

Returns to farm child labor in Tanzania

P. André, E. Delesalle, C. Dumas

Discussion Paper 2019-5

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain



Returns to farm child labor in Tanzania

Pierre ANDRE* Esther DELESALLE† Christelle DUMAS‡§

January 30, 2019

Abstract

In developing countries, the opportunity costs of children's time can significantly hinder universal education. This paper studies one of these opportunity costs: we estimate the agricultural productivity of children aged 10 to 15 years old using the LSMS-ISA panel survey in Tanzania. Since child labor can be endogenous, we exploit the panel structure of the data and instrument child labor with changes in the age composition of the household. One day of child work leads to an increase in production value by roughly US\$0.89. Children enrolled in school work 26 fewer days than nonenrolled children. Compensating enrolled children for loss in income can be accomplished with monthly payments of \$1.92. However, a complete simulation of a hypothetical conditional cash transfer shows that even \$10/month transfers would fail to achieve universal school enrollment of children aged 10 to 15 years old.

Keywords: Child labor, Human capital investment, Conditional cash transfer, Farm household, Production function, Tanzania

JEL Codes: D1, O1, J3

*Université de Cergy-Pontoise, THEMA, 33 Boulevard du Port, F-95011 Cergy-Pontoise. Email: pierre.andre@u-cergy.fr

†IRES, Université catholique de Louvain, esther.delesalle@uclouvain.be

‡Corresponding author. University of Fribourg, Department of Political Economy, Bd de Pérolles 90, CH-1700 Fribourg. Email: christelle.dumas@unifr.ch

§We wish to thank Martin Ravallion, Paul Glewwe, Marc Bellemare, Jason Kerwin, Isabel Günther, and Lorenzo Casaburi for their useful comments, as well as the participants in research seminars at the following institutions: the University of Minnesota, the University of Cergy-Pontoise, the University of Fribourg, and DIAL. We also wish to thank the participants at the following conferences for questions and suggestions: the German Economic Association, the Nordic Conference in Development Economics, the DIAL Conference on Development Economics and the Summer School of Development Economics.

1 Introduction

Worldwide, 144 million children aged 5 to 14 years old were economically active in 2012. Child labor is primarily concentrated in agriculture (58.6%) and is mostly performed within households (68.4% of child laborers are unpaid family workers).¹ Although work opportunities are a key determinant of children's time allocation and hence a key aspect of any fight for universal education, we know very little about the economic contributions of children. Indeed, children who provide wage work are a minority among working children and a very selected sample of this population. We should not infer anything from the wages that they earn to a broader population. This article contributes to the understanding of barriers to education by estimating the productivity of children in household farming in rural Tanzania. For most children, this estimate of the opportunity cost is better than the children's average wage in the country.² To do so, we estimate several production functions, compute the productivity of children and exploit the features of the LSMS-ISA panel dataset for identification. Finally, we use these results to simulate the effect of a hypothetical conditional cash transfer.

Our article contributes to two different strands of the literature. The first is the estimation of the opportunity cost of children, about which the literature has been scarce and has mostly focused on paid employment. IPEC (2007) studied different sectors employing children and concluded regarding the diversity of situations in terms of children's return to labor that it ranges from children who are as productive as adults and are paid as such to children who have significantly lower productivity or the same productivity but who are paid much less.³ To our knowledge, this study was the only one evaluating

¹Diallo et al. (2013).

²Indeed, in our Tanzanian sample, 80.65% of child workers are employed in household farms.

³Children exhibit a productivity similar to that of adults in carpet manufacturing in India but are paid less (Levison et al., 1998; Anker et al., 1998). In the fishing sector in Ghana and the construction sector in Uganda, adults and children receive similar pay. However, in most other sectors considered in the study (chop bars in Ghana, the pyrotechnics industry in the Philippines and vehicle repair in Delhi), children are paid less than adults, and this reduction in pay is greater than the difference in productivity.

child productivity, and it did not cover farm labor. However, the large body of literature on market imperfections in developing countries has incidentally addressed this question (Jacoby, 1993; Skoufias, 1994; Lambert and Magnac, 1997) since authors have provided estimations of agricultural production functions. The opportunity cost of individuals who do not participate in the market is called the shadow wage and is equal to the marginal productivity of their on-farm labor. The papers by Jacoby (1993) (for Peru) and Skoufias (1994) (for India) did not compute the marginal productivity of children, but the parameters associated with child labor in the Cobb-Douglas production function are fairly low compared to those for adult labor, indicating that children contribute only marginally.

This work relates to a second strand of the literature: the effect of conditional cash transfers (CCTs) on children's time allocation. CCTs have been increasingly seen as an efficient tool for fighting current and future poverty. They have been implemented in almost all Latin American countries but remain rare in Africa, except for pilot studies. The evaluation of these programs leads to the conclusion that child labor supply reacts very heterogeneously to cash transfers. For instance, Edmonds and Schady (2012) found that a transfer amounting to approximately 7% of the GNI per capita leads to a sharp decline in paid employment by 10 percentage points and in unpaid economic activity by 19 percentage points in Ecuador.⁴ Large effects relative to the transfers (3% of GNI per capita) were also found in the Food for Education program in Bangladesh (Ravallion and Wodon, 2000). In comparison, Attanasio et al. (2010) found no effect of the program on child participation in economic activities in Colombia.⁵ Other Latin American programs lead to intermediate results, sometimes with much larger transfers.⁶ From these results, it is difficult to extrapolate the amount of transfer that would lead to nonnegligible increases

⁴This finding corresponds to a decline in paid employment by 41% and in unpaid employment by 34%.

⁵The transfer in Colombia is approximately 5% of the GNI per capita when a child is in secondary school.

⁶See, for instance: Edmonds and Shrestha (2014); Bourguignon et al. (2003); Ravallion and Wodon (2000); De Leon and Parker (2000); Skoufias et al. (2001); Carpio et al. (2016); Galiani and McEwan (2013); Dammert (2009); Barrera-Orsorio et al. (2011); Behrman et al. (2011); Schady and Araujo (2006).

in schooling and reductions in child labor, if one were to implement a conditional cash transfer in Africa. Bourguignon et al. (2003) showed that ex-ante evaluations are useful tools for designing programs; it is therefore of interest to identify the key parameters of these models ahead of the implementation of an economic policy. Indeed, de Hoop et al. (forthcoming) showed that child labor might actually increase when the cash transfer only provides a partial schooling subsidy. Kakwani et al. (2005) attempted to provide an ex-ante assessment of the implementation of a cash transfer program conditional on school attendance in 15 Sub-Saharan African countries. However, the model failed to reproduce the conditionality due to a lack of data on the opportunity costs of children.⁷ This work provides an ex-ante evaluation of a conditional cash transfer, based on our measure of the value of children's time.

The main contribution of this paper is the measure of the productivity of child labor. Provided that most working (African) children are unpaid family workers on farms, we focus on them and identify their productivity. The estimation of production functions is delicate since inputs are chosen to maximize profits and might be plagued by endogeneity issues. In our case, we exploit both the panel dimension of the data to control for unobserved and permanent determinants of household productivity, and we instrument child labor with the number of children belonging to the household. With an imperfect labor market, the household characteristics affect production choices, such as labor input (Singh et al., 1986). Provided that we control for household fixed effects, the identification strategy relies on the aging of children between rounds of the panel. We also provide robustness checks to confirm that the endogeneity of other inputs does not plague our estimates. When allowing for different production functions, we find consistent results for child productivity. One day of work performed by a child between 10 and 15 years old leads to an increase in production value by roughly US\$ 0.89, depending on specifications. Children enrolled in school work

⁷The model therefore provided an ex-ante evaluation of an unconditional cash transfer.

on average 26 fewer days per year than nonenrolled children. Compensating them for the loss of income could be accomplished by monthly payments of \$1.92. However, considering the income elasticity of education demand suggests that conditional cash transfers would have to be larger to truly bring children to school. We estimate that a monthly CCT of \$10 results in enrollment in school of only 14% of working poor children, and we discuss why this outcome occurs.

The remainder of the paper is organized as follow: section 2 describes the dataset, and section 3 presents the specifications and the identification strategy, while section 4 provides the results. Section 5 assesses the productivity heterogeneity by child age and gender, and section 6 computes the compensation that should be offered in a CCT.

2 Data

2.1 LSMS-ISA data

Our analysis is based on the Tanzanian LSMS-ISA (LSMS-Integrated Surveys on Agriculture) panel data. The panel is composed of three rounds: 2008-2009, 2010-2011 and 2012-2013.⁸ It is representative at the national, the urban/rural and the major agro-ecological zones levels. The original sample size is 3,265 households, spread over 409 enumeration areas across Tanzania and Zanzibar. Roughly two-thirds of these households are located in rural areas. The subsequent rounds revisit all households and include potential split-offs. Given the large rate of split-off and the extremely low attrition rate in the panel (only 4.84% of households surveyed in 2008 were not observed in 2010 or 2012), the third round of the panel interviewed 5,015 households. The panel dimension of the dataset is crucial to our analysis; therefore, it is extremely important to rely on a panel in which attrition is not an issue.

⁸The data collection occurred from October 2008 to October 2009 for the first wave, from October 2010 to December 2011 for the second wave and from October 2012 to December 2013 for the third wave. The fourth wave, collected in 2014-15, is a new sample.

The data are also particularly relevant for our analysis because they gather detailed information both about the production side and about the household side. All inputs and outputs are reported for each plot. Most importantly, who has provided labor is recorded so that we are able to compute how many days of work have been provided by each household member (and a general category for nonhousehold members). The estimation of production functions is always delicate because it might be difficult to observe all inputs, and failure to observe all inputs increases the risk of biased estimates. In the case of LSMS-ISA, a large set of inputs is collected in addition to labor days: organic and inorganic fertilizers, pesticides, and irrigation information. Land area is measured with a GPS, which is considered much more accurate than estimations.⁹ Self-reported land quality, as well as erosion of the plot, is also recorded. In addition, the LSMS-ISA data are matched with information about rainfall, greenness and temperature, obtained from satellite measurements. Appendix A describes the variables in detail, and Table A1 provides descriptive statistics about the sample.

2.2 Production

Although inputs and outputs are collected at the plot level, the panel does not allow for matching of plots from one date to the next, and we aggregate the information at the household level. As a consequence, our unit of analysis is a household observed in a given year. Household farm output is the self-reported monetary value of the harvest (in T-shillings), as declared by the household. We do not include farm outputs, such as fruit growing and cattle herding, in our analysis for the following reasons: 1) trees do not need much labor investment once they are planted; 2) from the data, it is not possible to compute the value of the cattle; and 3) in both cases, the labor time allocated to these activities is not collected. Neglecting cattle might be an issue in regard to the analysis of child

⁹In the first round, not all plots were measured with GPS. We use the GPS information when available and use the estimations when it is not.

labor since this cattle tending is one of the predominant activities of children. However, the household should equalize the marginal productivity of child labor between various possible farm activities and focusing on activities on the plots should not bias our estimate of child productivity.

Tanzania has two types of agriculture: in the North-North East, there is only one cropping season, which lasts roughly from November to May; the rest of the country has two cropping seasons, the short one in October-February and the long one in February-July. In the LSMS-ISA data, households were retrospectively interviewed about working time, inputs and production during the long and short rainy seasons separately. We aggregate the two seasons so that, for each household, we use the output and inputs for the full year.

2.3 Child labor in the data

Much information is available about child labor: we know whether the child worked in the week before the survey (and how many hours), the hours of domestic work on the day before the survey and, as previously said, the days of agricultural work over the previous year. The data confirm the importance of agricultural work for children living in rural areas of Tanzania. Among the 5 to 15 year old children who live in rural areas, 76% did not work in the previous 7 days. Only 2% of children in this age range provided some wage work. When they did so, they earned an average daily wage of 2,696 TSh (equivalent to 2016 US\$2.6).¹⁰ Among children aged 5 to 15 years old who belong to a land-owning household, 22.4% participated in farming in the previous year. There was no difference by gender: 22.7% of boys participated, compared to 22.1% of girls. Two-thirds of children did not perform any domestic work on the day before the interview. Among those who provided domestic work, the median duration was half an hour. As expected, girls are more

¹⁰1,000 TSh in 2008 = US\$ 0.95 in 2016; authors' computation based on World Bank series of inflation rates in consumer price index and official average exchange rates. Sources: http://donnees.banquemondiale.org/indicateur/FP.CPI.TOTL.ZG?locations=TZ&name_desc=true.

likely to perform domestic chores (41% of them participate compared to 25% of boys) but when they do so, they do not spend more hours than boys.

The ILO definition of child labor includes all economically active children aged 5 to 15 years old. However, children from 5 to 9 years old work only marginally: Table 1 shows that the share of working children significantly increases with age and that older children also devote more time to agricultural work.^{11,12} Estimating the labor productivity of young children is difficult because few of them actually work. We therefore focus on the productivity of children older than 10 and younger than 15. However, we provide estimates of the productivity differential by child age.

