(5): 120-140.

Basu, K. (1999) “Child Labor: Cause, Consequence, and Cure with Remarks on International Labor Standards.” *Journal of Economic Literature* 37: 1083-1119.

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti (2006) “Child Labor and Agricultural Shocks.” *Journal of Development Economics* 81(1): 80-96.

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti (2005) “Why Should We Care about Child Labor? The Education, Labor Market, and Health Consequences of Child Labor.” World Bank Policy Research Working Paper 3479. CEPR Discussion Paper 4443. NBER Working Paper No. 10980.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon (2006a) “Kagera Health and Development Survey 2004 Basic Information Document.” World Bank.

Beegle, Kathleen, Joachim De Weerdt, Stefan Dercon (2006b) “Poverty and Wealth Dynamics in Tanzania: Evidence from a Tracking Survey.” manuscript.

Behrman, Jere R., Alexis Murphy, Agnes Quisumbing, Usha Ramakrishna, and Kathryn Young. (2006) "What is the Real Impact of Schooling on Age of First Union and Age of First Parenting? New Evidence from Guatemala.” World Bank Policy Research Working Paper no. 4023.

Bezerra, Márcio Eduardo G., Ana Lúcia Kassouf, Mary Arends-Kuenning. 2007. “The Impact of Child Labor and School Quality on Academic Achievement in Brazil.” mimeo.

Boozer, Michael, and T. Suri (2001) “Child Labor and Schooling Decisions in Ghana.” manuscript.

Cavalieri, C. (2002) “The Impact of Child Labor on Educational Performance: An Evaluation of Brazil.” manuscript.

Emerson, P., and A. Portela Souza (2006) “Is Child Labor Harmful? The Impact of Working Earlier in Life on Adult Earnings.” manuscript.

Heady, C. (2003) “The Effect of Child Labor on Learning Achievement.” *World Development*, 31: 385–398.

Horowitz, Andrew, and Jian Wang (2004) “Favorite Son? Specialized Child Laborers and Students in Poor LDC Households.” *Journal of Development Economics* 73: 631-642.

International Labour Organization (2003) *Investing in Every Child: An Economic Study of the Costs and Benefits of Eliminating Child Labour*. Geneva: International Labour Office.

International Labour Organization (2002). *A Future Without Child Labour*. Geneva: International Labour Office.

Krutikova, Sofya (2006) "Impact of Child Labor on Education Attainment in Adulthood: Evidence from Rural Tanzania." manuscript.

Morduch, Jonathan (1995) "Income Smoothing and Consumption Smoothing." *Journal of Economic Perspectives* 9 (3): 103-14.

O'Donnell O., F. Rosati, and E. Van Doorsaler (2005) "Health Effects of Child Work: Evidence from Rural Vietnam."*Journal of Population Economics* 18(3): 437-467.

Patrinos, H. A., and G. Psacharopoulos (1995) "Educational Performance and Child Labor in Paraguay." *International Journal of Educational Development* 15: 47–60.

Patrinos, H. A., and G. Psacharopoulos (1997) "Family Size, Schooling and Child Labor in Peru—An Empirical Analysis." *Journal of Population Economics* 10: 387– 405.

Pörtner, Klaus (2006) "Gone with the Wind? Hurricane Risk, Fertility and Education." Manuscript.

Ravallion, Martin and Michail Lokshin (2005) "Lasting Local Impacts of an Economywide Crisis." World Bank Policy Research Working Paper No. 3506.

Ravallion, M. and Q. Wodon (2000) "Does Child Labour Displace Schooling? Evidence on Behavioral Responses to an Enrollment Subsidy." *The Economic Journal* 110: 158-175.

Ray, R., and G. Lancaster (2003) "Does Child Labour Affect School Attendance and School Performance? Multi Country Evidence on SIMPOC Data." manuscript.

Rosenzweig, M. and K. Wolpin (1985) "Specific Experience, Household Structure and Intergenerational Transfers: Farm Family Land and Labor Arrangements in Developing Countries." *Quarterly Journal of Economics* 100: 961-987.

