Child Labor and Market Access: Identifying Excluded Households

Samia Badji*

January 2016

Abstract

The paper aims at identifying households making their children work more because their access is constrained on the land and the labor markets. Data from the EPM 2005 collected in Madagascar provide information on the amount of hours worked by each household member along with measures for market imperfections. A simple theoretical model highlights that land when constrained on the land and the labor markets and wage when not, have respectively no effect on child hours of work. Using a switching regression model with unknown sample separation to classify households in the two regimes (constrained or not), this paper shows that not belonging to the largest ethnic group at the local level significantly decreases access to the market. The same result holds for religion highlighting the importance of the informal market.

Keywords: child labor, market imperfections, wealth paradox, sub-Saharan Africa

JEL classification: I1, 012, D13, J13, J82

*GATE Lyon-Saint-Etienne CNRS UMR 5824, 93 chemin des Mouilles, 69130 Ecully (France).Tel: +33 (0)4 72 86 61 16. Fax: +33 (0)4 72 86 60 90. Email: badji@gate.cnrs.fr

1 Introduction

There is a growing global concern for children's rights and more particularly on child labor as evidenced by the 2014 Nobel Prize attribution. Though the worst forms of child labor are the most covered by the media, most child work actually consist in family work on the farm. This type of work is mostly associated with unpaid work and very early entry ages (ILO, 2010). It is mainly found in less developed economies, and beyond the obvious ethical concerns, there exist other reasons for why one should be concerned with it. Child farm labor often has present and future consequences on the child's health, leading to less productive workers in the future, and a less productive economy as a whole. In addition, reducing child labor might allow families to escape the intergenerational poverty trap by creating a virtuous circle: households sending their children to school allow them to be more educated and therefore to obtain a higher income. In other words, these future parents are more able to forgo child labor and thus their children can get education and so on and so forth. In that sense, fighting child labor is also a way to fight inequalities. As children often inherit the land of their parents, it is important for them to acquire the necessary skills to run a farm and therefore a certain level of child labor might be wished. Nevertheless the amount of hours worked found empirically suggests that child labor, is, in general, not a choice but almost an obligation. It does not necessarily originate from poverty but rather from a lack of access to the land and the labor markets, preventing parents from adjusting their land quantity or labor demand.

Access to the land and the labor markets is a major issue in developing countries where the agricultural sector typically plays a key role. More than two thirds of the poor living in rural areas have agriculture as their main source of income. However, these two markets are usually imperfect leading to relatively high transaction costs (Sadoulet, de Janvry,

Fafchamps, 1991) that deprive the most vulnerable households from interacting with them. Helping households access these markets is critical to avoid suboptimal outcomes that generally lead to more poverty and more child labor in particular (Bhalotra and Heady, 2003[4]; Dumas, 2013[11]). The literature has indeed identified two main determinants for child labor namely poverty and market imperfections (see Basu, Das and Dutta, 2010; Grootaert and Kanbur, 1995 or Fors, 2012 for a more recent review). The assumption behind the poverty hypothesis is simple: parents would make their children work to meet a subsistence level. Empirical evidence is hard to gather as the relationship between wealth and child labor is often endogenous. In addition, even when a negative relationship between child labor and exogenous wealth is estimated, it simply indicates that child leisure is a normal good and not necessarily that households make their children work because they are poor. Bhalotra (2007) argues that if the child wage elasticity is negative then child work is compelled by poverty. Indeed, if children work to meet a threshold, then they would need to work less as their wage increases. She finds a negative elasticity though only for boys. To the extent that the number of observations is limited in the study and that wage work is a limited sample for child labor in general, it is difficult to conclude that poverty is a main factor.

In general, poverty alone may not imply that children will be sent to work as parents could borrow against the children's enhanced earning (when sent to school). In the presence of credit market imperfections however, such borrowing is threatened and only land-rich households can provide the collateral[1]. Against previous explanations, the empirical literature has shown that some countries were displaying the "wealth paradox", that is, child labor is actually more predominant in land-rich households than in land-poor households. Bhalotra and Heady (2003) found the wealth paradox in Pakistan and Ghana (2007, 2013),

[1]Of course, in the ultimate case of absence of credit market, richer households would obviously be less likely to be in the subsistence situation described by the poverty hypothesis just before.

Dumas in Burkina Faso and Madagascar , Congdon Fors (2007) in rural India, Basu et al. (2010) in Northern India and Mueller (1984) in rural Botswana. The formers claim that it may be the result of market imperfections on both the land and the labor markets. When examining the link between land and child labor, two important opposing effects are at stake. The (negative) wealth effect where wealthier households are more able to forgo child labor. The second effect (the price or substitution effect) where the more land to which one has access, the costlier it is not to solicit child labor. Indeed, the marginal productivity of labor increases with land size and more land must be farmed. That is, one hour of labor spent on a large plot leads to higher yields than one hour of labor on a smaller plot. This increasing opportunity cost, however, strikes only for households with no other alternative but to use their children's work.

Bhalotra and Heady (2003) claim that credit market imperfections induce a negative effect of land on child labor. Indeed land serves as collateral and the higher the amount of land, the lower the interest rate so that credit market imperfections induce less child labor. Another channel, not discussed by Bhalotra and Heady, is through credit use. If the money obtained from the loan is used to invest in productive assets, then credit market imperfections will also induce a positive relationship with child labor. To the extent that one can control for these newly bought productive assets, it should be of limited importance (the extra productivity of one hour of work on the land can be captured by controlling for these assets). On the labor market, if households have more to farm than adult on-farm-labor supply can manage, the household will be willing to hire external workers. If external workers cannot be found, the household might rely on his/her own labor supply to face this extra work. The adult workers will work more and the household might rely on its child labor supply in order to face the labor shortage. The need for workforce will be even higher as land increases because the marginal productivity of labor is increasing

with land quantity. The opportunity cost of not farming the land increases with land size and that is why, land and child labor will be positively correlated. On the land market, if we are in a situation where the household has more land to farm than it can take care of, it can still sell the extra plots. However if the land market is not properly functioning, the household will be unable to sell its plots. The opportunity cost of not farming the plot that cannot be sold also increases as the amount of land increases and for the same reasons, this might induce a positive relationship between land and child labor. Obviously, if the land market solely fails, any amount of land that cannot be sold or rented out will be farmed by external workers. Conversely, when the labor market alone fails, preventing landlords from hiring external workers, the plots can simply be sold or rented out. In both cases, farm size will play little or no role on the quantity of child labor used. However, when both markets fail, households having lots to farm are much more likely to rely on child labor, inducing a positive relationship between land size and amount of child farm work, once the wealth effect is taken into account.

The reasons as to why the land and the labor markets could fail are manifold. Market failures in the agricultural sector are very widespread, especially in developing countries, as markets are typically very "local". Individuals in the same area might farm similar types of plants. These individuals and their neighbors will need the same type of work at the same period. For example they will all need unskilled labor for harvesting their plots at the very same time. This is called the simultaneity issue. Asymmetry of information also leads to market failures. When there is asymmetry of information, the labor hired for one's farm is not a perfect substitute for the labor provided by the household. In theory, hired-in labor and family labor should both be remunerated at their marginal productivity. However farm labor's output does not solely depend on effort and as effort cannot be observed (moral hazard), people will prefer to use their own labor since it does not encompass supervision

costs. The lack of property rights (Ray, 1998) might hamper transactions on the land market. Buyers not certain that the plot they want to purchase will truly be theirs are deterred from buying plots. The few buyers that might take the risk of buying land are willing to pay less than the market value of the land (which can be interpreted as the present value of the future profits generated). The uncertainty regarding land ownership can be worsened by new land reforms themselves leading former and new property systems to coexist and sometimes conflict. Besides land has collateral value that is only available to the owner. That is, in addition to its value *per se*, land has an extra value as collateral for loans and the owner will benefit from better interest rates when it owes the land and uses it for making a loan. Along the collateral value, it could be argued that land also has some value to the seller who might keep it for his children. However, a buyer might also think of land as being an investment for his children so that the only remaining difference could simply be in the seller's attachment to its land (in the sense that it potentially has belonged to the family for generations and should stay so). The mismatch between the value for the buyer and the value for the seller hampers the sales.

To circumvent these issues, the household might decide to rent the plot. However, the aforementioned failures on land purchase also apply for land rental. Asymmetries of information regarding land usage (will the land still be fertile after two years of potential intensive agriculture?) make most households unwilling to rent their plots. Sharecropping might solve some of the problem of hiring labor or selling/renting the plots. Sharecropping consists in having the owner of the plot and the person who cultivates it deciding together on a share of the production against the "free" use of the land. Yet, if the owner cannot observe the amount produced, and therefore that the right amount of production is received, the incentives for sharecropping will clearly be reduced or could even disappear.