Table 1: Agricultural child labor by age

Age	Share of working children	Nb. of days of working children	Observations
5	1.50%	ns	21
6	2.26%	ns	32
7	4.68%	ns	65
8	7.40%	31.11	103
9	16.27%	28.88	211
10	19.90%	35.95	277
11	27.11%	41.15	331
12	35.34%	37.48	463
13	41.83%	42.81	504
14	49.10%	46.35	629
15	51.86%	51.20	598

We restrict our analysis to farm labor because it encompasses the main activity of children and because it is the only activity for which the output is measurable. Our child labor variable is therefore the number of farming days in the previous year performed by children aged 10 to 15 years old. At the household level, we observe that, on average, 16 days of farming are provided by children, compared to 185 days by adults. Child labor therefore

¹¹Very few children are economically active in other sectors than agriculture: among the 5 to 9 year olds, 0.10% earn a wage, and 0.06% are self-employed.

¹²This finding is also consistent with 99% of 5 to 9 year old children being enrolled in school, while 87% of those 10-15 years old are enrolled.

constitutes 8% of total farm labor and as much (in days) as nonhousehold workers.¹³ Roughly one-third of child farm labor is devoted to preparing and planting, one-third to weeding and one-third to harvesting. Adults from the household have a fairly similar work allocation to them; if anything, adults spend slightly more time on the preparation of the field and less on harvesting (Figure A1).¹⁴ Finally, children mostly allocate their time to the same crops as their parents (see Figure A2), except for cassava: 34% of their working days are allocated to cassava, compared to only 20% for the adults. Cassava is known to be an easy crop to cultivate and might be an “easy” task to give to a child.

2.4 Sample

Our sample is composed of households farming land in which at least one adult member participates in the farming. Because we implement a household fixed-effect strategy, we only retain households surveyed at least twice. Households that are resurveyed but that have moved to a new location are discarded since we want to control for unobserved soil characteristics. Our total sample is composed of 4,986 households x dates (1,702 surveyed in 2008, 1,707 in 2010 and 1,577 in 2012).¹⁵

3 Production functions and identification

Estimating the productivity of children on the farm entails estimating a production function, which leads to a series of choices. First, different functional forms are available with different properties and sometimes different ways of estimating these functional forms. Sec-

¹³These numbers are stable across years.

¹⁴Nonhousehold members spend significantly more time on weeding (almost 40%) and significantly less on harvesting (22%).

¹⁵Special attention must be paid to households that split during the panel (20% of households in each wave). We treat differently households considered as original households and the split-offs. In practice, the “original” household after the split is mostly composed of members from the original household (85%), while the split-off is only composed of 35% original members, i.e., on average fewer than 2 members from the original household. Very often, the split-off household declares itself in a new location, while the original household is in the same location. As a consequence, we treat the split-off as a totally new household entering the panel.

ond, we must choose the set of inputs; third, we must address null values of inputs and output; fourth, we must also choose what type of productivity is reported; and finally, inputs on the farm are the result of a choice made by households and could therefore be endogenous. In this section, we describe our set of choices and explain how we manage the question of identification.

Before turning to this set of choices, we should also clarify why we have chosen to estimate production functions rather than profit functions. Indeed, profit functions can be expressed as functions of input prices rather than input levels, which is convenient because prices are more naturally assumed to be exogenous to the household, while input levels are the result of precisely the profit maximization and should therefore be considered endogenous. However, estimating profit functions run into two additional issues. First, collected prices must reflect relevant prices for the households, which is not necessarily the case if markets are imperfect, particularly for some inputs that the household can acquire without the market (seeds, for instance). Second, some households end up with negative profits. While this outcome could be due to shocks on production, it could also be due to incorrect estimation of the production costs; therefore, it is difficult to properly address this situation.

3.1 Production function

There is a large choice set in regard to specifying a production function. Insofar as we are interested in estimating child labor productivity, particularly by comparison with adult labor productivity, we have allowed for specifications that varied the degree of substitutability between the two types of labor. Much less emphasis has been placed on the analysis of the link between labor as a whole and the other inputs: we posit a Cobb-Douglas relationship between (total) labor and the other inputs. The elasticities of substitution between labor and other inputs are therefore assumed constant equal to one.

We focus our attention on the possible substitution between child and adult labor. A Cobb-Douglas specification for the different types of labor does not seem attractive since it entails a zero production when child labor is not used, which is not supported by the data. A flexible functional form would be a CES function:

$$Y = A[L_a^{-\rho} + \gamma_c L_c^{-\rho}]^{-\alpha/\rho} X^\beta \quad (1)$$

where L_a and L_c are, respectively, adult and child labor, X are the other inputs (land, fertilizers...), and A is a productivity factor. $\rho \in [-1, +\infty[$ is the substitution parameter between the two types of labor, $\gamma_c \in [0, 1]$ is the relative productivity of child labor compared to adult labor, and $\alpha, \beta < 1$. Eq. (1) can be log-linearized:

$$\log Y = \log A - \frac{\alpha}{\rho} \log(L_a^{-\rho} + \gamma_c L_c^{-\rho}) + (\log X)\beta \quad (2)$$

but it cannot be linearly estimated and often leads to very unstable results (Henningsen and Henningsen, 2012).

Based on this general specification, we offer three different specifications and linearizations.

Perfect substitutes The first assumes that child and adult labor are perfect substitutes ($\rho = -1$). In this case, the total amount of efficient labor is $L_a + \gamma_c L_c$. The corresponding production function is:

$$\log Y = \log A + \alpha \log(L_a + \gamma_c L_c) + (\log X)\beta \quad (3)$$

When $\gamma_c L_c \ll L_a$, it can be linearly approximated by:

$$\log Y \approx \log A + \alpha \log L_a + \alpha \gamma_c \frac{L_c}{L_a} + (\log X)\beta \quad (4)$$

Imperfect substitutes The second venue consists of departing from the perfect substitution approach ($\rho > -1$). Assuming that $\gamma_c \left(\frac{L_c}{L_a}\right)^{-\rho} \ll 1$ and that ρ remains close to -1, Equation (2) can be approximated by a Taylor development as:

$$\log Y \approx \log A + \alpha \log L_a + \alpha \gamma_c \frac{L_c}{L_a} + \alpha(1 + \rho) \gamma_c \frac{L_c}{L_a} \left(1 - \log \frac{L_c}{L_a}\right) + (\log X)\beta \quad (5)$$

If the fourth term is equal to 0 (but $\alpha \gamma_c \neq 0$), then both types of labor are perfect substitutes.

Translog Finally, we can simply use a Translog function, which is very flexible and allows us to avoid any assumptions about the parameters.¹⁶ The Translog function has also been demonstrated to approximate the CES function when $\rho \approx 0$ (Kmenta, 1967).

$$\begin{aligned} \log Y \approx \log A + \alpha_a \log L_a + \alpha_c \log L_c + \alpha_{aa}(\log L_a)^2 + \alpha_{cc}(\log L_c)^2 \\ + \alpha_{ac} \log L_a \log L_c + (\log X)\beta \end{aligned} \quad (6)$$

Obviously, the Translog specification is the most flexible among the three offered specifications, but the coefficients cannot be interpreted as structural parameters. This step determines how we compare the results of the three specifications.

¹⁶In this specification, we adjust the Translog function to our hypothesis that the substitution between labor and other inputs is constant and equal to 1. Robustness tests show that the results are similar with a complete Translog function. The results are available upon request.

3.2 Child labor productivity

We want to compute a child's productivity per day of farm labor. We are more interested in the average productivity than in the marginal productivity. Indeed, if marginal returns are decreasing, marginal productivity informs us about the production obtained on the last day of work. A cash transfer would aim to reduce substantially, perhaps even suppress, child labor. To do so, families would have to be compensated for a larger number of days, and average productivity on those days is the relevant concept.

As a consequence, we provide for each specification the *average semi-elasticity of output with respect to days of child labor*. More precisely, our estimates of Equations (4), (5) and (6) allow us to compute the expected production in the absence of child labor and with the actual number of days of child labor for each farm for their observed level of other inputs. We compute for each household:

$$E_{Y_{cit}} = \frac{\log \widehat{Y}_{it}(L_c = L_{cit}^o) - \log \widehat{Y}_{it}(L_c = 0)}{L_{cit}^o} \quad (7)$$

where L_{cit}^o is the number of days of child labor observed in household i on date t . $E_{Y_{cit}}$ is the average labor productivity of children in household i on date t , estimated for households that use child labor.

We then weight households by amount of child labor and calculate the average of the individual semi-elasticities to obtain an aggregate measure of child labor productivity:

$$\overline{E}_{Y_c} = \frac{1}{\sum_{i,t} L_{cit}^o} \sum_{i,t} L_{cit}^o \cdot E_{Y_{cit}} \quad (8)$$

Because adult productivity is a benchmark against which child productivity should be evaluated, we also compute it. Adult on-farm labor is never null in our sample; therefore, to compute the average productivity on a meaningful margin, we use as a base the 10th

percentile of the adult labor distribution ($\underline{L}_a$ equals 38 labor days per year).¹⁷ The adult semi-elasticity $\bar{E}_{Y_a}$ is computed as:

$$E_{Y_{ait}} = \frac{\log \widehat{Y}_{it}(L_a = L_{ait}^o) - \log \widehat{Y}_{it}(L_a = \underline{L}_a)}{L_{ait}^o - \underline{L}_a} \quad (9)$$

$$\bar{E}_{Y_a} = \frac{1}{\sum_{i,t} L_{ait}^o} \sum_{i,t} L_{ait}^o \cdot E_{Y_{ait}} \quad (10)$$

3.3 Inputs

The inputs that are included in the estimation are the following: adult labor, child labor, land area, use of organic and inorganic fertilizer, use of pesticides, erosion of the plot, irrigation of the plot and productive assets. Nonhousehold labor days are aggregated to household adult labor days (there is hardly any nonhousehold child labor on the farms). Several inputs, such as child labor, fertilizers and pesticides, frequently have null values, raising the issue in our specifications of logs. To avoid restricting the sample to households with positive values of inputs, which would lead to selection bias, we follow MacKinnon and Magee (1990), Burbidge et al. (1988), and Pence (2006) and use a modified function of the logarithm, which is defined in (11):

$$\log^M(x) = \log \frac{1}{2} \left(x + \sqrt{1 + x^2} \right) \quad (11)$$

This function behaves similarly to the log function when x is large. As a consequence, for all inputs that have large values, the estimated coefficient reflects the increase in the production (expressed as a percentage) associated with an increase by 1% of the input. Given that child labor is often equal to 0 in our data, we cannot use this approximation for the interpretation of the coefficient. The same also holds for the other inputs that tend to be close to zero. The semi-elasticities of production to labor are computed considering

¹⁷The estimated production function fits the data only for the range of adult labor that is observed. We do not want to extrapolate outside of this range.

that the $\log^M$ function (instead of $\log$) is used for the estimation. The details for these computations are provided in Appendix B.

In the data, we also have a nonnegligible number of households declaring null production despite nonzero inputs because of disasters, such as droughts and pests. We must retain these observations with null production to avoid selection bias. We therefore use the same modified function. Given that the expected value of output is always large, we consider our function to be well approximated by the logarithm function and interpret it accordingly (the effect of one additional unit of input is expressed as a percentage increase in production).

3.4 Identification and specifications

Ordinary Least Squares (OLS) estimates of Equations (4), (5) and (6) can suffer from endogeneity bias for several reasons. First, unobserved permanent household characteristics (wealth, abilities, education, networks, etc.) can both influence a household's production and its labor allocation decisions. In addition, households observe the circumstances of production much more clearly than econometricians and therefore might adapt their farm allocation decisions to determinants of production that we do not observe (sunlight, pests, etc.).

Our solution to this endogeneity issue combines three techniques. First, we exploit the panel dimension of our data and control for household fixed effects. Second, we control for observed household time-varying characteristics, such as rainfall, temperature, greenness increase, idiosyncratic and covariate shocks that the household has declared. We also control for village-year fixed effects, which should also capture common shocks. Third, because unobserved idiosyncratic time-varying determinants can still be correlated with labor decisions and productions, we instrument child labor.

According to the agricultural household model (Singh et al., 1986; De Janvry et al., 1991), when markets are complete and competitive, the household production decision is

separable from its consumption decision, and labor demand depends only on inputs and output prices. The literature testing whether markets are perfectly competitive in developing countries has almost unanimously rejected the hypothesis (Jacoby, 1993; Skoufias, 1994; Chennareddy, 1967; Chavas et al., 2005; Le, 2009). In addition, it has used household composition as a determinant of farm labor supply. For instance, Benjamin (1992) tested whether the household labor demand is independent from the family composition in rural Java and rejected the separable assumption. We exploit the labor market imperfections in rural Tanzania¹⁸ and use the number of children aged 10 to 15 years old in the household as an instrument for child labor. Given that we control for household fixed effects, it amounts to predicting variations in farm labor, based on within household variations in household members' ages.