World Bank (2004) "User's Guide to the Kagera Health and Development Survey Datasets." mimeo.

World Bank. (2007) "World Development Report 2007: Development and the Next Generation." World Bank.

Table 1: Summary statistics

	Mean	SD
Baseline sample		
Hours worked in last 7 days	16.79	13.42
Chore hours in last 7 days	10.54	9.05
Any crop lost	0.34	0.47
Female	0.49	0.50
Age	10.91	2.60
Number of observations		4,746
Panel sample: 1991-2004		
Mean hours at baseline	16.79	10.56
Mean hours (predicted) at baseline	16.85	8.44
Female	0.49	0.50
Age at wave 4 baseline	11.44	2.77
Age in 2004	22.65	3.17
2004 Outcomes:		
School years	6.36	2.77
Completed primary	0.78	0.41
Stayed in/near village	0.69	0.46
Farming in past 12 months	0.76	0.43
Growing cash crop	0.55	0.50
Wage/salary job in past 12 months	0.37	0.48
Married	0.51	0.50
Number of observations		1,313

Notes: Baseline sample is restricted to children in school at baseline or less than 10 years of age and not yet enrolled. It includes children who are measured up to 4 times in the baseline panel (1991-1994). Hours includes hours working in economic (income generating) activities and in chores. Panel sample is the subset of children in baseline sample who are re-interviewed in 2004.

Table 2: 1st Stage estimation of child labor hours

	(1)	(2)
	Hours	Chore hours
Any crop lost	-0.727	1.940
	(3.078)	(2.018)
Asset value (log per capita)	0.553*	0.496**
	(0.268)	(0.175)
Assets * any crop lost	-0.419	-0.340*
	(0.267)	(0.175)
Any crop lost * age	0.478***	0.170*
	(0.143)	(0.094)
Female	1.767***	3.093***
	(0.425)	(0.279)
Age	6.781***	3.736***
	(0.732)	(0.480)
Age squared	-0.231***	-0.118***
	(0.033)	(0.022)
Dad:1-6 yrs of education	-0.399	-0.684
	(1.245)	(0.816)
Dad:7 yrs of education	0.073	-0.452
	(1.281)	(0.840)
Dad:8+ yrs of education	0.446	0.046
	(1.639)	(1.075)
Mom:1-6 yrs of education	-0.290	0.592
	(1.056)	(0.692)
Mom:7 yrs of education	-0.583	0.241
	(1.024)	(0.672)
Mom:8+ yrs of education	-0.464	1.398
	(2.522)	(1.654)
Number of observations	4,746	4,746

Notes: Household-fixed effects regressions from waves 1-4 at baseline for restricted sample of children described in text ages 7-15. Standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%. Hours includes hours working in economic (income generating) activities and in chores.

Table 3: Long-run shock effect on household wealth

	(1) Physical assets in wave 5	(2) Business assets in wave 5	(3) Durables assets in wave 5	(4) Farm equipment in wave 5	(5) Land in wave 5	(6) Occupied dwellings in wave 5
Shock between waves 1-4	-1.701	-0.041	-0.161	1.642	-0.474	-0.310
	(1.787)	(0.340)	(0.478)	(1.433)	(2.424)	(1.537)
Assets in wave 1[(a)]	0.157	0.035	0.074	0.189	0.138	-0.029
	(0.121)	(0.053)	(0.050)	(0.118)	(0.119)	(0.126)
Assets in wave 1 [(a)] x Shock	0.128	0.028	0.095***	-0.210	0.010	0.016
	(0.137)	(0.063)	(0.057)	(0.158)	(0.196)	(0.140)
Number of observations	2,163	2,163	2,163	2,163	2,163	2,163

Notes: (a) The control for assets in wave 1 differs by column. It refers to the wave 1 value of the dependent variable for each column.