Clearly, households that have just enough adult labor force to farm their land will not

need to have their children work on the farm. More generally, the positive effect of land on child labor will not only occur if land and labor markets fail, but if households are constrained by the two failing markets in the sense that they would actually need to trade land or labor in order to have their labor/land demand match the supply. In other words, they need the market if their own labor/land demand does not match their own supply. We cannot assess beforehand which households are constrained. We can only evaluate the probability that they are constrained by looking at measures for market imperfections.

To identify adapted policies to improve land and labor market access, one must first identify the households that lack access to these two markets. Yet it does not mean that all households are unable to access them. The wealth paradox is an average effect on all households. It does not mean that each and every households are excluded. As a matter of fact, imperfections are not market-specific but household-specific (Sadoulet, de Janvry, Fafchamps, 1991). Strictly speaking, a household lacks access when the cost of a transaction through the market creates greater disutility than the gain it generates, with the result that the market is not used for the transaction. To recapitulate, a household is constrained if a) the land and the labor markets cannot be fully accessed for that particular household, b) the household would like/need to interact on the market(s). We rely on the preceding analysis to identify the households rationed (or "constrained") by the market imperfections. Controlling for the wealth effect, there is no effect of land on child labor supply unless both markets fail (and the household would need market access).

This paper relies on the relationship between land and child labor to identify households constrained on the land and labor markets. The model allows not only to identify households excluded from the formal markets but also from the informal markets. The relationship between child labor supply and land holds as long as one has no other alternatives and these alternatives can be on the formal or informal sectors. In order to avoid

the transaction costs associated with the formal access to the land and the labor markets, a number of individuals move part of their economic activities into the informal sector. where the role of culture, religion, and ethnicity might play a strong role (Godfrey, 2011). One might indeed wonder whether kin relationships help in case of labor shortage. As households from a certain religion/ethnicity might have different preferences, we need to separate the effect linked with household preferences. Distinguishing the two is beyond the scope of this paper. We will simply assess whether being in the largest ethnic group (at the local level) affects the probability to belong to the "constrained" households and similarly for religion.

The purpose of this paper is to identify the households that cannot access the land and the labor markets when they need to and who, as a result, have their children do more farm work. The paper further looks at whether minorities at the village level are more likely to be constrained using a switching regression model. A dataset on the number of hours worked by each family member above five and information on wealth and productive assets helps show that mother education and wealth decreases the number of hours worked by children. The ratio of girls decreases the probability to be constrained whereas it has a negative effect on the average child supply at the household level suggesting that girls in the household might serve more as an adjustment variable through domestic work. We find evidence of a restricted access for households that do not belong to the dominant (in number) ethnic group of their local area or the dominant religion.

The remainder of this paper is organized as follows. Section two describes the data. Section three presents the theoretical model. Section four explains the econometric implementation. Section five shows the empirical results and the last section concludes.

2 Data and Descriptive Statistics

The data are drawn from a household survey entitled "Enquête auprès des Ménages 2005" (EPM) collected by the National Institute of Statistics of Madagascar (INSTAT). It is a sample representative of the Malagasy population and includes 54995 individuals belonging to 11781 households along 561 cities. The sample was divided into 44 strata, comprised of the rural and urban area of the 22 regions. For each locality, 21 households were surveyed. The dataset provides information not only on the amount of hours worked by any above-six-year-old family member but also on market imperfections. Information on titling, registration on a land cadastre, renting and sharecropping is given for each plot. For each household, financial institution availability in the village and type (formal: bank, microfinance institution or informal: friends, moneylenders,...) is inquired.

The population in Madagascar is very young and rural. Half of Malagasy are under 20 years old and 78% live in rural areas. They have a low level of education with as high as a quarter of the population who did not finish primary school. 90% of Malagasy are believers and religions are mainly grouped into the "traditional" religion, Catholicism and Protestantism. Other religions reflect a much smaller fraction of the population. Madagascar has originally 18 ethnic groups, each having their own dialect[2]. The distribution for religions and ethnic groups is displayed in tables 15, 16, 17 and 18. In the following subsections, a description of the credit market and more particularly the land and the labor markets is undertaken.

[2]Note that the dialects are not very different so that Malagasy from different ethnic groups still understand each other. Another six ethnic groups are formed by more recent migrations. Ethnic groups and religions had to be aggregated and the distribution of religions and ethnicities can be found in the appendix.

2.1 Credit Market

The credit market in Madagascar, at least for the formal sector, is largely underdeveloped. A few 3.6% of households are part of a mutual savings bank, less than 3% have a bank account, a mere 8.5% intend to open a bank account and a few 10.5% intend to become members of a mutual savings bank.

Focusing on households who borrowed money in the 12 months preceding the interview (4.6% of total households), we find that half of them inquired neighboring households, a quarter relatives and 6.6% a moneylender. Less than a fifth had recourse to the formal sector (13.7% through a microfinance institution and 5.5% through a bank). It seems that the credit market suffers from a lack of information. As high as 95% of households do not know how to get a loan from a formal institution.

To see how imperfections on the credit market might have an impact on the household, we investigate the reasons for inquiring a loan and find that among the 619 loans made by the 532 households in the year before the interview, 40% were contracted in order to buy food and a quarter for professional activity including 13% for agriculture. This points not only to the high vulnerability of households in Madagascar but also that only a few loans are agriculture-related.

2.2 Labor Market

In Madagascar, the labor market is quite limited despite a particularly high participation rate as described in table 1[3]. Participation rates are as high as 82% for women and 90% for men. Most of this work consists in farm work for adults, and almost only farm work for the 6 to 15 years old. More than 97% of the children in the agricultural sector work in their families which goes against the assumption of having a child farm wage. Not surprisingly,

[3]Participation is simply understood as working. It is not limited to formal salary employment.

98% of children never tried to look for employment suggesting that they work when their parents want and/or need them to. When engaged in any activity, most individuals are in the primary sector (78.5%), 96% of which is agriculture. Slight imperfections on this sector can clearly have tremendous consequences. Despite the huge participation rates, the share of adult population engaged in salaried work plummets to barely 16%. For the primary sector, it sinks to less than 5% of salaried work. A third of the Malagasy rely on a second activity where they are mostly independent workers (35%) or salaried (22% of those having a secondary activity) workers which in that latter case involve a temporary activity for 83% of the individuals. More generally, 84% of the main activity is a permanent activity but because most of the work is from self employment (38%) or family work (46%), it is hard to interpret it as an indicator of reliable or safe employment due to the intrinsic volatility in agricultural work.

Note that all the participation statistics were based on the seven days prior to the survey and therefore one might be worried that they do not reflect the reality especially as one might think of agricultural as typically seasonal. Stifel et al. (2007) compared surveys for Madagascar with similar questions with the 2004 survey (where the question did not only rely on the last seven days) and found that participation rates were very similar. He found that seasonal employment accounted for less than 1% of total employment and we can therefore rely on the last seven days for our estimations[4].

Table 2 gives the number of hours worked in a typical week[5] for participating adults and children disaggregated by type of work, age and gender. Adults work on average 45 hours a week or a bit less than 40 when in the agricultural sector. They spend almost the same amount of time in domestic chores. When focusing on children, we find that one

[4]The previous statistics imply that there is no particular seasonality in the *participation* rates which is different from saying that there is no seasonality in the *number* of hours worked.

[5]A typical week is what we are interested in and it most likely differ from the number of hours worked in the last seven days.

Table 1: Participation rates in percentages by gender, age category and type of work

	Work		Farm work		Domestic chores	
	Male	Female	Male	Female	Male	Female
6-10	14.5	13.5	13.5	12.9	80.2	82.2
11-15	30.3	29.4	27.0	26.9	89.3	93.1
16-60	89.1	81.9	68.8	65.0	89.3	98.2

Work includes farm work. Individuals may participate to several activities (domestic chores and work) so that adding the different activities does not sum to 100.

child out of five is working. 93% of young children (aged 6-10) and 88% of older children (aged 11-15) are in the agricultural sector indicating that child labor is clearly dominated by child farm labor.

Table 2: Hours of work disaggregated by gender, age category and type of work for individuals who worked at least one hour on the farm in the seven days preceding the interview

	Work		Farm work		Domestic chores	
	Male	Female	Male	Female	Male	Female
6-10	24.6	21.1	22.2	19.9	7.1	7.8
11-15	31.9	30.1	28.6	27.3	8.9	12.6
16-60	48.1	40.5	40.5	34.8	9.6	20.2

In the survey, domestic work is given in minutes for a typical day; (farm) work for a typical week. Minutes of domestic work were converted to hours and multiplied by seven for the sake of comparability. Work includes farm work.