More precisely, for each household, we build a pool of children 10 to 15 years old at some point during the course of the panel. This pool of children consists only of offspring of at least one member of the household. Fostered children are excluded for endogeneity reasons: they could have been fostered in the household precisely because the household requires human resources.¹⁹ Ideally, we would use all offspring of adult household members. However, the dataset lacks information about those children: we do not know whether adults have children living elsewhere. The pool of children is therefore composed of all offspring of household members, who were recorded as belonging to the household on least at one date during the panel. The underlying assumption is that this pool of children constitutes all children who might belong to the household in case of shocks (offspring who never belonged to the household during the six years of the panel are assumed not to be available for help). This pool of children $\mathcal{N}_{ci}$ for household i is then fixed for the entirety of the panel, and N_{cit} is the number of children who belong to this pool and are aged 10

¹⁸Dumas (forthcoming) showed that only 5 % of households hire an external workforce. She also showed that positive rainfall shocks increase child labor, suggesting that the separability assumption does not hold.

¹⁹Safir (2009) showed that the household composition (in Senegal) reacts to shocks.

to 15 in year t :

$$N_{cit} = \sum_{k \in \mathcal{N}_{ci}} 1(10 \leq \text{age in year } t \text{ of child } k \leq 15). \quad (12)$$

This procedure is useful because only the aging of children provides variation in N_{cit} , not the actual presence of children in the household, which could be correlated with the needs of the household. Therefore, only the entry/exit of children into/out of the 10-15 age bracket provides the exogenous variation used in the IV estimation. Children who provide time variation in a given household are those included in the pool and who cross the age limit between two rounds. Since rounds are spaced by two years, they are aged 10 or 11 when they enter the age bracket, and they are aged 14 or 15 before exiting the age bracket. The productivity estimation thus relies on a mixed composition of children aged 10, 11, 14 and 15 years old.²⁰ When no heterogeneity in productivity is allowed, we estimate the average productivity on children of the previously mentioned ages, which should be close to the average productivity of children between 10 and 15 years old.

However, despite our instrument not being manipulated by the household, we could still face violations of the exclusion restriction if this entry/exit of children into/out of the age range had other consequences in terms of production. Starting with entry into the 10-15 age range, one additional child in the 10-15 age range is associated with one fewer child in the 5-9 age range. If children in this age range actually work and are productive, then our production function is misspecified (because labor performed by young children is not included), with consequences for the identification. This issue is unlikely to be serious since young children provide little labor. Conversely, exit out of the age range is in fact associated with an increase in adult labor time. Here, the problem is slightly different since we do control for adult labor. However, this control is imperfect if adult productivity is

²⁰We checked that the children in the pool were balanced with respect to age.

actually heterogeneous according to age. Allowing for heterogeneous productivity by child age helps us to check whether the exclusion restriction is violated, but we postpone this analysis until section 5.

To summarize, our specifications are the following, with i indexing households in village v and t indexing dates:

$$\log Y_{it} = \alpha \log L_{ait} + \alpha \gamma_c \frac{L_{cit}}{L_{ait}} + (\log X_{it})\beta + \mu_i + \zeta_{vt} + \varepsilon_{it} \quad (\text{Perfect})$$

$$\begin{aligned} \log Y_{it} = & \alpha \log L_{ait} + \alpha \gamma_c \frac{L_{cit}}{L_{ait}} + \alpha(1 + \rho) \gamma_c \frac{L_{cit}}{L_{ait}} \left(1 - \log \frac{L_{cit}}{L_{ait}} \right) \\ & + (\log X_{it})\beta + \mu_i + \zeta_{vt} + \varepsilon_{it} \quad (\text{Imperfect}) \end{aligned}$$

$$\begin{aligned} \log Y_{it} = & \alpha_a \log L_{ait} + \alpha_c \log L_{cit} + \alpha_{aa}(\log L_{ait})^2 + \alpha_{cc}(\log L_{cit})^2 \\ & + \alpha_{ac} \log L_{ait} \log L_{cit} + (\log X_{it})\beta + \mu_i + \zeta_{vt} + \varepsilon_{it} \quad (\text{Translog}) \end{aligned}$$

with μ_i standing for household fixed-effects and ζ_{vt} for village-year fixed effects (not systematically included). The first stage Equations are the following:

$$\frac{L_{cit}}{L_{ait}} = \delta_1 \frac{N_{cit}}{N_{ait}} + \delta_2 \log L_{ait} + (\log X_{it})\delta_3 + \nu_i + \xi_{vt} + \eta_{it} \quad (\text{Perfect and Imperfect})$$

$$\log L_{cit} = \delta_1 N_{cit} + \delta_2 \log L_{ait} + \delta_{22} \log L_{ait}^2 + (\log X_{it})\delta_3 + \nu_i + \xi_{vt} + \eta_{it} \quad (\text{Translog})$$

with ν_i household fixed-effects and ξ_{vt} village-year fixed effects. N_{cit} is named the number of children from 10 to 15, but it is actually the variable defined in Equation (12). N_{ait} is built similarly to N_{cit} : it is number of adults from 16 to 65 years old on date t among the adults observed at least once in household i during the panel.

The demographic structure is a valid instrument if:

$$E\left(\frac{N_c}{N_a} \cdot \varepsilon | \log(X), \log(L_a), \log(L_a), \mu, \zeta\right) = 0 \quad (\text{Perfect and Imperfect})$$

$$E(N_c \cdot \varepsilon | \log(X), \log(L_a), \mu, \zeta) = 0 \quad (\text{Translog})$$

A second issue to consider in specifications Imperfect and Translog is that the term to be instrumented appears more than once and with different interactions or functional forms. It is not advisable to instrument each of these terms because it would lead to an extremely low partial R-squared for each of the instrumented terms. Instead, we follow Wooldridge (2015) and implement a control function approach, which amounts to predicting $\hat{\eta}_{it}$ in the first stage equation and including it as a control in the main equation. The underlying idea is that $\hat{\eta}_{it}$ captures the endogeneity of the household behavior and then is controlled for, in the same spirit as the inverse-Mills ratio in a selection equation. Wooldridge (2015) showed that this parsimonious control function approach leads to more efficient estimates than the IV estimates when the left-hand-side variable is not linear in the endogenous variable.

So far, we have only discussed how to address the endogeneity of child labor. However, the households are also expected to choose other inputs. Part of the identification strategy for child labor productivity already answers most endogeneity concerns for other inputs. In particular, household and village-year fixed effects likely control for a substantial amount of joint determination issues. In addition, $\hat{\eta}$ should also pick up unobserved and idiosyncratic shocks occurring to households, which are also relevant for the determination of other inputs used, particularly adult labor.

Nevertheless, it is worth discussing the consequences of possible violations of exogeneity for the other inputs. Notably, it is important to recognize that adult productivity is in itself a result of interest, to which child labor productivity will be compared. If endogeneity issues remain for adult labor, then we should be cautious in the comparison. However, we can

provide additional tests to assess the extent of this remaining endogeneity. First, we test for child labor exogeneity, conditional on household and village-year fixed effects. If exogeneity is not rejected for child labor, then it is less likely that adult labor is endogenous, provided that we also condition on the same fixed effects. Second, neglecting the endogeneity of other inputs could lead to a bias in the child labor estimates under certain circumstances. According to the Zellner property, the IV estimates can be biased only if the instrument (N_c or $\frac{N_c}{N_a}$) is correlated with the other inputs, conditional on household and village-year fixed effects. This condition could be invalidated if, for instance, households anticipate that their children age and adjust for other inputs based on the increased productivity of the children. However, additional regressions show that inputs are barely correlated with N_c , i.e., the number of children aged 10 to 15 years old.

4 Results

4.1 OLS estimations

We start with OLS estimations of the three production functions.

Perfect substitutes Table A2 provides estimates when child labor and adult labor are assumed to be perfect substitutes. The α parameter estimate is 0.663 (coefficient of $\log L_a$) when not controlling for household fixed effects and 0.802 with household fixed effects, which is consistent with decreasing marginal returns to labor (since the coefficients are less than 1). The relative productivity of children γ_c compared to adults is estimated to $1.22 = 0.814/0.663$ when we do not control for household fixed effects (column 1). However, once we control for household fixed effects, the relative productivity of children decreases: γ_c is estimated at $0.71 = 0.569/0.802$, suggesting that more productive households are those that tend to make their children work, which is consistent with previous evidence that households with more work opportunities are those that employ children (Bhalotra

and Heady, 2003; Dumas, 2007). When controlling for observable shocks on the production (column 3), the estimate remains the same, and it somewhat decreases when allowing for unobservable shocks at the village level. The other inputs that display significant, positive effects in the specification with household fixed effects are: land area, inorganic fertilizer, and pesticides. Covariate shocks impact negative household production. Rainfall does not, perhaps because declared shocks are more accurate. The estimated semi-elasticities, computed with formulas (8) and (10), are provided in Table 2, Panel A. They are precisely estimated, but we find in this specification that the average productivity of children is similar to that of adults. On average, one additional day of child work is associated with an increase of 0.3% in production.

Imperfect substitutes Table A3 provides the estimates for production in which adult labor and child labor are imperfect substitutes. The adult productivity is very close to that estimated earlier; again, controlling for household fixed effects lowers the child productivity. In Table 2, Panel B, we find that, on average, one day of adult work is associated with a 0.66% increase in production, while one day of child work is associated with a 0.45% increase in production. This outcome is consistent with the Perfect estimates since child productivity is estimated to be two-thirds of adult productivity. The last column, however, provides lower semi-elasticities of work.

Given that the Imperfect model is an overmodel of the Perfect model, we can test whether the Perfect model is rejected. The bottom part of Panel B shows that, in all specifications but one, we can reject the Perfect model. However, allowing for the imperfect term does not improve the estimation since the R-squared is the same in Tables A2 and A3.

Translog Finally, Table A4 provides the estimates for the Translog specification. Table 2 Panel C shows that the child's semi-elasticity obtained from the Translog is strikingly the

Table 2: Average semi-elasticities of labor (based on estimates in Tables A2, A3 and A4)

	(1)	(2)	(3)	(4)
Panel A: Perfect substitute specification				
Adult	-0.0002 (0.0009)	0.0027** (0.0011)	0.0029** (0.0011)	0.0028** (0.0013)
Child	0.0045*** (0.0007)	0.0031*** (0.0009)	0.0029*** (0.0009)	0.0026*** (0.0010)
Panel B: Imperfect substitute specification				
Adult	0.0035** (0.0016)	0.0066*** (0.0016)	0.0066*** (0.0016)	0.0036 (0.0023)
Child	0.0057*** (0.0010)	0.0047*** (0.0013)	0.0045*** (0.0013)	0.0029* (0.0015)
Test $\alpha(1 + \rho)\gamma_c = 0$	6.030	6.036	5.589	0.109
Prob>F	0.015	0.015	0.019	0.741
Panel C: Translog specification				
Adult	0.0025*** (0.0008)	0.0047*** (0.0010)	0.0048*** (0.0010)	0.0037*** (0.0012)
Child	0.0057*** (0.0010)	0.0047*** (0.0015)	0.0045*** (0.0015)	0.0028 (0.0017)
Test $\alpha_{aa} = \alpha_{cc} = \alpha_{ac} = 0$	9.982	5.889	5.063	5.065
Prob>F	2.60×10^{-6}	0.001	0.002	0.002
Household FE		×	×	×
Climatic factors			×	
Village-year FE				×
Computation sample	1480	1480	1458	1480

Note: The computation sample is made of households with positive adult labor, child labor and farming area. Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means, respectively, that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. Climatic factors are rainfall, the number of idiosyncratic and covariate shocks, the mean temperature and the greenness increase in days.

same as that obtained from the Imperfect specification.

Based on this specification, we can test whether the Cobb-Douglas specification is rejected. This test amounts to evaluating the joint significance of the interacted terms between $\ln L_a$ and $\ln L_c$ (α_{aa}, α_{cc} and α_{ac} in Equation (6)), and the test is provided in the bottom part of Panel C. We clearly reject the Cobb-Douglas specification, as expected. An alternative specification with interacted terms between each labor variable and the other inputs has been estimated and provides similar results.

4.2 First-stage

Table 3: Effect of the number of children N_c on child labor (first stage)

Dependent variable	Perfect and imperfect		Translog	
	(1) L_c/L_a	(2) L_c/L_a	(3) $\log(L_c)$	(4) $\log(L_c)$
$\frac{N_c}{N_a}$	0.203*** (0.028)	0.187*** (0.031)		
N_c			0.559*** (0.060)	0.566*** (0.059)
R-squared	0.077	0.268	0.078	0.299
F-test	52.45	35.76	87.16	91.63
Household FE	×	×	×	×
Climatic factors	×		×	
Village-year FE		×		×
Observations	4912	4982	4914	4984

Note: The computation sample is made of households with positive adult labor, child labor and farming area. Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means, respectively, that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. Climatic factors are rainfall, the number of idiosyncratic and covariate shocks, the mean temperature and the greenness increase in days.

We now turn to the identification of child labor productivity based on the exogenous variation provided by children aging. We first check whether aging of the children in the pool provides sufficient explanatory power. From now on, we systematically control for household fixed effects. Columns (1) and (3) of Table 3 control for shocks, while columns (2) and (4) control for village-year fixed effects. We find that household composition predicts

well the labor supply, even when conditional on household and village-year fixed effects. In particular, if adults have one more child aged 10-15 years old, it increases child labor days by 60%. The F-statistics are high accordingly (ranging from 36 to 92, depending on specifications). We also attempt to predict total child labor using the number of children of each age in the pool, but doing so does not improve the F-statistics.