Table 4: Impact of Child Labor in Waves 1-4 on Outcomes in Wave 5: OLS

	(1) School years	(2) Completed primary	(3) Stayed in/near village	(4) Farming in past 12 months	(5) Growing cash crop	(6) Wage/salary job in past 12 months	(7) Married
Boys and girls							
Mean child labor hours, waves 1-4	-0.011	-0.001	-0.001	0.001	0.001	-0.001	0.002
	(0.009)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Girls							
Mean child labor hours, waves 1-4	-0.011	-0.002	-0.003	0.000	0.000	0.001	0.004*
	(0.013)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Boys							
Mean child labor hours, waves 1-4	-0.010	-0.001	0.002	0.002	0.003	-0.002	0.000
	(0.012)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Table 5: Impact of Child Labor: 2SLS of 2004 outcomes

	(1) School years	(2) Completed primary	(3) Stayed in/near village	(4) Farming in past 12 months	(5) Growing cash crop	(6) Wage/salary job in past 12 months	(7) Married
Boys and girls							
Mean predicted child labor hours, waves 1-4	-0.033*	-0.004*	-0.000	0.004*	0.001	-0.000	0.004*
	(0.012)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Girls							
Mean predicted child labor hours, waves 1-4	-0.011	-0.003	-0.004	-0.000	0.001	0.003	0.005**
	(0.017)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Boys							
Mean predicted child labor hours, waves 1-4	-0.055*	-0.006**	0.003	0.007**	0.001	-0.003	0.003
	(0.018)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)

Table 6: 2SLS robustness checks: sample, rainfall variability, movers vs. non-movers

	(1) School years	(2) Completed primary	(3) Stayed in/near village	(4) Farming in past 12 months	(5) Growing cash crop	(6) Wage/salary job in past 12 months	(7) Married
All kids sample							
Mean hours (predicted) at baseline	-0.070*	-0.009*	-0.001	0.004**	0.001	-0.000	0.004*
	(0.012)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Number of observations	1,460	1,460	1,460	1,460	1,460	1,460	1,460
Controlling for rainfall variability							
Mean hours (predicted) at baseline	-0.032**	-0.004**	-0.000	0.004***	0.001	-0.000	0.004**
	(0.013)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of observations	1,313	1,313	1,313	1,313	1,313	1,313	1,313
Non Movers							
Mean hours (predicted) at baseline	-0.020	-0.001		0.005**	-0.001	-0.003	0.007**
	(0.020)	(0.003)		(0.002)	(0.003)	(0.003)	(0.003)
Number of observations	684	684	684	684	684	684	684
Movers							
Mean hours (predicted) at baseline	-0.041***	-0.007***	-0.001	0.001	0.002	0.002	0.000
	(0.015)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Number of observations	629	629	629	629	629	629	629

Notes: OLS estimates with bootstrapped standard errors are in parentheses. *** indicates significance at 1%; ** at 5%; and, * at 10%. Additional controls (not reported) are as in Table 6.

Table 7: 2SLS robustness checks: magnitude of the shocks

	(1) School years	(2) Completed primary	(3) Stayed in/near village	(4) Farming in past 12 months	(5) Growing cash crop	(6) Wage/salary job in past 12 months	(7) Married
Crop shocks<5% of total							
Mean hours (predicted) at baseline	-0.028**	-0.004***	-0.001	0.003	0.001	-0.000	0.003
	(0.011)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of observations	1189	1189	1189	1189	1189	1189	1189
Crop shocks<10% of total							
Mean hours (predicted) at baseline	-0.033*	-0.004**	-0.000	0.004**	0.001	0.000	0.004***
	(0.012)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of observations	1214	1214	1214	1214	1214	1214	1214
Crop shocks<20% of total							
Mean hours (predicted) at baseline	-0.036*	-0.005*	-0.001	0.004**	0.001	0.001	0.004**
	(0.011)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of observations	1242	1242	1242	1242	1242	1242	1242

Note: Each row of coefficients comes from a different regression, where only losses up to respectively 5, 10, and 20% of total harvest are considered shocks. In particular, the instrument shock is coded 0 if there was no shock, 1 if the household had a shock that implied the loss of 5, 01 or 20% of total crop, missing otherwise. Predicted child labor hours are then used in the second stage regression with controls as in Table 6.