2.3 Land Market

To understand how the land market functions, let us look at table 3 which portrays the different uses of land in Madagascar. It shows that as high as 80% of plots are owned and used/farmed by the same household. If a household has land and if most workers are paid at piecework, we should probably not observe such a high amount of self-use. Households may simply have the exact amount of land they need thanks to a very well functioning land market yet the tables afterwards cast some doubt on this explanation. Sharecropping

and renting together account for almost 10% of land use, that is almost one household out of 10 uses the previously described strategies of renting and sharecropping in order to circumvent the market failures which is quite high.

Table 3: Use of agricultural land

Land use	Freq.	Percent
Own farm	20,323	80.26
Sharecropping	1,120	4.42
Renting	1,030	4.07
Free use	1,954	7.72
Temporary occupation	596	2.35
Not farmed	298	1.18
Total	25,321	100

Table 4: Owned plots' status

Status	Freq.	Percent
Purchased	2,804	13.42
Inherited	12,852	61.50
Donated	1,487	7.12
Clearing	3,439	16.46
Other	316	1.51
Total	20,898	100

Table 4 shows another facet of market imperfections. It indicates how owned plots were acquired and depicts quite low mobility across land ownership with more than 60% of plots acquired through inheritance[6]. Less than 15 % of plots were bought and there were actually more plots acquired through clearing than buying. If we focus on how each household (as opposed to each plot, so that the figures are not driven by a few who would have bought lots of plots) acquired their plots, we find similar results with 55% of households who

[6]Only riceland and *tanety* land are presented but the distribution of farmable land (riceland and *tanety* land) is very close to that of non farmable plots.

acquired all of their plots through inheritance[7], more than four households out of five did not buy any of their plots and a few 10% bought all of their plots. We now turn to table 5 which indicates from whom plots were bought for each plot. Almost one out of five was bought from a close relative. About 30% from a friend or a distant relation which amount to almost 50% of plots bought through relatives which explains why there is such a low mobility for land.

Table 5: Purchased plots' origin

Origin	Freq.	Percent.	Cumul.
Close relative	489	17.56	17.56
Friend or distant relation	816	29.30	46.86
Villagers	1,353	48.58	95.44
Other	127	4.56	100
Total	2,785	100	

It is possible that most of the time inheritance schemes are adjusted so that the households are left with the amount of land that they need. However we do believe that market imperfections occur preventing households from adjusting their quantity of land as can be understood from the statistics on land titling given in table 6. The status of each plot of land owned is indicated for farmable land (a very close distribution is found for non farmable land). We find that most of the land is ancestral land which can be associated to the previous finding that most owned plots were inherited. As high as 12% of lands that were owned are actually state land. Less than a fourth of plots are bound by a contract, 16% being registered on a cadaster and not even 5 % are titled. Among titled plots, 35% are not titled under the name of any household member and a fifth are under the name of a deceased person. Focusing on titled land, we find that it takes on average five years and

[7] 35% acquired all of their plots without inheritance, that is a mere 10% of households have a mixture of inherited and non inherited farmable plots.

a half to have a plot titled, although for 50% of titled plots it took less than two months suggesting strong ownership issues on the land market.

Table 6: Plot status for farmable land

Legal status	Freq.	Percent
Titled	1,002	4.73
Cadastre	2,549	12.04
Legal contract	1,376	6.50
Written contract	499	2.36
Verbal contract	265	1.25
Ancestral land	12,137	57.35
State land	2,625	12.40
Not known	435	2.06
Others	276	1.30
Total	21,164	100

All in all, a strong case can be made that the land market is extremely rigid and that we are likely to observe households constrained on both the land and the labor markets. Their behavior is made plain in the following section.

3 Theoretical model

To model the differences in the response function of the amount of child labor to the amount of land, we use the classical farm household model (Singh, Squire and Strauss offer a first review in 1986) where the household is a consumer, a worker, and a producer. We consider here a unitary framework with one adult and one child[8] and solve the problem under two alternative assumptions: 1) markets work well enough so that the household uses them when needed and 2) the opportunity costs are so high that the household does not want to trade on the market. Dumas' model (2013) allows for four regimes depending on where

[8]The econometric implementation allows for more flexibility in the estimation with household composition being controlled for.

the marginal productivity of labor of the household's farm is located with respect to two wages; the first wage being the wage at which the adult members of the household can be hired and the second wage is the wage at which the household can hire labor. The setting presented here cannot be straightforwardly associated with the regimes in Dumas' paper as the price band is not made explicit: we are assuming that there is a continuum of wages corresponding to the productivity of each and every individual. Furthermore, households can be constrained in the quantity wished to trade more than in the actual access. As described before we need the household to be unable to trade on both markets to see any effect. We chose here to assume no land market so that when the household cannot access the labor market, it will be unable to access both the land and the labor markets[9].

3.1 General framework with market access

We assume a unitary framework where the household maximizes its utility under its budget and technology constraint. Its utility is a function of household consumption x, adult leisure l^a and child leisure l^c. We assume the latter are additively separable[10] so that the utility function to maximize is the following:

$$u(x) + u_a(l^a) + u_c(l^c) \tag{1}$$

where the subscripts a stands for adults and c for children. u, u_a and u_c are all concave. Adults can choose between spending their time working on the family farm denoted as h^a, working off the family farm t^a and not working at all l^a. Children do not have the opportunity to work outside the farm and can either work on the family farm an amount h^c or not work l^c. Children and adults might have different time endowments so that their

[9]The descriptive section emphasizes how little role the land market is playing so that assuming no land market is not a strong assumption.

[10]Not assuming separability leads to the same results at a price of a very tedious model.

time constraints write as[11]:

$$\bar{E_c} = h^c + l^c \tag{2}$$

$$\bar{E_a} = h^a + l^a + t^a \tag{3}$$

Let $v_a(s) \equiv -u_a(-s + \bar{E_a})$ and $v_c(s) \equiv -u_c(-s + \bar{E_c})$, then the utility function rewrites as:

$$u(x) - v_a(h^a + t^a) - v_c(h^c) \tag{4}$$

The standard assumptions that the marginal disutility of work increases with the amount of work and the marginal utility of consumption decreases with consumption translate into: $v_c' > 0$, $v_a' > 0$, $u' > 0$, $v_a'' > 0$, $v_c'' > 0$, $u'' < 0$. We assume a general production function with two inputs: land k and labor h . When the labor market is not rationed for a household, then that household can freely hire external labor denoted as h^i. The household can also choose to have the children work. We assume that all types of labor are perfect substitutes. That is, we assume that there is no labor that an adult cannot undertake without children and there is no type of on farm work that a child cannot undertake without parents. The same holds for external workers. Although it is not difficult to argue that any type of agricultural work a child undertakes can be undertaken by an adult and that adult will not need children to work, it might seem a bit difficult to assume the opposite. For the sake of simplicity, we will however maintain this assumption. In the end, adult farm work, child farm work and external work only differ by their efficiency δ. This allows not only to distinguish child labor and adult labor by their productivity but it also allows to model differences between hired in labor and off-farm labor with labor being remunerated

[11] Assuming $E_c = E_a$ does not change the results.

at its marginal productivity for an interior solution. This assumption is not strong in the sense that labor is often paid on piecework and therefore, regardless of the amount of time spent on the farm, the wage received will be the same for harvesting a plot. In the end, it comes down to paying a worker per hour relative to their productivity. The technology of production follows the standard assumptions of a decreasing marginal productivity of labor with labor and an increasing marginal productivity of labor with land:

$$q = g(\delta_a h^a + \delta_c h^c + \delta_i h^i; k) \text{ with } g_k > 0, g_h > 0, g_{hh} < 0, g_{hk} > 0 \tag{5}$$

Fixing the price of output at p, denoting y exogenous transfers, that is revenue outside work, and w_a the off family farm wage, we obtain the following budget constraint:

$$y + w_a \bar{E}_a + pq = x + w_i h^i + w_a h^a + w_a l^a \tag{6}$$

x can be seen as expenditures on goods or the amount of goods purchased if we normalize the respective price to one. Equation (6) represents the budget constraint. Total wealth, that is, exogenous transfers, the value of adult time and the revenue generated by farm output sales must equal total expenditures, that is, expenditures on goods, hired-in labor, and adult time not spent working outside the farm (as an opportunity cost).