4.3 IV estimations

Table 4: Average semi-elasticities of labor, based on OLS and IV estimations

Variables	Perfect		Imperfect		Translog	
	OLS	IV	OLS	IV	OLS	IV
Panel A: Without village-year FE						
Adult	0.0029** (0.0011)	0.0053 (0.0054)	0.0066*** (0.0016)	0.0096* (0.0055)	0.0048*** (0.0010)	0.0048*** (0.0010)
Child	0.0029*** (0.0009)	0.0009 (0.0044)	0.0045*** (0.0013)	0.0021 (0.0044)	0.0045*** (0.0015)	0.0032 (0.0065)
$\hat{\eta}$		0.3874 (0.8711)		0.4612 (0.8658)		0.0242 (0.1226)
Climatic factors	×	×	×	×	×	×
Household FE	×	×	×	×	×	×
Computation sample	1458	1458	1458	1458	1458	1458
Panel B: With village-year FE						
Adult	0.0028** (0.0013)	0.0033 (0.0060)	0.0036 (0.0023)	0.0042 (0.0074)	0.0037*** (0.0012)	0.0037*** (0.0010)
Child	0.0026*** (0.0010)	0.0022 (0.0050)	0.0029* (0.0015)	0.0025 (0.0050)	0.0028 (0.0017)	0.0026 (0.0068)
$\hat{\eta}$		0.0744 (0.9751)		0.0922 (1.0017)		0.0035 (0.1293)
Household FE	×	×	×	×	×	×
Computation sample	1480	1480	1480	1480	1480	1480

Note: The estimation sample is made of households with positive adult labor and farming area while the computation sample is constrained to households with positive adult labor, child labor and farming area. Standard errors reported in parentheses are clustered at the ward level and are bootstrapped for IV estimates. ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. For the perfect and the imperfect specification, the IV is $\frac{N_c}{N_a}$. For the translog specification, the IV is N_c .

We now turn to the IV estimations. The full specifications are reported in Table A5. Table 4 reports tests of the specifications and the estimated labor semi-elasticities with IV (but it also reports OLS results for ease of comparison). The test of exogeneity is

simply given by the significance of the coefficient of $\hat{\eta}$. Quite surprisingly, we never reject the exogeneity hypothesis of child labor. However, we lose much in terms of the precision of the estimates, which become nonsignificant, including in the Translog specification, in which the F-statistic was particularly large. In the specification without village-year fixed effects, we find that the semi-elasticity of production to child labor is roughly divided by two when controlling for first-stage residuals. The estimates are consistent with each other in the various specifications: on average, one day of child labor increases production by 0.2%. In the second part of the panel, precision is again an issue, but the estimates are very close to those obtained in Panel A and to the OLS estimates. These OLS semi-elasticities of production to child labor range from 0.26% to 0.29%, which is a very small range. In comparison, the estimated adult labor semi-elasticities range from 0.28% to 0.37%. We must be more cautious about this last set of estimates, for which we have not properly addressed the question of endogeneity. However, the ratio between the two, ranging from 0.76 (Translog) to 0.93 (Perfect), is plausible and points to high efficiency of child labor, which might be driven by children working on-farm being only household children, while adult work encompasses a more diverse category: household males, households females and nonhousehold members. This last category of workers could have lower productivity due to a lack of incentives. In addition, adult labor includes all labor provided by individuals aged between 16 and 65 years old, and older workers are presumably less productive than younger ones.²¹

5 Heterogeneity in productivity

We now want to explore the heterogeneity in productivity by child characteristics, namely, age and gender.

²¹The tests for comparing the different functional forms provide the same results as previously: we reject the Perfect substitutes production function to the benefit of the Imperfect function, and we cannot reduce the Translog to a Cobb-Douglas.

5.1 Specification

We do so based on the Perfect specification. While we have shown that the data reject the hypothesis that both types of labor are perfect substitutes, we have also seen that the estimates are strikingly close from one functional form to the other. Testing for heterogeneity by children's characteristics is straightforward with the Perfect functional form.

Starting back from Equation (3) but allowing for a different γ_c by child's gender and then linearly approximating, we obtain:

$$\log Y = \log A + \alpha \log(L_a + \gamma_b L_b + \gamma_g L_g) + (\log X)\beta \quad (13)$$

$$\log Y \approx \log A + \alpha \log L_a + \alpha \gamma_b \frac{L_b}{L_a} + \alpha \gamma_g \frac{L_g}{L_a} + (\log X)\beta \quad (14)$$

where L_b stands for labor performed by boys and L_g for labor performed by girls.

For heterogeneity by age, we allow for a parsimonious specification according to which productivity depends linearly in age. For limited variations in age, this assumption should be reasonable.

$$\log Y = \log A + \alpha \log \left(L_a + \sum_{o=10}^{o=15} (\gamma_c + \gamma_{c'} \cdot (o - 15)) L_{co} \right) + (\log X)\beta \quad (15)$$

$$\log Y \approx \log A + \alpha \log L_a + \alpha \sum_{o=10}^{o=15} (\gamma_c + \gamma_{c'} \cdot (o - 15)) \frac{L_{co}}{L_a} + (\log X)\beta \quad (16)$$

where L_{co} stands for labor time provided by children of age o .

Given that the data do not reject the hypothesis of exogeneity in the allocation of child labor, we simply treat each of the child labor variables as exogenous.²²

²²We have also exploited the composition of the pool of children to instrument the child labor variables, but again, exogeneity is never rejected and leads to comparable, but less precise, results.

5.2 Results

Table 5 provides the semi-elasticities of production to child days of work by gender. We find that boys are markedly more productive than girls, which could be either due to differences in strength or due to girls spending only a proportion of their day in the field (because they have to also provide domestic work, for instance), while boys spend their full day on the plot. We do not have information about farming hours per day on an annual basis. However, we know the number of hours devoted to each activity in the week before the survey. Figure A3 provides the statistics for economically active children. Boys spend more than 18 hours per week on agricultural tasks, while girls provide only 13 hours. We do not have information about the number of *days* spent farming in the previous week, so we cannot compute a number of hours per day of work. However, we know that girls spend, on average, 54.3 days per year conditional on participation, compared to 48.9 days for boys. This small discrepancy suggests that girls indeed provide less farm labor per day of work. If parents allocate child working hours to equate the marginal productivity in various activities, then the productivity of boys is a better estimate of the full productivity of a child.

Table 5 provides the semi-elasticities by child age. We confirm that older children are more productive than younger ones. Fifteen year old children are three times more productive than 11 year old children, and the semi-elasticity for 10 year old children is not significantly different from 0, consistent with few children younger than 10 working on farms.

This finding also informs us about the validity of our instrumentation strategy. Indeed, entry in and exit out of the age range are also associated with changes in the other inputs if 1) children younger than 10 actually contribute; or 2) children older than 15 have different productivity than adults. Given that children younger than 10 have been shown not to be economically productive, 1) is excluded. We also provide in Table A6 a specification with

Table 5: Average semi-elasticities by gender and by age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gender					Age		
	Boy	Girl	10	11	12	13	14	15
Panel A: Without village-year FE								
	0.0046***	0.0015	0.0006	0.0012	0.0018	0.0024**	0.0029***	0.0035**
	(0.0014)	(0.0010)	(0.0036)	(0.0027)	(0.0018)	(0.0011)	(0.0009)	(0.0014)
Household FE	×	×	×	×	×	×	×	×
Climatic factors	×	×	×	×	×	×	×	×
Computation sample	1458	1458	1458	1458	1458	1458	1458	1458
Panel B: With village-year FE								
	0.0042***	0.0013	-0.0003	0.0004	0.0012	0.0019	0.0027***	0.0034*
	(0.0016)	(0.0014)	(0.0049)	(0.0036)	(0.0024)	(0.0013)	(0.0010)	(0.0020)
Household FE	×	×	×	×	×	×	×	×
Computation sample	1480	1480	1480	1480	1480	1480	1480	1480

Note: The estimation sample includes all households with adult labor while the simulation sample is constrained to households with adult and child labor. Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%.

labor provided by children younger than 10 years old. It confirms the results. We also find that the productivity of 15 year old children is not significantly different from adults' productivity, which excludes 2). The instruments therefore seem to satisfy the exclusion restriction.

6 How much should children be compensated?

We are now equipped with estimates of child labor productivity. Based on this information, we can compute two measures of interest. The most direct one is the children's opportunity cost of time, which is obtained by simply expressing these productivity estimates as the value for one day of labor. We can also infer how much would be needed to bring children to school, i.e., the amount of cash transfer that would be required to convince parents to enroll their children in school. To do so, we must additionally estimate the education demand, which comes with additional assumptions but provides a key figure for a potential economic policy.

Table 6: Average value of one day of labor in US\$, by gender and by age

	(1)	(2)	(3)	(4)	(5)	(6)
	Adult	Child	Girl	Boy		
	3.252***	0.890***	0.375	1.471***		
	(0.188)	(0.306)	(0.389)	(0.457)		
Age	10	11	12	13	14	15
	-0.0406	0.221	0.466	0.697*	0.915***	1.122**
	(1.682)	(1.175)	(0.746)	(0.412)	(0.311)	(0.501)
Household FE	×	×	×	×	×	×
Computation sample	1458	1458	1458	1458	1458	1458

Note: Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%.

The previous specifications have displayed remarkable homogeneity in results. Given that we want to keep a specification allowing for heterogeneity by children's characteristics, we simply use in this last section a Perfect production function. The OLS estimates of the Perfect specification with climatic factors are those obtained in Panel B of Table 4; we thus proceed with these estimates.²³

6.1 Child opportunity cost of time

We compute the value of one day of child labor in 2016 US\$. We find that the child productivity is roughly \$0.89.²⁴ Gertler and Glewwe (1992) computed the opportunity cost of education by differencing work duration between enrolled and nonenrolled children. In our data, nonenrolled children work 26 days more than enrolled children, on average. Therefore, the opportunity costs of education are roughly equal to \$23, which could be compensated by monthly payments of \$1.92.

As previously said, using the child wage on the market to assess the child time opportunity cost is difficult to defend. However, it is interesting to compare the adult agricultural wage with our estimates of adult productivity. In our data, wages can be reported per

²³The main discrepancy is the lower estimated productivity for adults.

²⁴The Translog specification, which provides the higher estimates, predicts that the child productivity is \$1.25.

day or per month at the discretion of the respondent. The average per day wage among adults who are paid by the day is 4,818 TSh (\$4.6).²⁵ This amount is slightly greater than most of the adult elasticities obtained in Table 6, which is expected since the existence of labor market imperfections would lead to a discrepancy between observed wages and shadow wages. In particular, household workers would presumably have difficulties being hired full time on the market, and considering the likelihood of remaining unemployed on a given day reduces the opportunity cost of time and increases on-farm employment. This discrepancy confirms that we cannot completely rely on child wages to inform us about child time opportunity costs, but the difference between adult wages and adults' estimated productivity suggests that our estimates are meaningful.

We can also compute the child earnings per day by age and gender. As previously discussed, boys have higher estimated productivity than girls, and 15 year old children are found to provide \$1.12 per day of work.

6.2 Simulation of the effect of a conditional cash transfer

Our simulation of the opportunity cost of time, however, does not lead to a clear policy recommendation. Indeed, how much is needed to convince parents to enroll their children in school also depends on the income elasticity of education demand, which has not yet been estimated. The (conditional) cash transfer to provide might well be lower than the opportunity cost if parents value education but simply cannot afford it, but it could also be higher if the parents dislike education and/or attach a specific value to child labor. We therefore perform a simulation exercise to evaluate the effect of a CCT. To do so, we replicate the simulation provided in Bourguignon et al. (2003), which is one of the few papers computing the ex-ante effects of a CCT. Their paper was applied to Brazil.

²⁵When performing a similar computation for adults paid by the month, we obtain a monthly average agricultural wage of 96,334 TSh, corresponding to 20 days of the daily wage. Therefore, there does not seem to be a striking difference between pay per day and pay per month for 20 days of work in a month, which seems sensible.

Compared to their method, we only change the simulation of child income: they relied on a Mincer equation function to predict unobserved child wages, while we rely on our production function (with heterogeneity in productivity by age). The rest of the specifications are similar.

The method boils down to estimating an occupational choice (S_{ci}) with 3 options as follows: 0 if the child does not attend school, 1 if the child goes to school and works on the farm and 2 if the child goes to school and does not work on the farm. The choice of child c in household i (S_{ci}) therefore takes on three values. When $S_{ci} = 0$, the child is assumed to work full time, when $S_{ci} = 1$, the child cannot provide as much economic labor since s/he also attends school, and when $S_{ci} = 2$, s/he might also be employed in domestic activities, along with school attendance. The child's occupation is chosen to maximize the utility (U_{ci}), and we assume that:

$$U_{ci}(j) = Z_{ci} \cdot \gamma_j + (Y_{-i} + y_{cij})\alpha_j + v_{cij} \quad (17)$$

where $j = 0, 1, 2$ are the alternatives, Y_{-i} is the total income of household members other than the child, and y_{cij} is the contribution of the child in alternative j , Z_{ci} , all of which are nonincome explanatory variables, while v_{cij} is a random variable that reflects unobserved preferences for schooling and labor. Once the parameters are estimated, we can simulate the occupation choice with a household income increased by the transfer amount, conditional on school enrollment (i.e., for alternatives 1 and 2) and on being in a poor household.

Appendix D provides more details about the method, the computation of the incomes, and the intermediate results for the parameters estimates. Table 7 provides the estimated effect of a monthly \$10 conditional cash transfer.²⁶ The effect of the transfer is small. The

²⁶Such a transfer targeted to children aged 10-15 years old and below the poverty line would cost approximately 0.6% and 0.8% of the GDP, respectively, depending on whether only rural areas or the whole territory is treated. In 2014, the government expenditures on education amounted to 3.48% of the GDP.

Table 7: Simulated effect of a conditional monthly transfer (\$10) on schooling and work status of children aged 10-15

Actual Status	New status			Total
	Work	Work and Attend	Attend	
All children				
Work	88.89	6.21	4.90	100.00
Work and Attend	0.00	99.74	0.26	100.00
Attend	0.00	0.00	100.00	100.00
Before the transfer	6.90	35.37	57.73	100.00
After the transfer	6.14	35.71	58.15	100.00
Only children below the poverty line				
Work	86.23	7.69	6.07	100.00
Work and Attend	0.00	99.75	0.25	100.00
Attend	0.00	0.00	100.00	100.00
Before the transfer	7.59	36.16	56.25	100.00
After the transfer	6.54	36.65	56.80	100.00

Note: The conditionality is based on schooling attendance and being below the poverty line.

bottom panel of the table states that, among the 7.59% of children who do not attend school and are below the poverty line, only 13.77% of them would switch to school enrollment (7.69% combining the two activities and 6.07% only attending school). As a result, the share of nonenrolled children diminishes only slightly to 6.54%. When using the whole sample, this targeted transfer has a smaller effect, as expected, since the transfer is conditional on poverty. This outcome is in stark contrast to the results obtained in Bourguignon et al. (2003), despite a very similar methodology. The corresponding figure for Bolsa Escola (first figure of bottom panel) is 41.3%, while the transfer per child is lower (2016\$ 6.67 per month). Of course, additional assumptions come with the simulation. In particular, we estimate the occupational choice model and its elasticity to income without considering potential endogeneity issues (as Bourguignon et al. (2003) did). However, instrumenting the income variable with a wealth index based on durable ownership does not change the results. One must recognize, however, that the conclusion that we reach in this section bears not only on our estimates of the child earnings but also on the validity of the occupation

choice specification.

It is interesting to try and understand the discrepancy between the two figures provided in this paper. On the one hand, we find that children make only a marginal contribution to their households (see Table A7) and that having them enrolled in school only costs, on average, \$26 per year. On the other hand, we find that, with a yearly CCT of \$120, only a small proportion of nonenrolled children would attend school. It therefore seems that households opting for child labor are reluctant to stop using it. How can we explain this finding? First, the average child productivity hides considerable heterogeneity. Examining the characteristics of the compliers and of the noncompliers is very informative. Table A8 displays the characteristics of the children who are never enrolled, who enroll due to the program and who are always enrolled. It shows that the children who comply are actually poorer and rely less on farming than the noncompliers. Actually, the estimated farm production of children who do not comply is \$213 per year, compared to \$46 for those who comply and \$80 for those who are enrolled, computed for a similar number of work days (full time). These outcomes clearly suggest that, for efficiency purposes, the transfer would have to be higher for children who have higher opportunity costs (more land area). This finding raises ethical issues. Interestingly, the land area is similar for the compliers and for the always-takers. The former, however, have fewer other sources of income. We also find that children who do not comply have left school for longer periods, limiting their possibility of returning to school.

Second, it might also be that parents value child labor per se, not only for the income it provides, which could be the case if child labor is actually an investment in specific human capital. Additionally, if parents anticipate higher transfers from their grown-up children if they continue farming (and remain close to their parents), they might decide not to enroll them in school. Both explanations fit the stylized fact that education demand is not very

sensitive to household income,²⁷, but in the absence of information on returns to investment in the two types of human capital, we cannot disentangle the two explanations.

This result also incidentally speaks to the debate between proponents and opponents of CCT schemes. Proponents argue that the conditionality allows sending more children to school and therefore reducing poverty in the long-run. Opponents argue that there are hidden costs associated with “forcing” the household to send their children to school and that some households might decide not to comply, hampering the fight against poverty. Baird et al. (2011) provided mixed evidence for the relative efficiency of conditional versus unconditional cash transfers. They also showed that, when there is a large proportion of non-compliers (with the CCT), it is best to opt for an unconditional cash transfer. Evaluating the impact of our hypothetical program on poverty reduction is beyond the scope of our paper, but it seems that, in our setting, an unconditional cash transfer would be preferable. Bauchet et al. (2018) reached a similar conclusion when observing that, in Bolivia, the most marginalized group of children was less likely to participate in educational CCTs. In our case as well, many children would not comply and belong to the poorer households. Of course, opting for an unconditional cash transfer incurs the expense of any improvement in schooling (see Table A11 for the simulated effect of an unconditional transfer), but one must remember that only 1% of children living in poverty change their schooling status due to CCTs.

7 Conclusion

In this paper, we estimate the economic contributions of children as unpaid family workers on Tanzanian farms. To do so, we exploit the panel dimension of the data to control for unobserved permanent household and time-varying village characteristics. We also

²⁷The estimated effect of household income on education demand is roughly half that obtained in Brazil by Bourguignon et al. (2003).

instrument child labor by the variation in the number of children of household members over time; however, exogeneity is never rejected. Our results are consistent despite the use of various functional forms and specifications. The results show that one day of labor performed by a child aged 10 to 15 years old increases production by roughly \$0.89.

This outcome informs us about the opportunity costs of children in agricultural households. To our knowledge, this paper is the first to provide an estimate of the child contribution per day of work for a setting reflecting the most common form of child labor in African countries.

We also explore the heterogeneity of child productivity by gender and age. It appears that children begin being productive from 12 years old and reach adult productivity by 15 years old. In addition, that boys spend more time working on farms and contribute less than girls to domestic activities could explain why their daily productivity is higher.

Considering the discrepancy in the amount of work that enrolled and nonenrolled children perform (26 days per year), we find that households should be compensated on average with \$1.92 per month for enrolling their children. However, when simulating a hypothetical CCT program, we find that even larger cash transfers (\$10 per month) would fail to bring nonenrolled children into school. The heterogeneity in productivity as well as the low sensitivity of education demand to income seems to explain this result. Our methodology and our findings are especially relevant when assessing the value of conditional cash transfer programs.

References

- Anker, R., Barge, S., Ashraf, S., and Levison, D.** (1998) Economics of child labour in india’s carpet industry, in R. Anker, S. Barge, S. Rajagopal, and M. Joseph (eds.), *Economics of Child Labour in Hazardous Industries of India*, Center for Operations Research and Training.
- Attanasio, O., Fitzsimons, E., Gomez, A., Gutierrez, M. I., Meghir, C., and Mesnard, A.** (2010) Children’s schooling and work in the presence of a conditional cash transfer program in rural colombia, *Economic development and cultural change*, **58**, 181–210.
- Baird, S., McIntosh, C., and Özler, B.** (2011) Cash or condition? evidence from a cash transfer experiment, *The Quarterly Journal of Economics*, **126**, 1709–1753, doi: 10.1093/qje/qjr032.
- Barrera-Osorio, F., Bertrand, M., Linden, L. L., and Perez-Calle, F.** (2011) Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia, *American Economic Journal: Applied Economics*, **3**, 167–95.
- Bauchet, J., Undurraga, E. A., Reyes-García, V., Behrman, J. R., and Godoy, R. A.** (2018) Conditional cash transfers for primary education: Which children are left out?, *World Development*, **105**, 1–12, doi:10.1016/j.worlddev.2017.1.
- Behrman, J. R., Parker, S. W., and Todd, P. E.** (2011) Do Conditional Cash Transfers for Schooling Generate Lasting Benefits?: A Five-Year Followup of PROGRESA/Oportunidades, *Journal of Human Resources*, **46**, 93–122.
- Benjamin, D.** (1992) Household composition, labor markets, and labor demand: testing for separation in agricultural household models, *Econometrica*, 287–322.

- Bhalotra, S. and Heady, C.** (2003) Child farm labour: the wealth paradox., *World Bank Economic Review*, **17**, 197–227.
- Bourguignon, F., Ferreira, F. H. G., and Leite, P. G.** (2003) Conditional Cash Transfers, Schooling, and Child Labor: Micro-Simulating Brazil’s Bolsa Escola Program, *World Bank Economic Review*, **17**, 229–254.
- Burbidge, J. B., Magee, L., and Robb, A. L.** (1988) Alternative transformations to handle extreme values of the dependent variable, *Journal of the American Statistical Association*, **83**, 123–127.
- Carpio, X. V. D., Loayza, N. V., and Wada, T.** (2016) The impact of conditional cash transfers on the amount and type of child labor, *World Development*, **80**, 33 – 47.
- Chavas, J.-P., Petrie, R., and Roth, M.** (2005) Farm household production efficiency: Evidence from the Gambia, *American Journal of Agricultural Economics*, **87**, 160–179.
- Chennareddy, V.** (1967) Production efficiency in south indian agriculture, *Journal of Farm Economics*, **49**, 816–820.
- Dammert, A. C.** (2009) Heterogeneous impacts of conditional cash transfers: Evidence from Nicaragua, *Economic Development and Cultural Change*, **58**, 53–83.
- de Hoop, J., Friedman, J., Kandpal, E., and Rosati, F. C.** (forthcoming) Child schooling and child work in the presence of a partial education subsidy, *Journal of Human Resources*, doi:10.3368/jhr.54.2.0317.8627R1.
- De Janvry, A., Fafchamps, M., and Sadoulet, E.** (1991) Peasant household behaviour with missing markets: Some paradoxes explained, *The Economic Journal*, **101**, 1400–1417.

- De Leon, J. and Parker, S.** (2000) The impact of anti-poverty programs on children's time use: The case of Progresa in Mexico, Working paper, Progresa.
- Diallo, Y., Etienne, A., and Mehran, F.** (2013) Global child labour trends 2008 to 2012, Tech. rep., International Labour Office, International Programme on the Elimination of Child Labour (IPEC), Geneva.
- Dumas, C.** (2007) Why do parents make their children work? A test of the poverty hypothesis in rural areas of Burkina Faso., *Oxford Economic Papers*, **59**, 301–329.
- Dumas, C.** (forthcoming) Productivity shocks and child labor: the role of credit and labor markets, *Economic Development and Cultural Change*, doi:10.1086/701828.
- Edmonds, E. V. and Schady, N.** (2012) Poverty alleviation and child labor, *American Economic Journal: Economic Policy*, **4**, 100–124.
- Edmonds, E. V. and Shrestha, M.** (2014) You get what you pay for: Schooling incentives and child labor, *Journal of Development Economics*, **111**, 196–211.
- Galiani, S. and McEwan, P. J.** (2013) The heterogeneous impact of conditional cash transfers, *Journal of Public Economics*, **103**, 85 – 96.
- Gertler, P. and Glewwe, P.** (1992) The willingness to pay for education for daughters in contrast to sons: Evidence from rural peru, *The World Bank Economic Review*, **6**, 171–188.
- Henningsen, A. and Henningsen, G.** (2012) On estimation of the CES production function - Revisited, *Economics Letters*, **115**, 67–69.
- IPEC** (2007) Child labour wages and productivity: Results from demand-side surveys, Tech. rep., International Labour Office.

- Jacoby, H. G.** (1993) Shadow wages and peasant family labour supply: An econometric application to the peruvian sierra, *The Review of Economic Studies*, **60**, 903–921.
- Kakwani, N., Soares, F. V., and Son, H. H.** (2005) Conditional cash transfers in African countries, international Policy Centre for Inclusive Growth Working Papers No. 9.
- Kmenta, J.** (1967) On estimation of the CES production function, *International Economic Review*, **8**, 180–189.
- Lambert, S. and Magnac, T.** (1997) Implicit prices and recursivity of agricultural households' decisions, Crest working paper no. 9731.
- Le, K. T.** (2009) Shadow wages and shadow income in farmers' labor supply functions, *American Journal of Agricultural Economics*, **91**, 685–696.
- Levison, D., Anker, R., Barge, S., and Ashraf, S.** (1998) Is child labour really necessary in india's carpet industry?, in R. Anker, S. Barge, S. Rajagopal, and M. Joseph (eds.), *Economics of Child Labour in Hazardous Industries of India*, Center for Operations Research and Training.
- MacKinnon, J. G. and Magee, L.** (1990) Transforming the dependent variable in regression models, *International Economic Review*, **31**, 315–39.
- Pence, K.** (2006) The Role of Wealth Transformations: An Application to Estimating the Effect of Tax Incentives on Saving, *The B.E. Journal of Economic Analysis & Policy*, **5**, 1–26.
- Ravallion, M. and Wodon, Q.** (2000) Does child labor displace child schooling? evidence on behavioral responses to an enrollment subsidy, *Economic Journal*, **110**, C158–C176.