Combining equation (2) to (6), the maximization program to solve is the following:

$$\max_{h^i, h^a, h^c, t^a} u(y + w_a t^a + pg(\delta_a h^a + \delta_i h^i + \delta_c h^c, k) - w_i h^i) - v_a(h^a + t^a) - v_c(h^c)$$

3.1.1 Interior solutions

For interior solutions, the first order conditions follow:

$$\frac{\partial \mathcal{P}}{\partial t^a} = 0 \Leftrightarrow u' w_a = v'_a \tag{7}$$

$$\frac{\partial \mathcal{P}}{\partial h^a} = 0 \Leftrightarrow u' p g_h \delta_a = v'_a \tag{8}$$

$$\frac{\partial \mathcal{P}}{\partial h^c} = 0 \Leftrightarrow u' p g_h \delta_c = v'_c \tag{9}$$

$$\frac{\partial \mathcal{P}}{\partial h^i} = 0 \Leftrightarrow u' p g_h \delta_i = u' w_i \Leftrightarrow p g_h \delta_i = w_i \tag{10}$$

Equation (9) can be rewritten as: $u' w_c = v'_c \leftarrow u'/1 = v'_c / w_c$. It indicates that at the optimum one euro spent on leisure brings the same utility as one euro spent on consumption with w_c representing the opportunity cost of not working. Unfortunately we do not observe the amount of external work hired in the farm, nor do we observe profit. We can however derive the m-demand (Browning, 1999) of child labor from the first order condition. M-demands simply consist in modeling the demands derived from the marginal rates of substitution, as a function of prices and the quantity of a reference good. The reference good must be normal and here child labor demand is written conditioning on x which can be seen as representing the intertemporal wealth of the household. Using the first order condition for child labor $u'(x)\delta_c w = v'_c(h^c)$ with $w = pg_h$, child labor supply can be written as a function of w and x (child productivity δ_c is already included in the function f below):

$$h^c = f(w, x)$$

The partial derivatives are given hereafter:

$$\frac{\partial h^c}{\partial x} \equiv h^c_x = \frac{u''}{v''}\delta_c w < 0, \qquad \frac{\partial h^c}{\partial w} \equiv h^c_w = \frac{u'}{v''}\delta_c > 0. \tag{11}$$

An increase in the total amount of expenditure, that is an increase in wealth, will lead to less child labor. If the local wage increases however, child labor will tend to rise. If the local wage becomes higher, hired-in labor becomes more expensive and therefore the household has a stronger incentive to have children work instead of hiring labor.

Obviously, if h^c is a function of solely w and x, then, controlling for w and x, land will no longer impact child labor. In other words, controlling for the local wage in the primary sector and the total expenditures of the household, land will not impact child labor so that when regressing child amount of work on the local wage, total expenditures and land, the coefficient for land should be zero.

3.1.2 Corner solutions

The results previously given hold for an interior solution. For corner solutions, the appendix shows in greater details how child hours of work will respond to land quantity. The following summarizes the results.

The main reason that leads to an absence of effect of land on child labor, is that an exogenous wage (as opposed to a shadow wage) can be defined. In the general case for interior solutions, one of the adult[12] first order conditions is clearly redundant. Thus, any corner solution where only one adult variable is zero will give the same m-demand for child labor as in the general case for interior solutions and land will not impact child labor.

Cases for which $h^c = 0$ are not investigated because if the quantity of child labor is

[12] What is understood by adult is either adult off farm labor t^a, adult on farm labor h^a or (adult) hired-in labor h^i.

fixed, then it will not vary with respect to land (as it will not vary at all). Ultimately, the cases of interest are $t^a = h^i = 0$, $h^a = t^a = 0$, $h^i = h^a = 0$ and $t^a = h^i = h^a = 0$.

If $t^a = h^i = 0$ (and as well if $t^a = h^i = h^a = 0$), then there is no adult working outside the farm and no labor is hired. Note that $t^a = 0 \Leftrightarrow pg_h\delta_a > w_a$ and $h^i = 0 \Leftrightarrow pg_h\delta_i < w_i$. This is not possible in the framework with market access as it implies both $pg_h > w$ and $pg_h < w$ so that households who are neither working outside the farm, nor hiring labor will be assigned to the regime with limited market access (the next section).

If $t^a = h^a = 0$, since there is still access to external labor, the m-demand for child labor will still be written as in the general case, that is $h^c = f(w, x)$ and land will not impact the level of child labor even though adult members of the households are not working.

Finally $h^i = h^a = 0$ corresponds to a case where only children are working on the farm. Since $h^i = 0$, it means that $pg_h < w$, so that if there is market access, we necessarily have $h^a = 0$ (as presented). In this case external workers cannot be hired because the marginal productivity of an hour of their work is lower than the wage to pay them. In the same way, adults have no incentive to work their own farm as they can earn more by working off the farm. In this particular situation, the m-demand for child labor will be such that $h^c = f(pg_h, w)$ and hours of child work will increase with land quantity, despite market access. In this situation, the productivity of the land is so low that it is not worth farming for anyone but the children since they do not have the option of working outside the farm. Though theoretically possible, empirically, more than 98% of households who have land and whose adult members are not working their farm would have their children farm it (that is $h^c > 0$ and $h^a = 0$). As mentioned information on hired-in labor is lacking, so that 98% should be seen as a lower bound for the case where in addition households would not hire in labor (that is $h^c > 0$, $h^a = 0$ and $h^i = 0$). For that reason, we believe that this

later case can be ignored and therefore we can conclude that even for corner solutions land will not impact child labor if the household can fully access the labor market.

3.2 Model with labor market imperfections for adults

We assume that the household cannot rely on the market to satisfy its needs regarding work so that we have : $h^i = \bar{h^i}$ and $t^a = \bar{t^a}$ (note that $h^i = t^a = 0$ is a particular case of the model). The function to maximize becomes:

$$\max_{h^a,h^c} u(y + w_a\bar{t^a} + pg(\delta_a h^a + \delta_i\bar{h^i} + \delta_c h^c, k) - w_i\bar{h^i}) - v_a(h^a + \bar{t^a}) - v_c(h^c)$$

We obtain the following FOCs:

$$\frac{\partial \mathcal{P}}{\partial h^a} = 0 \Leftrightarrow u' pg_h\delta_a = v'_a, \tag{12}$$

$$\frac{\partial \mathcal{P}}{\partial h^c} = 0 \Leftrightarrow u' pg_h\delta_c = v'_c \tag{13}$$

Here we do not have anymore that $pg_h = w$ or a constant, so we define a non constant shadow wage as:

$$\tilde{w} = pg_h$$

Clearly, the supply of child farm work will be a function of this shadow wage and total consumption (through u) and we can rewrite the supply function as:

$$\boxed{h^c = f(\widetilde{w}, x) = f(pg_h, x)}$$

with f increasing in its first argument and decreasing in its second argument. The difference between this case and the previous case with market access, is that here controlling for x and $\tilde{w}$ is not feasible as the shadow wage cannot be observed. Of course if we

could keep x and $\tilde{w}$ constant, we would have a zero effect of land. Yet if land increases, g_h increases as well that is even holding total expenditure x constant, land will impact the amount of child labor through the shadow wage.

$$\frac{\partial h^c}{\partial k} \equiv h_k^c = f_{\widetilde{w}} \frac{\partial \widetilde{w}}{\partial k} = f_{\widetilde{w}} p \delta_c (g_{hh} \delta_c h_k^c + g_{hk}) \Leftrightarrow h_k^c = \frac{f_{\widetilde{w}} \delta_c g_{hk} p}{1 - f_{\widetilde{w}} p g_{hh} \delta_c^2} > 0 \tag{14}$$

$h_k^c > 0$[13] indicates that when the household is constrained by the land and the labor markets[14], land has a positive effect on child labor after controlling for expenditure x.

The previous results hold for interior solutions and corner solutions except when child labor is at zero. If we assume that child labor is fixed at zero, then despite a restricted access to the market, land quantity will not affect child labor (since it is fixed at zero). The only corner solution left is then such that $h^a = 0$ and $h^c > 0$. It is obvious that in this case the m-demand will be such that $h^c = f(pg_h, x)$ and therefore land quantity will affect child labor in the same way as for an interior solution.

On a more general note, we expect productive assets to have the same effect as land on child labor. That is, as they increase, child labor should increase only if the household has no access to the land and the labor markets.

3.3 Parametrization

The previous model, though simple, allows for a finer interpretation of the estimated reduced-form. Clearly, the estimation enables more flexibility by introducing more controls. The following shows how the parameters in the theoretical model and the estimated

[13] The denominator is necessarily positive since all terms are positive except for g_{hh} which is negative so that the denominator is actually greater than one.

[14] Recall that we assumed no land market for simplification so that when the labor market is not functioning for the household then both the land and the labor markets are not functioning for the household.

model relate. Assuming an exponential form (CRRA) for u and v_c and assuming that all the parameters are positive (so that the previous assumptions on the utility functions hold), one obtains:

$$\begin{aligned} u &= -ae^{-\alpha x} \\ v_c &= be^{\beta h^c} \end{aligned}$$

with α and β indicating the curvature for the functions u and v_c respectively.