- Safir, A.** (2009) Who leaves, who moves in? The impact of positive and negative income shocks on migration in Senegal, PSE Working Papers No. 2008-76.
- Schady, N. and Araujo, M. C.** (2006) Cash transfers, conditions, school enrollment, and child work : evidence from a randomized experiment in Ecuador.
- Singh, I., Squire, L., and Strauss, J.** (1986) *Agricultural Household Models*, chap. 10, Washington, D.C.: World Bank., 79–83.
- Skoufias, E.** (1994) Using shadow wages to estimate labor supply of agricultural households, *American Journal of Agricultural Economics*, **76**, 215–227.
- Skoufias, E., Parker, S. W., Behrman, J. R., and Pessino, C.** (2001) Conditional cash transfers and their impact on child work and schooling: Evidence from the progresas program in mexico [with comments], *Economia*, **2**, 45–96.
- Wooldridge, J.** (2015) Control Function Methods in Applied Econometrics, *Journal of Human Resources*, **50**, 420–445.

A Variables definition

Production: Total household production on farm. Each crop’s output is the output valued at the median price in the district. The value of the production is expressed in current Tanzanian shillings.

Area: Total surface area (in acres) of farmed plots.

Adult labor L_a : Number of farming days in the last year performed by adults aged 16 to 65 in the household.

Child labor L_c : Number of farming days in the last year performed by children aged 10 to 15 in the household.

Organic fertilizer: Dummy variable equal to 1 if the household uses organic fertilizer for at least one plot.

Inorganic fertilizer: Total spending (in Tanzanian shillings) on inorganic fertilizer during the year of the survey.

Pesticide: Total spending (in Tanzanian shillings) on pesticide during the year of the survey.

Rainfall: We use data from the NOAA to compute a standardized rainfall index from July to June of the year prior to the survey: http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.CPC/.FEWS/.Africa/.DAILY/.ARC2/.daily/.est_prdp/datafiles.html.

Number of idiosyncratic shocks: Number of productivity shocks (drought, flood, crop disease, crop pests and severe water shortages) that severely affected the household and/or some households in the village in the year of the survey.

Number of covariate shocks: Number of productivity shocks (drought, flood, crop disease, crop pests and severe water shortages) that severely affected most or all households in the community in the year of the survey.

Temperature: Annual mean temperature in degrees Celsius, multiplied by 10.

Greenness increase: Average change in greenness (integral of daily EVI values) during growing season.

Productive assets: Based on the ownership status of 25 farming assets (carts, tractors, harvesting machine, etc.), we adopt Principal Component Analysis (PCA) to construct a productive asset index.

Land erosion: Dummy variable equal to 1 if the household had an erosion issue for at least one plot during the year of the survey.

Land irrigation: Dummy variable equal to 1 if the household has at least one irrigated plot.

Gender of the household head: Dummy variable equal to 1 if the household head is a woman.

Crops: Variables equal to the percentage of maize, paddy, sorghum, cassava, sweet potatoes, beans, cowpeas, sunflowers, groundnuts and cotton on the farm (these crops constitute the ten main crops in Tanzania).

B Semi-elasticities with function $\log^M$

Dependent variable

Equations (4), (5) and (6) describe the functional forms used in this paper (respectively, for the specifications of Perfect, Imperfect and Translog). These equations follow the form $\log Y = f(L_c, L_a, X)$. We estimate these equations with the function $\log^M$ since we need to consider the observations that report an absence of output: $\log^M Y = f(L_c, L_a, X)$.

This decision affects marginally the interpretation of the semi-elasticities. Indeed, the monetary unit of Tanzania is small ($1\$ \approx 1,000\text{TSh}$) and accordingly, $\log^M Y \approx \log Y$ as soon as $Y > 0$. In addition, we write the fundamental model $Y = u \exp f(L_c, L_a, X)$, where $u \geq 0$ is a multiplicative error term independent from L_c, L_a, X .

$$\log^M(Y) \approx f(L_c, L_a, X) + \log u \text{ when } u > 0 \quad (18)$$

$$= 0 \text{ when } u = 0 \quad (19)$$

$$E(\log^M(Y)|L_c, L_a, X) \approx P(u > 0)E(\log Y|u > 0, L_c, L_a, X) \quad (20)$$

The semi-elasticities (7) and (9) are based on the difference in $\log Y$ between two situations with and without labor. The corresponding differences in $\log^M(Y)$ are differences in $\log Y$ multiplied by the chance of having production. In other words, we measure an additional production per day as a percentage multiplied by the likelihood of having production.

Explanatory variables

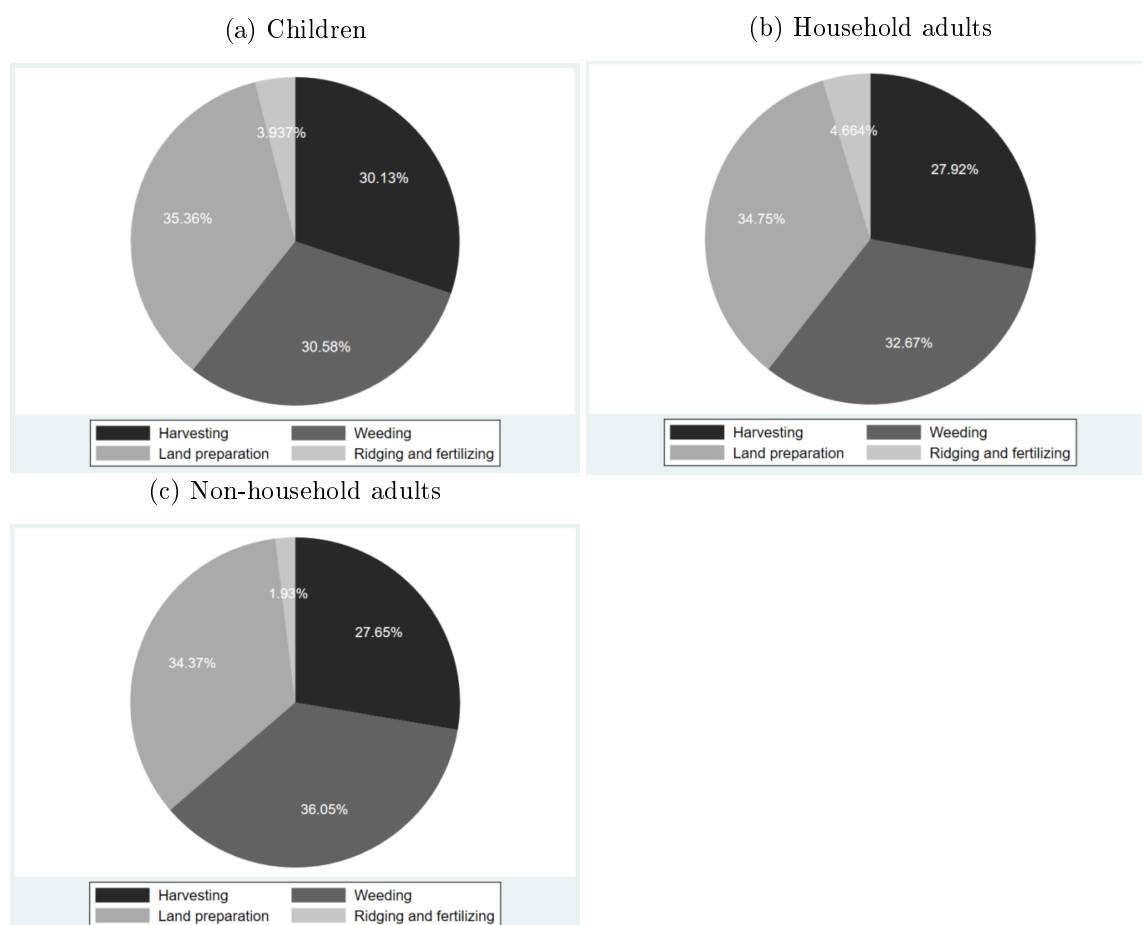
The semi-elasticities (7) and (9) depend on the expected production for two different levels of L_c, L_a . When log enters the estimated equations, we use the function $\log^M$ in the estimation and in the computations following the estimation; hence, the simulated function is the same as the estimated function.

C Additional tables and figures

Table A1: Descriptive statistics on the sample

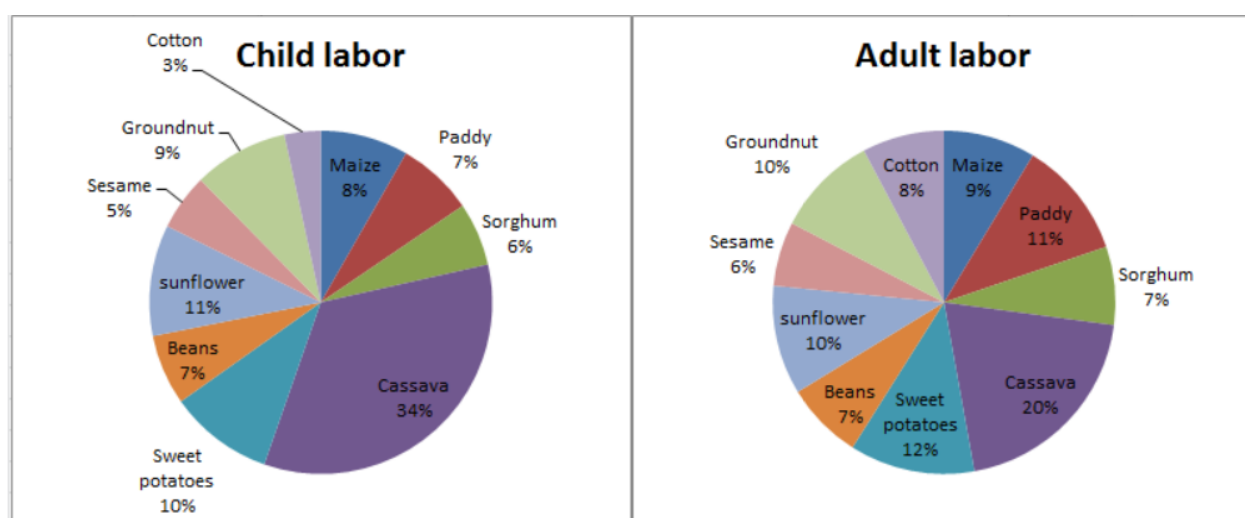
	(1) N	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
$\log^M(\text{production}) \approx \log(\text{production})$	4986	11.12	3.69	-0.69	18.16
$\log^M(\text{area}) \approx \log(\text{area})$	4986	1.16	0.97	-0.69	6.41
$\log^M(L_a) \approx \log(L_a)$	4986	4.88	0.96	0.75	7.90
$\log^M(L_c) \approx \log(L_c)$	4986	0.52	1.98	-0.69	6.47
N_c	4986	1.05	1.16	0	16
N_a	4986	3.49	2.25	0	28
Organic fertilizer	4986	0.22	0.41	0	1
$\log(\text{inorganic fertilizer})$	4986	1.30	4.39	-0.69	16.05
$\log(\text{pesticide})$	4986	0.87	3.73	-0.69	13.76
Number of idiosyncratic shocks	4986	0.07	0.28	0	3
Number of covariate shocks	4986	0.27	0.51	0	3
Land erosion	4985	0.17	0.37	0	1
Land irrigation	4985	0.03	0.17	0	1
Rainfall	4986	-0.02	0.81	-2.08	2.66
Temperature	4915	227.38	26.85	146	278
Greenness increase	4944	129.26	18.76	-12.8	166
Age of household head	4986	48.84	14.46	19	107
Productive assets	4986	0.08	1.84	-0.34	113.54
Percentage of maize	4986	0.45	0.38	0	1
Percentage of paddy	4986	0.10	0.23	0	1
Percentage of sorghum	4986	0.05	0.17	0	1
Percentage of cassava	4986	0.02	0.11	0	1
Percentage of sweet potatoes	4986	0.03	0.12	0	1
Percentage of beans	4986	0.14	0.27	0	1
Percentage of cowpeas	4986	0.03	0.13	0	1
Percentage of sunflower	4986	0.03	0.13	0	1
Percentage of groundnuts	4986	0.05	0.16	0	1
Percentage of maize	4986	0.02	0.10	0	1

Figure A1: Activities in agricultural work, by category of worker



Note: Distribution of total labor days by activity. Data for 2012.

Figure A2: Labor distribution among main crops



Note: Distribution of total labor days by crop, only plots without intercropping

Table A2: Perfect substitutes: OLS estimation

	(1)	(2)	(3)	(4)
$\log L_a$	0.6628*** (0.0711)	0.8024*** (0.0933)	0.7903*** (0.0903)	0.7394*** (0.1014)
$\frac{L_c}{L_a}$	0.8140*** (0.1254)	0.5695*** (0.1587)	0.5310*** (0.1569)	0.4787*** (0.1835)
$\log(\text{area})$	1.0828*** (0.0730)	0.6786*** (0.1040)	0.6808*** (0.1046)	0.5330*** (0.1096)
Organic fertilizer	0.4186*** (0.1260)	0.1137 (0.1400)	0.1185 (0.1395)	0.2423 (0.1695)
$\log(\text{inorganic fertilizer})$	0.1095*** (0.0112)	0.0659*** (0.0176)	0.0662*** (0.0177)	0.0694*** (0.0203)
$\log(\text{pesticide})$	0.0234 (0.0162)	0.0291** (0.0144)	0.0286* (0.0146)	0.0154 (0.0167)
Number of idiosyncratic shocks	-0.0206 (0.1621)	0.0066 (0.2080)	0.0474 (0.2060)	-0.1401 (0.2280)
Number of covariate shocks	-0.1272 (0.0791)	-0.1845** (0.0883)	-0.1923** (0.0891)	-0.0071 (0.1012)
Land erosion	0.2255* (0.1179)	-0.0781 (0.1219)	-0.0696 (0.1236)	-0.0392 (0.1496)
Land irrigation	0.6326** (0.2498)	0.5095 (0.3904)	0.5241 (0.3875)	0.3514 (0.4083)
Productive assets	0.0034 (0.0113)	-0.0052 (0.0118)	0.0112 (0.0125)	0.0344 (0.0604)
Rainfall			0.0553 (0.0734)	
Temperature			0.0410 (0.0439)	
Greenness increase			0.0153** (0.0069)	
R-squared	0.380			
Within household R-squared		0.233	0.231	0.466
Household FE		×	×	×
Village-year FE				×
Observations	4984	4984	4914	4984

Sample: The estimation sample is made of households with positive adult labor and farming area. Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. We also control for the crops variables, the household head's gender and age, the use of plot GPS information, survey month and year dummies.

Table A3: Imperfect substitutes: OLS estimation

	(1)	(2)	(3)	(4)
$\log L_a$	0.6587*** (0.0711)	0.8009*** (0.0932)	0.7886*** (0.0902)	0.7394*** (0.1014)
$\frac{L_c}{L_a}$	0.5175*** (0.1444)	0.2686* (0.1495)	0.2429 (0.1481)	0.4235* (0.2176)
$\frac{L_c}{L_a} * (1 - \log \frac{L_c}{L_a})$	0.5069** (0.2064)	0.5653** (0.2301)	0.5459** (0.2309)	0.1067 (0.3228)
$\log(\text{area})$	1.0795*** (0.0730)	0.6741*** (0.1039)	0.6769*** (0.1046)	0.5317*** (0.1093)
Organic fertilizer	0.4131*** (0.1260)	0.1191 (0.1398)	0.1242 (0.1391)	0.2432 (0.1697)
$\log(\text{inorganic fertilizer})$	0.1094*** (0.0111)	0.0657*** (0.0176)	0.0659*** (0.0177)	0.0694*** (0.0203)
$\log(\text{pesticide})$	0.0239 (0.0161)	0.0289** (0.0144)	0.0285* (0.0146)	0.0155 (0.0167)
Number of idiosyncratic shocks	-0.0207 (0.1616)	0.0047 (0.2075)	0.0454 (0.2055)	-0.1404 (0.2279)
Number of covariate shocks	-0.1309* (0.0791)	-0.1897** (0.0885)	-0.1976** (0.0893)	-0.0081 (0.1012)
Land erosion	0.2253* (0.1173)	-0.0748 (0.1220)	-0.0662 (0.1236)	-0.0389 (0.1497)
Land irrigation	0.6366** (0.2493)	0.5189 (0.3873)	0.5332 (0.3848)	0.3538 (0.4065)
Productive assets	0.0034 (0.0113)	-0.0053 (0.0117)	0.0109 (0.0125)	0.0342 (0.0607)
Rainfall			0.0552 (0.0732)	
Temperature			0.0413 (0.0437)	
Greenness increase			0.0151** (0.0069)	
R-squared	0.381			
Within household R-squared		0.234	0.231	0.466
Household FE		×	×	×
Village-year FE				×
Observations	4984	4984	4914	4984

Sample: The estimation sample is made of households with positive adult labor and farming area. Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. We also control for the crops variables, the household head's gender and age, the use of plot GPS information, survey month and year dummies.

Table A4: Translog: OLS estimation

	(1)	(2)	(3)	(4)
$\log L_a$	1.1940*** (0.3700)	1.4674*** (0.5022)	1.4126*** (0.5215)	1.5193*** (0.5291)
$\log L_c$	0.6496*** (0.1203)	0.5680*** (0.1442)	0.5300*** (0.1439)	0.5562*** (0.1560)
$(\log L_a)^2$	-0.0586 (0.0372)	-0.0713 (0.0504)	-0.0665 (0.0524)	-0.0837 (0.0535)
$(\log L_c)^2$	-0.0070 (0.0152)	-0.0285 (0.0184)	-0.0306* (0.0184)	-0.0107 (0.0205)
$\log L_c \cdot \log L_a$	-0.1000*** (0.0243)	-0.0719*** (0.0272)	-0.0639** (0.0265)	-0.0893*** (0.0303)
$\log(\text{area})$	1.0791*** (0.0730)	0.6667*** (0.1040)	0.6673*** (0.1048)	0.5220*** (0.1099)
Organic fertilizer	0.3864*** (0.1263)	0.0991 (0.1396)	0.1012 (0.1389)	0.2373 (0.1704)
$\log(\text{inorganic fertilizer})$	0.1082*** (0.0110)	0.0665*** (0.0178)	0.0667*** (0.0178)	0.0701*** (0.0203)
$\log(\text{pesticide})$	0.0262 (0.0161)	0.0327** (0.0144)	0.0323** (0.0146)	0.0188 (0.0167)
Number of idiosyncratic shocks	-0.0463 (0.1607)	-0.0186 (0.2078)	0.0224 (0.2059)	-0.1618 (0.2278)
Number of covariate shocks	-0.1200 (0.0789)	-0.1774** (0.0883)	-0.1840** (0.0891)	-0.0063 (0.1012)
Land erosion	0.2266* (0.1162)	-0.0875 (0.1222)	-0.0791 (0.1238)	-0.0456 (0.1500)
Land irrigation	0.6230** (0.2506)	0.5434 (0.3801)	0.5566 (0.3778)	0.3906 (0.4012)
Productive assets	0.0066 (0.0127)	-0.0046 (0.0120)	0.0108 (0.0126)	0.0366 (0.0634)
Rainfall			0.0636 (0.0725)	
Temperature			0.0465 (0.0448)	
Greenness increase			0.0142** (0.0069)	
R-squared	0.383			
Within household R-squared		0.237	0.235	0.469
Household FE		×	×	×
Village-year FE				×
Observations	4984	4984	4914	4984

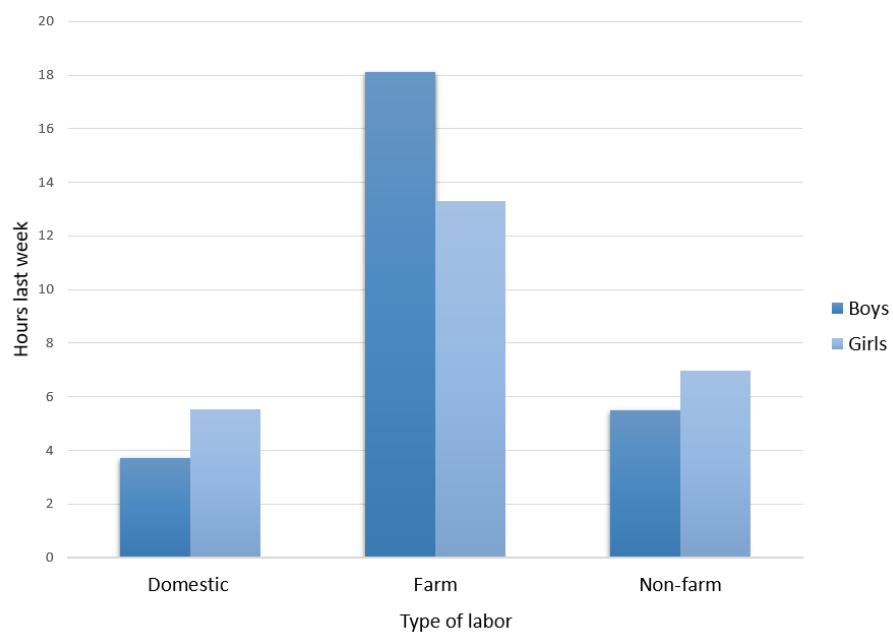
Sample: The estimation sample is made of households with positive adult labor and farming area. Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. We also control for the crops variables, the household head's gender and age, the use of plot GPS information, survey month and year dummies.

Table A5: IV estimations of the production functions

	Perfect		Imperfect		Translog	
	(1)	(2)	(3)	(4)	(5)	(6)
$\log L_a$	0.7760*** (0.0880)	0.7358*** (0.0974)	0.7715*** (0.0881)	0.7349*** (0.0979)	1.4176*** (0.5142)	1.5200*** (0.4036)
$\frac{L_c}{L_a}$	0.1657 (0.7336)	0.4078 (0.9208)	-0.1977 (0.7941)	0.3343 (1.0244)		
$\frac{L_c}{L_a} * (1 - \log \frac{L_c}{L_a})$			0.5569** (0.2618)	0.1094 (0.3531)		
$(\log L_a)^2$					-0.0667 (0.0511)	-0.0838** (0.0421)
$\log L_c$					0.5085** (0.2046)	0.5531*** (0.1998)
$(\log L_c)^2$					-0.0305 (0.0218)	-0.0106 (0.0168)
$\log L_c \cdot \log L_a$					-0.0642** (0.0259)	-0.0893*** (0.0298)
$\log(\text{area})$	0.6877*** (0.0985)	0.5350*** (0.0949)	0.6849*** (0.0988)	0.5342*** (0.0946)	0.6706*** (0.1169)	0.5229*** (0.0951)
Organic fertilizer	0.1187 (0.1464)	0.2423 (0.1723)	0.1246 (0.1455)	0.2432 (0.1728)	0.1036 (0.1408)	0.2377 (0.1737)
$\log(\text{inorganic fertilizer})$	0.0659*** (0.0193)	0.0694*** (0.0187)	0.0657*** (0.0193)	0.0694*** (0.0186)	0.0668*** (0.0169)	0.0701*** (0.0188)
$\log(\text{pesticide})$	0.0308** (0.0153)	0.0160 (0.0153)	0.0310** (0.0154)	0.0161 (0.0153)	0.0327** (0.0155)	0.0188 (0.0155)
Number of idiosyncratic shocks	0.0498 (0.2128)	-0.1396 (0.2164)	0.0483 (0.2122)	-0.1397 (0.2161)	0.0246 (0.2041)	-0.1613 (0.2153)
Number of covariate shocks	-0.1963** (0.0890)	-0.0072 (0.0894)	-0.2025** (0.0898)	-0.0084 (0.0895)	-0.1848** (0.0868)	-0.0062 (0.0907)
Land erosion	-0.0609 (0.1347)	-0.0369 (0.1353)	-0.0565 (0.1352)	-0.0365 (0.1353)	-0.0780 (0.1217)	-0.0456 (0.1370)
Land irrigation	0.5177 (0.3831)	0.3474 (0.3374)	0.5259 (0.3795)	0.3491 (0.3364)	0.5513 (0.3581)	0.3894 (0.3352)
Productive assets	0.0119 (0.0516)	0.0354 (0.0581)	0.0117 (0.0515)	0.0355 (0.0579)	0.0113 (0.0558)	0.0371 (0.0638)
Rainfall	0.0557 (0.0779)		0.0557 (0.0778)		0.0633 (0.0808)	
Temperature	0.0155** (0.0074)		0.0153** (0.0074)		0.0145** (0.0065)	
Greenness increase	0.0410 (0.0449)		0.0412 (0.0447)		0.0458 (0.0421)	
$\hat{\eta}$	0.3874 (0.8711)	0.0744 (0.9751)	0.4612 (0.8658)	0.0922 (1.0017)	0.0242 (0.1226)	0.0035 (0.1293)
Household FE	×	×	×	×	×	×
Village-year FE		×		×		×
R-squared	0.091	0.075	0.091	0.075	0.096	0.080
Observations	4986	4986	4986	4986	4914	4984

Sample: The estimation sample is made of households with positive adult labor and farming area. Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%. We also control for the crops variables, the household head's gender and age, the use of plot GPS information, survey month and year dummies.

Figure A3: Working hours by activity in the last week, among children who work



Sample: Children who have an economic activity and who live in a farm household.

Table A6: Average semi-elasticities by age for children 5 to 15 years old

Age	(1)	(2)
	OLS	
5	0.000656 (0.00350)	0.00214 (0.00454)
6	0.000934 (0.00306)	0.00221 (0.00394)
7	0.00121 (0.00263)	0.00228 (0.00335)
8	0.00149 (0.00220)	0.00235 (0.00277)
9	0.00177 (0.00179)	0.00242 (0.00220)
10	0.00204 (0.00140)	0.00248 (0.00166)
11	0.00232** (0.00107)	0.00255** (0.00121)
12	0.00260*** (0.000860)	0.00262*** (0.000944)
13	0.00288*** (0.000863)	0.00269*** (0.00104)
14	0.00315*** (0.00108)	0.00276* (0.00142)
15	0.00343** (0.00141)	0.00283 (0.00192)
Adult 16-65	0.00209 (0.00179)	0.00250 (0.00250)
Household FE	×	×
Climatic factors	×	
Village-year FE		×
Computation sample	1545	1569

Note: The estimation sample is made of households with positive adult labor and farming area while the computation sample is constrained to households with positive adult labor, child labor and farming area. Standard errors are clustered at the ward level and are reported in parentheses. ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%.