For the first regime with market access, using equation (10) from the FOC, equation (9) translates into $u'w_c = v'$ and becomes: $\alpha a e^{-\alpha x} w_c = \beta b e^{\beta h^c}$. In the regression we will estimate an additive form so that we take the logarithm of the previous equation and obtain[15] for h_1^c (which stands for h^c in regime 1, i.e. with market access) :

$$h_1^c = \frac{1}{\beta}\ln\left(\frac{\alpha a}{\beta b}\right) - \frac{\alpha}{\beta}x + \frac{1}{\beta}\ln w_c \tag{15}$$

$$\text{or equivalently: } h_1^c = \underbrace{\frac{1}{\beta}\ln(\frac{\alpha a \delta_c}{\beta b})}_{A_1} - \underbrace{\frac{\alpha}{\beta}}_{B_1} x + \underbrace{\frac{1}{\beta}}_{C_1} \ln w \tag{16}$$

A_1, B_1 and C_1 are the coefficients from the regression. Clearly β is identified through C_1 and then α through C_1 and B_1. Because the utility is ordinal we can set $a = 1$ without loss of generality and one obtains $A_1 = \frac{\alpha\delta_c}{\beta b}$ so that only the ratio δ_c/b can be identified and unsurprisingly, child productivity cannot be identified.

If we now look at the case where the household is constrained (regime 2) we have:

[15] see appendix for details

$$\underbrace{p\delta_c g_h}_{\tilde{w}_c} u' = v'$$

Assuming that the production function is of the following form:

$$g(h,k) = \frac{ck^{\gamma_k}}{\gamma_h}(1 - e^{-\gamma_h h}) \text{ with } g_h > 0, g_k > 0 \text{ and } g_{hk} > 0$$

then $\tilde{w}_c$ will write as:

$$\tilde{w}^c = p\delta_c c k^{\gamma_k} e^{-\gamma_h h^c}$$

with all parameters positive so that $-\gamma_h h^c$ and $\gamma_k k$ reflect the fact that the marginal productivity of labor decreases with labor but increases with land. We obtain (denoting now h^c as h_2^c for the regime 2, that is the regime where households are constrained):

$$\Leftrightarrow h_2^c = \underbrace{\frac{1}{\gamma_h + \beta} \ln \frac{a\alpha p \delta_c c}{b\beta}}_{A_2} - \underbrace{\frac{\alpha}{\gamma_h + \beta}}_{B_2} x + \underbrace{\frac{\gamma_k}{\gamma_h + \beta}}_{C_2} \ln k \tag{17}$$

Since α and β are identified thanks to the previous regime, we get γ_h from B_2 and then γ_k from C_2. If we set again $a = 1$ and $p = 1$, we can identify δ_c/b from the previous regime and therefore we can identify c. Again, we cannot identify separately δ_c from b.

The following section will give the results for the estimation of h_2^c and h_1^c. Clearly these equations have an additive form and corresponds to the functional form estimated in the next section.

4 Econometric Implementation

4.1 General framework

The strategy of the paper consists in identifying the households that are constrained in their access to the land and the labor markets using their response function regarding child labor supply. A switching regression model with unknown regime separation will help us in the assessment of the two regimes. The system is the following:

$$\begin{cases} C^* = \mathbf{Z}\Gamma_0 + \alpha_0 \text{ ln k} + \beta_0 \text{ ln w} + \mathbf{G}\Gamma_c + \epsilon & \text{Switching equation} \\ h_1^{c*} = \mathbf{Z}\Gamma_1 + \alpha_1 \text{ ln w} + u_1 \text{ if } C^* < \overline{C} & \text{Regime 1, not constrained} \\ h_2^{c*} = \mathbf{Z}\Gamma_2 + \beta_2 \text{ ln k} + u_2 \text{ if } C^* > \overline{C} & \text{Regime 2, constrained} \end{cases}$$

The first equation determines regime membership (the "switching" equation) where characteristics for market imperfections as well as household head characteristics are included. Some of these variables are common to all equations and are represented by vector **Z**. **Z** includes ethnicity, religion[16], parent's education, sex of the household head, family composition, total expenditure and a dummy indicating whether the household lives in a rural or an urban area. ln w is the logarithm of the wage (only included in the switching equation and the first regime) and ln k is the logarithm of owned land (only included in the switching equation and the second regime according to the theoretical model). **G** are all the variables that are likely to influence regime membership but not the quantity of hours worked (so that they are only in the switching equation). **G** includes a dummy indicating whether the household is part of the local ethnic minority at the community level, a similar dummy for religious minority, the percentage of plots titled at the community level, the percentage of plots registered on a cadaster at the community level, the rate of salaried

[16] If in very rare cases, different religions or ethnic groups were found in a same household, then the religion or ethnicity of the household head was chosen

workers in the primary sector and the rate of salaried workers in the non primary sector. The first regime is composed of households who either do not need to trade on the land and the labor markets or who *need and can* trade on them (the "unconstrained" regime). The second regime is made up of households rationed on both markets and who would need to trade on these two markets (the "constrained" regime).

It is important to bear in mind that this switching strategy is used because identifying which households are constrained is not possible ex ante. It is sometimes possible to observe whether households trade on a market or look at market imperfection variables, however it is not possible to know whether those households needed to trade on the market or whether their own (within household) labor supply or amount of land was sufficient. In addition to that, some households might make transactions on the market and still be constrained with respect to the quantity they wished to trade (especially for the labor market). One must simultaneously estimate the three aforementioned equations (the two regimes and the switch).

4.2 The regimes

As mentioned, the exclusion variables for the system will be land and wage with land appearing only in the regime where households are constrained in their market access and wage appearing only in the regime where they are not. The wage will be defined as the wage prevalent where the household lives or local wage. In the survey, the module on employment is rich and allows for the record of a secondary activity for all individuals. The wage of an individual is simply the average of the earnings in the principal and secondary activities weighted by their respective number of hours. However, focusing on wage in the primary sector we find that among adults working in the primary sector, only 13.6% are salaried. For the rest of the population, the market wage can be recovered by using

the prediction from a regression of the wage on individual and community characteristics from the population for which the wage is available. Since the probability to observe the wage is likely not random, a selection issue might appear and the Heckman procedure is implemented.

To implement the procedure, a variable that influences the probability to observe the wage (that is, a variable that influences the decision to work outside one's farm) but not its value is required. Land is one such variable: as land quantity increases, the marginal productivity of labor increases as well. Land-poor households will be more likely to work outside the family farm as it offers a higher wage than the shadow wage they would obtain on their farm. The wage offered to workers is clearly independent of the quantity of land such workers have. Yet, to the extent that the wage might be partly determined at the local level, one might argue that the local primary sector wage is influenced by the quantity of land available at the local level and that the land quantity of the household influences that same local wage. Still, if each household's quantity of farmable land is small enough (atomistic agent assumption), their land amount should not influence the local wage market[17]. The basic correlation between farmable land quantity and local wage in the primary sector is of 0.0345 and is not significant at the 10% level. The regression of the primary wage on land showed that land was not a significant predictor for local wage[18].

Table 7 displays the estimation for the wage with the Heckman procedure. Regional dummies were introduced to account for the differences in market development. As expected the coefficient for land is positive, that is, the more land the household members have access to, the lower the probability to observe their wage since the marginal produc-

[17] Note that local wage is defined at the household level: it is the average income received by all workers but the workers from the household.

[18] Results are shown in the appendix, table 19 where the logarithm of the wage is regressed on land. Results were similar if the wage as opposed to its logarithm was taken as the dependent variable.

tivity of labor on the farm is very high. The prediction shows a lower wage for women and households in rural areas. We are less likely to observe the wage for educated people probably because they chose to work in other sectors. When they chose to work in the primary sector, they receive higher wages. The reference group are the 16 to 20 year old and the other age categories have a higher likelihood to be engaged in salaried work in the primary sector as well as receiving a higher wage, except for 55 to 60 year old people who are less likely to be salaried (and probably to work at all) and who do not significantly earn more than the 16 to 20 year old.

The theoretical model showed that for an individual i the wage w_i was such that $w_i = \delta_i p g_h$. Hence the regional dummies and the area (rural or urban) can be interpreted as proxies for the price p of the output. g_h is independent of the workers characteristics so that age, gender, education, religion and ethnicity are related to δ_i, the efficiency of worker i. According to the estimation, women and uneducated individuals are less efficient (or at least less paid). It seems that any household whose religion is not the traditional one is more efficient (and also more likely to engage in salaried wage). The reference ethnic group are the Merina, the largest ethnic group in Madagascar, mainly spread around the capital where most of the riceland (more productive) is. Ethnicities are not very widespread on the map so that controlling for the region hides some effect of the ethnic group. That said, we can see that the non Merina are less likely to be salaried in the agricultural sector yet when they are, they seem to earn more. Again, because of the little variability in the geographical distribution of ethnic groups, these results should be interpreted with caution.