D Simulation exercise

We provide a brief description of the simulation exercise. More details and justifications can be found in Bourguignon et al. (2003).

D.1 Detail of the simulation method

The objective is to estimate an occupation choice model, as in Equation 17, and then simulate the effect of the CCT. To do so, we must obtain estimates of household and child incomes in the various alternatives.

If y_{ci0} denotes the child's contribution to the household income in alternative 0; then, the child's contribution in the other alternatives j is:

$$y_{ci1} = My_{ci0}; y_{ci2} = Dy_{ci0} \quad (21)$$

where M can be observed, and D is not. M reflects the difference in child income when s/he attends school instead of only working. D reflects the difference in child contribution when s/he does not perform farm work, instead of only working. Such a contribution is expected if children engage in domestic activities, for instance.

Plugging in these incomes into the utilities under the various alternatives (eq. 17) yields:

$$U_{ci}(j) = Z_{ci} \cdot \gamma_j + \alpha_j Y_{-i} + \beta_j y_{ci0} + v_{cij} \quad (22)$$

with $\beta_0 = \alpha_0; \beta_1 = \alpha_1 M; \beta_2 = \alpha_2 D$. Bourguignon et al. (2003) showed that a multilogit model identifies α s and D .

Once these parameters are obtained, we can simulate a conditional cash transfer, targeted on people below the poverty line (Y^0).²⁸ The simulation simply consists of assessing

²⁸This process also requires drawing residuals for the individuals that are consistent with the observed

the alternative that results in the greatest utility when a transfer T is given to the household, under the condition that the child is enrolled in school (alternatives 1 and 2).²⁹

$$\begin{aligned}
U_{ci0} &= Z_{ci} \cdot \gamma_0 + \alpha_0 Y_{-i} + \beta_0 y_{ci0} + v_{ci0} \\
U_{ci1} &= Z_{ci} \cdot \gamma_1 + \alpha_1 (Y_{-i} + T) + \beta_1 y_{ci0} + v_{ci1} & \text{if } Y_{-i} + M y_{ci0} \leq Y^0 \\
U_{ci1} &= Z_{ci} \cdot \gamma_1 + \alpha_1 Y_{-i} + \beta_1 y_{ci0} + v_{ci1} & \text{if } Y_{-i} + M y_{ci0} > Y^0 \\
U_{ci2} &= Z_{ci} \cdot \gamma_2 + \alpha_2 (Y_{-i} + T) + \beta_2 y_{ci0} + v_{ci2} & \text{if } Y_{-i} + M y_{ci0} \leq Y^0 \\
U_{ci2} &= Z_{ci} \cdot \gamma_2 + \alpha_2 Y_{-i} + \beta_2 y_{ci0} + v_{ci2} & \text{if } Y_{-i} + M y_{ci0} > Y^0
\end{aligned}$$

D.2 Computation of child and adult incomes

The child contribution is determined by two components: the intensity of work (number of days per year); and the productivity of work (obtained from the production function). The intensity of work depends on whether the child combines work with schooling or not (M parameter). We first identify this parameter, along with the estimate of the “full working time”, i.e., the average number of work days for children only engaged in economic activity. We estimate how child farming days change with age and enrollment in school with the following Equation:

$$L_{ci} = \mu_{age_{ci}} + \nu 1(S_{ci} = 1) + \xi_{age_{ci}} \cdot 1(S_{ci} = 1) + \epsilon_{ci} \quad (23)$$

performed for children who work ($S_{ci} \leq 1$). This specification provides M^{30} and the “full working time” ($\overline{L_c}$)³¹, which varies with age.

We now turn to the estimation of the child’s full-time income. We use the Perfect choice.

²⁹In the absence of information about actual attendance, we assume that children who are enrolled actually attend.

³⁰ $M = \frac{\hat{\mu}_{age} + \hat{\nu} + \hat{\xi}_{age}}{\hat{\mu}_{age}}$.

³¹ $\overline{L_c} = \hat{\mu}_{age_{ci}}$

Substitutes production function:

$$\log Y_{it} = \alpha \log L_{ait} + \alpha \gamma_c \frac{L_{cit}}{L_{ait}} + (\log X_{it})\beta + \mu_i + \zeta_{vt} + \varepsilon_{it} \quad (24)$$

The household production absent his/her contribution is estimated to be:

$$Y_{0ict} = \exp\left(\log(Y_{it}) - \hat{\alpha}\hat{\gamma}_c \frac{L_{cit}}{L_{ait}}\right) \quad (25)$$

where L_{cit} here stands for the labor of the child under scrutiny (keeping constant the labor provided by the other household children). The household production, if the child was contributing his/her full-time $\overline{L}_c$, would be:

$$Y_{fict} = \exp\left(\log(Y_{it}) - \hat{\alpha}\hat{\gamma}_c \frac{L_{cit}}{L_{ait}} + \hat{\alpha}\hat{\gamma}_c \frac{\overline{L}_c}{L_{ait}}\right) \quad (26)$$

The difference between the two is the full-time contribution of the child to the household income:

$$y_{ci0} = \exp\left(\log(Y_{it}) - \hat{\alpha}\hat{\gamma}_c \frac{L_{cit}}{L_{ait}} + \hat{\alpha}\hat{\gamma}_c \frac{\overline{L}_c}{L_{ait}}\right) - \exp\left(\log(Y_{it}) - \hat{\alpha}\hat{\gamma}_c \frac{L_{cit}}{L_{ait}}\right) \quad (27)$$

The definition of the child's full-time contribution does not depend on the actual child labor supply but solely on his/her age (which determines how much time can be allocated to the farm) and on the household production characteristics (adult labor inputs, farm size, flexible inputs, and household unobserved fixed productivity components). Landless households are incorporated into the computation sample with a child potential contribution equal to 0.

The adult income is made up of not only farm production but also income obtained from other activities. However, the survey does not allow to us to compute these additional

incomes because we do not have the information on an annual basis. As a result, we must use the household consumption measure, from which we subtract the actual child contribution Y_{0ict} .

D.3 Descriptive statistics on the simulation sample

Table A7 provides descriptive statistics on the simulation sample.

Table A7: Descriptive statistics on the simulation sample

	Mean	Std. Dev.	Min	Max
Household consumption minus child contribution (Y_{-i})	2565.59	2236.24	-1375.97	27935.29
Agricultural production	477.88	2107.35	0	51925.14
Estimated production of other household members	611.12	3380.24	0	84887.25
Estimated production of the child if working full time (y_{ci0})	86.38	501.25	0	25510.36
Estimated child work days, when working full time ($\bar{L}_c$)	51.73	4.61	45.42	59.01
Estimated production of the child (Y_{0ict})	18.57	66.27	0	2152.02
Estimated production of the child, conditional on participation	46.86	98.78	0	2152.02
Work days of the child (L_{cit})	16.60	42.51	0	645
Work days of the child, conditional on participation	41.89	59.16	1	645
Cultivated area	7.05	21.23	0	605
Age	12.32	1.70	10	15
Boys	0.50	0.50	0	1
Last grade	4.52	2.26	0	20
Experience ²	6.97	12.95	0	81
Age of household head	49.98	12.54	19	107
Education of household head	5.08	3.80	0	20
Number of children below 7	1.13	1.54	-2	17
Number of children 7 to 9	1.19	1.07	0	8
Number of children 10 to 15	2.12	1.16	1	10
Number of adults below 65	3.37	2.14	1	24
Number of adults 65 and more	0.18	0.44	0	2

Note: The simulation sample includes all households used in the estimation and the landless households. The variables used in the education demand and in the simulations are: household consumption minus child contribution and estimated production of the child if s/he was working full time. The other variables are given for the discussion. The statistics are computed on 4728 observations, except when conditional on labour participating children (1874 children). The monetary amounts are expressed in current 2016 \$ and all the variables are annual.

Table A8: Descriptive statistics depending on the education status before and after the transfer

	Never Enrolled		Enrolled after the transfer		Always Enrolled	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Household consumption minus child contribution (Y_{-i})	2000.19	1463.79	1365.78	483.35	2631.38	2275.69
Agricultural production	375.35	540.41	247.58	248.05	479.34	2096.34
Estimated production of other household members	415.24	624.25	296.40	328.71	611.81	3357.45
Estimated production of the child if working full time (y_{cit0})	212.88	1631.52	45.71	40.91	79.73	331.86
Estimated child work days, when working full time ($\bar{L}_c$)	55.16	3.98	54.21	4.29	51.47	4.53
Estimated production of the child (Y_{6ict})	68.53	105.99	39.68	48.31	16.44	63.85
Estimated production of the child, conditional on participation	68.53	105.99	39.68	48.31	43.26	97.85
Work days of the child (L_{cit})	56.68	68.99	46.47	49.52	14.90	40.02
Work days of the child, conditional on participation	56.68	68.99	46.47	49.52	39.22	57.13
Cultivated area	7.32	20.50	5.00	4.70	7.05	21.86
Age	13.58	1.47	13.24	1.58	12.23	1.67
Boys	0.57	0.50	0.56	0.50	0.49	0.50
Last grade	2.47	2.84	3.91	3.05	4.87	1.95
Experience ²	32.40	25.82	14.85	14.27	3.83	5.94
Age of household head	50.22	12.38	51.06	13.31	50.01	12.52
Education of household head	3.99	3.13	3.62	3.38	5.27	3.83
Number of children below 7	1.11	1.11	0.94	0.92	1.10	1.52
Number of children 7 to 9	1.28	0.97	1.15	0.82	1.16	1.06
Number of children 10 to 15	1.97	0.92	1.94	0.92	2.13	1.16
Number of adults below 65	2.88	1.28	2.85	1.21	3.39	2.16
Number of adults 65 and more	0.16	0.42	0.18	0.46	0.18	0.45
Observations	272		34		4127	

Note: The simulation sample includes all households used in the estimation and the landless households. The monetary amounts are expressed in current 2016 \$ and all the variables are annual.

D.4 Structural parameters of the occupation choice

Table A9 provides the multilogit estimates. Table A10 provides the values of the structural parameters for the occupational choice model, based on the multilogit estimation. Table A11 provides the mobility matrix if the transfer was implemented without school enrollment conditions.

Table A9: Occupational Choice Models for children aged 10-15 (reference: Only work)

	Work and Attend School	Attend School
Y_{-i}	0.0001** (0.0001)	0.0002*** (0.0001)
$y_{ci,j}$	-0.0004** (0.0002)	-0.0002** (0.0001)
Age of household head	-0.0123 (0.0089)	-0.0065 (0.0088)
Education of household head	0.0130 (0.0288)	0.0138 (0.0307)
Number of children below 7	-0.0332 (0.0800)	-0.0365 (0.0822)
Number of children 7 to 9	-0.0285 (0.0812)	-0.0931 (0.0818)
Number of children 10 to 15	-0.0575 (0.1224)	-0.1162 (0.1298)
Number of adults below 65	0.0364 (0.0813)	0.3073*** (0.0796)
Number of adults 65 and more	0.2055 (0.2419)	0.2153 (0.2416)
Birth order	0.4520** (0.1892)	0.5363*** (0.1919)
Boy	0.1955 (0.1827)	0.1781 (0.1817)
Last grade	0.5731*** (0.1748)	0.4935*** (0.1746)
Experience ²	-0.0516** (0.0243)	-0.0550** (0.0258)
Pseudo R-squared	0.183	
Observations	4433	4433

Additional controls: Year of the panel and age dummies. Experience measures the time since the child left school (and zero if the child is still enrolled in school). ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%.

Table A10: Structural parameters estimates in occupational choice model, children aged 10 to 15

α_0	α_1	α_2	M	D	β_0	β_1	β_2
0.0019*** (0.0007)	0.0020*** (0.0007)	0.0021*** (0.0007)	0.7314	0.8129*** (0.0418)	0.0019*** (0.0007)	0.0015*** (0.0005)	0.0017*** (0.0006)

Note: ***, **, * means respectively that the coefficient is significantly different from 0 at the level of 1%, 5% and 10%.

Table A11: Simulated effect of an unconditional monthly transfer (\$10) on schooling and work status of children aged 10-15

Actual Status	New status			Total
	Work	Work and Attend	Attend	
All households				
Work	99.35	0.65	0.00	100.00
Work and Attend	0.00	99.74	0.26	100.00
Attend	0.00	0.00	100.00	100.00
Before the transfer	6.90	35.37	57.73	100.00
After the transfer	6.86	35.33	57.82	100.00
Below the poverty line				
Work	99.19	0.81	0.00	100.00
Work and Attend	0.00	99.75	0.25	100.00
Attend	0.00	0.00	100.00	100.00
Before the transfer	7.59	36.16	56.25	100.00
After the transfer	7.53	36.13	56.34	100.00

Note: The conditionality is only based on being below the poverty line.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