Table 7: Wage estimation in the primary sector, Heckman procedure

	Selection	ln(wage)
land/1000	-0.401***	

	(0.056)	
Education	-0.213***	0.085***
	(0.010)	(0.025)
Female	-0.258***	-0.137***
	(0.024)	(0.033)
Rural	0.342***	-0.087**
	(0.025)	(0.041)
Age		
20-25	0.182***	0.087**
	(0.043)	(0.043)
25-30	0.278***	0.077
	(0.043)	(0.048)
30-35	0.287***	0.133***
	(0.045)	(0.049)
35-40	0.279***	0.122**
	(0.047)	(0.051)
40-45	0.087*	0.198***
	(0.051)	(0.047)
45-50	0.071	0.175***
	(0.053)	(0.049)
50-55	-0.005	0.157***
	(0.056)	(0.052)
55-60	-0.178**	0.114
	(0.072)	(0.074)
Religion		

Catholics	0.113**	0.146***
	(0.048)	(0.048)
Protestants	0.159***	0.099**
	(0.049)	(0.050)
Other	0.093*	0.008
	(0.053)	(0.055)
Ethnic group		
ethnic group 1	-0.144**	0.019
	(0.070)	(0.075)
ethnic group 2	-0.041	0.177***
	(0.067)	(0.066)
ethnic group 3	0.026	0.020
	(0.065)	(0.061)
ethnic group 5	0.019	0.096**
	(0.054)	(0.044)
ethnic group 6	-0.346***	-0.118
	(0.079)	(0.098)
Region		
Diana	-0.344***	0.260
	(0.131)	(0.170)
Sava	-0.278**	0.668***
	(0.115)	(0.146)
Itasy	1.068***	-0.231*
	(0.072)	(0.137)
Vakinankaratra	0.945***	-0.492***

	(0.073)	(0.128)
Bongolava	1.357***	0.087
	(0.074)	(0.157)
Sofia	-0.177	-0.308**
	(0.117)	(0.146)
Boeny	-0.097	0.068
	(0.103)	(0.121)
Betsiboka	0.675***	0.319***
	(0.082)	(0.112)
Melaky	-0.019	0.216*
	(0.105)	(0.119)
Alaotra-Mangoro	0.260***	0.244**
	(0.096)	(0.107)
Atsinanana	0.585***	-0.484***
	(0.094)	(0.123)
Analanjirofo	0.280***	-0.157
	(0.103)	(0.114)
Amoron'i Mania	1.386***	-0.398**
	(0.086)	(0.170)
Mahatsiatra Ambony	0.421***	-0.645***
	(0.091)	(0.104)
Vatovavy-Fitovinany	1.313***	-0.623***
	(0.091)	(0.166)
Atsimo-Atsinanana	-0.063	-0.805***
	(0.110)	(0.124)

Ihorombe	-0.519***	0.012
	(0.127)	(0.181)
Menabe	-0.157	-0.233*
	(0.110)	(0.134)
Atsimo-Andrefana	-0.333***	-0.592***
	(0.124)	(0.160)
Androy	-0.337**	0.266
	(0.146)	(0.201)
Anosy	0.665***	0.002
	(0.102)	(0.136)
Constant	-1.425***	6.910***
	(0.088)	(0.274)
Mills ratio		0.080
		(0.138)
Observations	25539	

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported.
Analamanga (region of the capital), the traditional religion,
Merina and the 15-20 years old are the omitted categories
of their respective group.

As shown in the descriptive statistics, one child out of five is working. That is, we do not observe child hours of work for 80% of the children. In addition, we only observe one regime at a time so that we observe the real amount of hours wished by the household for 10% of the children. To lower the censorship, we aggregate information at the household level.

That is, in both regimes, the dependent variable is the average number of hours worked by children (aged 6 to 15) per household. Note that since we are trying to classify households and not children (no child specific variable is in the switching equation), having information at the child level makes the estimation procedure more cumbersome and does not bring extra information. In a way, the results are improved because now all the households have the same weight regardless of the number of children they have[19]. The theoretical model only gives the amount of child work regardless of the number of children so that one can worry that the reduced-form equations differ. The form of the equation is not changed, but all the coefficients in each regime (not the switch) will need to be divided by the average number of 6-15 year old children per household in the economy before recovering the parameters.

When h^c equals zero, it is the result of a corner solution for the unobserved latent variable "average number of hours wished by each household and denoted by h^{c*}". For this reason, we use a Tobit model. We regress h^c on $\mathbf{Z}$. The variables included in Z were already enumerated and here only a short justification for their inclusion is provided. Household characteristics such as the parental level of education are important because educated parents usually value school more which can conflict with child work on the farm. Religion and ethnicity might induce different preferences towards child labor and the social stigma might differ among them. Household composition can clearly influence child labor quantity as it indicates the non-child workforce available, the amount of consumption needed and diversification strategies in terms of investing into some child(ren) (provided there are several children which is often the case). To the extent that child labor might be different across gender and because the information was aggregated at the household level, a variable indicating the proportion of girls among the 6-15 year old children is added. An

[19] If we do not collapse the information at the household level, then we give more weights to households with more children. Note that we always control for the number of children in the estimation.

index for productive assets was included and for the same reasons as land, we believe that productive assets should only influence child labor if the household is constrained. In the other regime, land and productive assets will have no effect because exogenous wealth is controlled for. The wage offered at the local level, will increase child labor as the theoretical model showed.

As shown in the theoretical model we need to control for expenditure. However to the extent that unobserved terms in the child labor supply function can be correlated with current expenditure (third variable potential bias), expenditure is an endogenous variable. Indeed, we are worried that there is a third variable bias issue. Child labor is indirectly bringing money by avoiding external labor to be hired or by freeing up adult household member time and allowing them to bring more money (which may in turn allow for higher expenditure) and richer households might be more able to forgo child labor through the wealth effect. Consequently expenditure is instrumented with permanent wealth (using an asset index, based on the methodology of Filmer and Pritchett, 2001[20]). as depicted in table 8. The first column displays the results of the regression of the number of hours of work of children before the instrumentation. The second column shows the same regression once the instrumentation has been undertaken and the last column shows the actual instrumentation. As can be seen, the wealth index is a very good predictor for total spending. It is positively related with spending and significant at the one percent level[21]. Interestingly, we can see that the endogeneity linked with spending led the local primary wage and the possession of a non farm enterprise to be downward biased. Other

[20]The asset index was computed using information on asset ownership (whether they own a chair, a table, a bed, a lamp, a fridge, a radio, other pieces of furniture, sewing machine, TV set, private car, bicycle, bike, camera, music instrument, fix phone, cellular phone, computer, a VCR, a video camera and other durable good. None of the productive assets were included (for example tractors, plow, carts,harrow, etc.). Housing information such as the material for the walls / floor/ ceiling, the source of lighting, the type of toilet, whether there is electricity or not, the source for the water, how the liter is getting rid of, the number of meter squares.

[21]The corresponding F-statistic is of 796.

variables were merely affected and have the expected sign in both regressions with the average age being positively correlated to the number of hours of work, education of the parents, in particular mother education having a negative impact on child work, Catholics and Protestants seem to distaste child work more than people following the traditional religion. In the end, we find that after instrumentation we can observe the effect of the local wage on child labor but not the effect of land.

Since land might have a non linear effect that might not be completely captured by the logarithmic transformation, we also classified land in three categories and run the same regressions which showed similar results except that the effect of land is now captured in the pooled regression as can be seen in table 9 and has now a positive effect on farm work[22]. That is, taking the nonlinear effect of land into account, it is possible to observe a positive relationship between land quantity and child labor, suggesting, as found in Pakistan and Ghana by Bhalotra and Heady (2003), Burkina Faso by Dumas (2007) and Mueller in rural Botswana (1984) that there are market imperfections on the land and the labor markets that affect a non negligible share of the households.

4.3 The switching equation

The equation determining in which regime households are more likely to belong to, is a function of different characteristics including market imperfection measures and household characteristics. Note that we do not observe the dependent variable in the switching equation. All we know is that the higher the value of this equation, the higher the probability that the household will be constrained[23]. Let C be the dependent variable of the equation, representing the amount of constraints. Whenever C is greater than some value $\bar{C}$ (that is

[22]We also categorize land in five groups and just as for the case with three categories, land is positive and significantly different from zero for high amount of land.

[23]It is so simply by construction as the higher C, the more likely to belong to the second regime, hence the switching equation represents the probability to be constrained

Table 8: Pooled estimation of the number of hours of child work with and without instrumentation, land in logarithm (then standardized with all other non dummy variables)

	Nb. Hours	Nb.hours	Spending
	No inst	IV: wealth index	
spending	-0.218***	-1.547***	
	(0.054)	(0.176)	
wage, prim. sector	0.046	0.195***	0.091***
	(0.039)	(0.046)	(0.010)
land, log	0.043	0.088	0.011
	(0.069)	(0.074)	(0.018)
prod. asset index	-0.007	-0.015	0.010
	(0.036)	(0.038)	(0.009)
age, years	0.687***	0.731***	0.016*
	(0.039)	(0.042)	(0.010)
education, mother	-0.439***	-0.356***	0.004
	(0.048)	(0.052)	(0.012)
education, father	-0.230***	-0.131***	0.019*
	(0.043)	(0.047)	(0.011)
nb children 6-15	0.308***	0.269***	-0.018
	(0.051)	(0.055)	(0.013)
household size	-0.039	0.221***	0.188***
	(0.053)	(0.065)	(0.014)
Catholics	-0.551***	-0.420***	0.042
	(0.120)	(0.129)	(0.032)
Protestants	-0.379***	-0.283**	-0.004
	(0.119)	(0.128)	(0.032)
Other religions	-0.095	-0.108	-0.018
	(0.116)	(0.125)	(0.032)
household head, female	-0.022	-0.083	-0.041
	(0.105)	(0.112)	(0.027)
rural	0.053	-0.018	-0.021**
	(0.039)	(0.042)	(0.010)
ratio of girls	-0.167*	-0.170	0.000
	(0.098)	(0.104)	(0.024)
non farm. entreprise	0.134	0.328***	0.161***
	(0.090)	(0.099)	(0.023)
wealth index			0.346***
			(0.012)
N	4966	4966	4966

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported. A two-step procedure was used for the instrumentation. Maximum likelihood estimation for the instrumentation and the pooled tobit gave very similar results. Ethnic dummies were included but are not shown. The omitted variables for religion and ethnicity are respectively the "Traditional" religion and the Merina.

Table 9: Pooled estimation of the number of hours of child work with and without instrumentation, land in three categories

	Nb. Hours	Nb.hours	Spending
	No inst	IV: wealth index	
spending	-0.213***	-1.539***	
	(0.054)	(0.176)	
wage, prim. sector	0.035	0.187***	0.092***
	(0.039)	(0.046)	(0.010)
land, group 2	0.265***	0.171*	-0.051**
	(0.089)	(0.095)	(0.023)
land, group 3	0.261***	0.256**	-0.010
	(0.095)	(0.101)	(0.024)
prod. asset index	-0.001	-0.012	0.008
	(0.036)	(0.038)	(0.009)
age, years	0.683***	0.727***	0.017*
	(0.039)	(0.042)	(0.010)
education, mother	-0.441***	-0.358***	0.004
	(0.048)	(0.052)	(0.012)
education, father	-0.230***	-0.132***	0.019*
	(0.043)	(0.047)	(0.011)
nb children 6-15	0.310***	0.270***	-0.018
	(0.051)	(0.054)	(0.013)
household size	-0.048	0.213***	0.189***
	(0.053)	(0.065)	(0.014)
Catholics	-0.537***	-0.412***	0.039
	(0.120)	(0.129)	(0.032)
Protestants	-0.362***	-0.274**	-0.008
	(0.119)	(0.128)	(0.032)
other religions	-0.100	-0.114	-0.018
	(0.116)	(0.125)	(0.032)
household head, female	0.012	-0.062	-0.047*
	(0.106)	(0.113)	(0.027)
rural	0.048	-0.020	-0.020**
	(0.039)	(0.042)	(0.010)
ratio of girls	-0.166*	-0.169	0.000
	(0.098)	(0.104)	(0.024)
non farm. entreprise	0.140	0.330***	0.160***
	(0.090)	(0.099)	(0.023)
wealth index			0.344***
			(0.012)
N	4966	4966	4966

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported. A two-step procedure was used for the instrumentation. Maximum likelihood estimation for the instrumentation and the pooled tobit gave very similar results. Ethnic dummies were included but are not shown. The omitted variables for religion and ethnicity are respectively the "Traditional" religion and the Merina. Land was divided in three categories, the first category is the omitted group which corresponds to little land.

set to zero without loss of generality), people will be assigned to the constrained regime. Otherwise they will be said to not be constrained (therefore assigned to first regime).

Variables included in the switch that do not directly impact child labor will help the statistical identification, The percentage of salaried households in the primary sector, the percentage of salaried households in other sectors, the share of inherited plots and the share of purchased plots are introduced in the switch but excluded from the regimes. Access to credit is also included through the availability (dummy variables) of a bank (formal) or a moneylender (informal sector). Though access to credit might have an impact on child labor, we believe that credit infrastructure's availability might lead to less exclusion but the extent to which it would influence the actual number of hours is doubtful.

4.4 Optimization Issues

The likelihood function to estimate is a sum of probability distributions for the two regimes, weighted by the probability to fall into each regime. The log of a sum is notoriously known to be quite troublesome to maximize and the EM or Expectation Maximization algorithm first introduced by Hartley (1958) and popularized by Dempster, Laird and Rubin (1977) is used. Hartley (1978) suggested its application for switching regression models and we draw on his work to apply it when regimes themselves are censored by zero corner solutions. The EM algorithm simply consists in writing the expectation that one would have if all data were available conditional on the data actually available, and then, maximizing it. The parameters obtained from the maximization are then introduced into the expectation which is again maximized and the process continues until convergence is achieved[24].

[24]Train (2003) provides a more general yet more detailed explanation of the EM algorithm. Gupta and Chen (2011) and McLachland and Krishnan (2008) give the theory and examples of applications of the EM algorithm. More recently Caudill (2012) applied the algorithm to a mixture of two censored normal distributions.

The (log) likelihood function is the following with f_j the probability density function in regime j and F_j the respective cumulative distribution function:

$$\log p(H^c|\theta) = \log \prod_{i=1, h_i^c \neq 0}^{n} (\Phi(-X_i\beta) f_1(h_i^c|\theta_1) + \Phi(X_i\beta) f_2(h_i^c|\theta_2)) \times \prod_{i=1, h_i^c = 0}^{n} (\Phi(-X_i\beta) F_1(h_i^c|\theta_1) + \Phi(X_i\beta) F_2(h_i^c|\theta_2))$$

If the wished amount of child working hours h_i^{c*} and the regime membership of each household r_i^*, could be observed, the likelihood function to maximize would be :

$$\begin{aligned} \log p(H^c, R^*, H^{c*}|\theta) &= \log \prod_{i=1}^{n} [p(h_i^c, h_i^{c*}|r_i^*, \theta) p(r_i^*|\theta)] \\ &= \log \prod_{i=1}^{n} [p(h_i^{c*}|r_i^*, \theta) p(r_i^*|\theta)] \end{aligned}$$

whose expectation can be maximized and, as shown in the appendix, it simplifies into maximizing separately three different terms corresponding to parameters for the three equations (the switch, the constrained regime and the non constrained regime)

As identification of the coefficients in the switching regression is not possible (they cannot be disentangled from the standard error σ), we impose that the regime-determining equation (or "switch") has normally distributed error terms of mean zero and standard deviation one. Note that setting σ to one will not have any implication in our structural model as it only appears in the switch. For the switch we mainly want to see the variables that significantly predict the constraint status of households and their sign so that setting the variance to one does not affect the interpretation.

5 Results

Table 10 provides the results for the switching regression. Panel A displays the output obtained from the switching regression strategy. Panel B and C uses the regime division in Panel A to reestimate the two regimes (and are discussed afterwards). From the results in Panel A, it seems that the regimes have been identified. A positive effect of land on child hours of work is found in the constrained regime and a positive effect of the local wage on child hours of work is found in the regime with full access. For the switch, we find that the richer the household, the less likely it is to be constrained. A high wage in the primary sector indicates higher difficulty for the household to find labor and therefore to be constrained. That is we are more likely to have households that cannot find workers than households that would be willing to work outside of the farm for farm work. This result is not surprising since we have chosen to focus on households with land who are by definition less likely to be constrained by low paid farm jobs than landless households.

Being in a rural area non surprisingly means less access to the land and the labor markets. The number of individuals could be positive because once the number of children are controlled for, the variation in the number of people might simply be other adults that are not able to work (if we believe that if they could work and earn enough money from their work, they would leave the household). As most households are made up of the parents and their children, most variation in the number of individuals in the household outside 6 to 15 year old children is more likely to come from under five year old or grand parents who are "dependent". Obviously there could also be children above 15 but the variation is more likely to come from the dependents[25]. In that sense having more individuals is more constraining.

[25]In the full, non pooled model, including variables with low variations as well as dummies had to be limited because of computational issues. For this reason the distinction among non 6-15 years old was not implemented.

The exclusion variables for the switch also seem to point to a higher likelihood to fall into the constrained regime as the latent variable C rises, thereby confirming that regimes have been identified. The proportion of salaried workers in the primary sector significantly decreases the probability to lack access to the markets. Obviously the more salaried workers there are, the easier it is to find workforce. The proportion of salaried workers in the other sectors increases the likelihood to be constraineds[26]. The proportion of inherited plots at the local level undoubtedly hinders transaction on the land markets. Having a high proportion of plots either titled or registered on a cadaster at the local level increases market access[27]. Finally, it seems that if one's ethnic group is not the (numerically) dominating group at the local level, then, the household is less likely to have a full access to the markets.

Looking at the regimes, we find that wealth has a negative impact on child labor. Controlling for this wealth effect, we find that land increases child hours of work when people lack access to the markets consistent with the theoretical model[28] and we find a slightly significant effect of land in the switch, indicating a higher probability to be constrained as land rises. The wage offered at the local level clearly increases the number of hours of work when households can access the market. As in the pooled regime in table 8, we find that as age increases, the number of hours increases, regardless of the regime considered. Mother education reduces child labor whether the household is constrained or not. The effect is always stronger than the effect of father education (which is only significant when access to the land and labor markets is limited). Having a non farm

[26] If a person is neither salaried in the primary sector, nor in the other sectors then that person could be not working at all or working but not salaried.

[27] Note that for each of the proportion variables computed at the local level (for salaried worker, sharecropping, or plot rental) we systematically excluded the household so that the variables show local characteristics independently of households' decisions.

[28] Note that from the implementation itself, land has no impact in the other regime, however it could have been that land still has no impact in the regime without market access. It was even more possible as the pooled tobit did not seem to show any significant impact of land on child labor.

enterprise does not seem to have an effect on child working hours regardless of whether the household has access to the markets[29].

Table 10 panel B shows the estimation taking into account the results from the switch. That is, we use C^* to determine to which regime households belong. Indeed the whole sample is used to estimate both regimes whereas only some of the households are constrained. To check if results are robust, it is important to reestimate the two tobit equations (one for each regime) only on the corresponding subsamples (as identified by the switching regression strategy). Reestimating the model also allows for more flexibility such as the use of locality fixed effects. The addition of locality fixed effects did not change the results and for that reason only the results without fixed effects are reported[30].

One of the main difference is also in the instrumentation as it was instrumented on the whole sample therefore having land or local wage as instruments depending on the regime. That said, we should emphasize that when the household is in the regime with full market access, land does not influence child labor. Consequently, land becomes a potential instrument for exogenous wealth[31]. In the constrained regime, land is only a control and since the local wage does not appear in the final regression, local wage becomes an instrument. Even if local wage has no influence on child labor it can be an indicator of the general wealth of the village and therefore predict expenditures without affecting

[29]Estimations without this variable led to similar results for all equations in all specifications.

[30]Once the switching model identified the two regimes, land was added to panel B in the regime with market access and it did not show any significant result. Wage was added when there is a limited access and it did not have any significant impact on the amount of child work either. Note that it is a limited way of testing the assumptions of the absence of impact. However, it is a result of the theoretical model which was based on a limited number of assumptions. Secondly, relaxing the assumption of no impact at the level of the switching regressions would lead to very different results: two regimes always have a better prediction that one regime so that the switching regression model will (almost) always significantly predict better than one. It could be that the switch identify outliers in one regime and the rest of the observations in another regime. For that reason it is necessary to have restrictions to identify the two regimes.

[31]Remember that we cannot know ahead what is the regime households belong to so we are instrumenting on the whole sample. Because we exclude land in the full access regime, then it is as if we had included it as an instrument. The case is similar for the regime with a restricted market access with wage instead of land.

Table 10: Switching regression estimation, land in logarithm

		Panel A		Panel B		Panel C	
	switch	regime 1	regime 2	regime 1	regime 2	regime 1	regime 2
constant	-0.426	0.272*	0.960				
	(0.366)	(0.149)	(0.798)				
land, log	0.119*	0.000	0.449***		0.480**		0.461**
	(0.066)	.	(0.169)		(0.220)		(0.220)
expenditures	-0.346	-1.017***	-1.953***	-1.362***	-0.896**	-1.471***	-0.863**
	(0.333)	(0.135)	(0.610)	(0.140)	(0.409)	(0.176)	(0.422)
wage, primary	0.267***	0.212***	0.000	0.293***		0.285***	
	(0.087)	(0.034)	.	(0.036)		(0.042)	
productive asset	-0.042		-0.090		-0.099		-0.076
	(0.061)	.	(0.148)		(0.094)		(0.095)
average age	0.105	0.467***	1.296***	0.575***	1.181***	0.579***	1.173***
	(0.075)	(0.039)	(0.209)	(0.041)	(0.093)	(0.049)	(0.093)
education, mother	-0.107	-0.117***	-1.655 ***	-0.120***	-1.126***	-0.120***	-1.125***
	(0.091)	(0.040)	(0.327)	(0.039)	(0.158)	(0.046)	(0.159)
education, father	-0.068	-0.024	-0.556 ***	-0.036	-0.511***	-0.017	-0.525***
	(0.079)	(0.037)	(0.200)	(0.036)	(0.133)	(0.044)	(0.133)
nb. children 6-15	-1.301***	0.000	0.055	-0.025	-0.065	0.016	-0.054
	(0.160)	(0.043)	(0.294)	(0.045)	(0.201)	(0.054)	(0.202)
nb. individuals	0.300**	0.203***	0.127	0.288***	-0.056	0.290***	-0.059
	(0.122)	(0.049)	(0.225)	(0.050)	(0.161)	(0.060)	(0.164)
Catholics	0.015	-0.236**	-1.083**	-0.354***	-1.084***	-0.361***	-1.088***
	(0.212)	(0.097)	(0.522)	(0.100)	(0.341)	(0.120)	(0.343)
Protestants	-0.097	-0.319***	0.222	-0.422***	-0.176	-0.389***	-0.204
	(0.212)	(0.093)	(0.475)	(0.099)	(0.325)	(0.120)	(0.326)
other religions	0.303	-0.150	0.265	-0.263***	0.234	-0.189	0.186
	(0.210)	(0.104)	(0.429)	(0.101)	(0.292)	(0.124)	(0.294)
ethnic group 1	0.560*	-0.401***	0.482	-0.447***	0.205	-0.552***	0.224
	(0.287)	(0.116)	(0.567)	(0.120)	(0.419)	(0.143)	(0.421)
ethnic group 2	0.251	-0.315***	-0.901	-0.359***	-0.702	-0.343***	-0.744
	(0.270)	(0.106)	(0.651)	(0.110)	(0.451)	(0.132)	(0.460)
ethnic group 3	0.700 ***	-0.588***	-1.624**	-0.775***	-1.173***	-0.852***	-1.152***
	(0.251)	(0.111)	(0.638)	(0.108)	(0.422)	(0.130)	(0.421)
ethnic group 5	0.871***	-0.215*	0.594	-0.236**	0.498	-0.179	0.456
	(0.253)	(0.113)	(0.565)	(0.111)	(0.386)	(0.132)	(0.391)
ethnic group 6	0.314	0.164*	0.054	0.235**	0.012	0.267**	-0.030
	(0.253)	(0.097)	(0.601)	(0.100)	(0.423)	(0.121)	(0.425)
household head female	-0.446**	-0.016	-0.776	-0.019	-0.555*	-0.007	-0.551*
	(0.193)	(0.081)	(0.473)	(0.086)	(0.304)	(0.103)	(0.305)
rural area	0.204***	-0.072**	0.421**	-0.109***	0.294***	-0.115***	0.302***
	(0.077)	(0.033)	(0.177)	(0.033)	(0.111)	(0.040)	(0.111)
ratio of girls	-0.273	0.106	-1.343***	0.157*	-1.117***	0.128	-1.114***
	(0.175)	(0.092)	(0.407)	(0.095)	(0.235)	(0.113)	(0.236)
non farm entreprise	-1.086 ***	0.043	-0.676	0.094	-0.401	0.118	-0.368
	(0.205)	(0.068)	(0.636)	(0.071)	(0.354)	(0.086)	(0.360)
ethnic minority	0.445***						
	(0.172)						
religion minority	0.258**						
	(0.124)						
bank	0.106						
	(0.139)						
moneylender present	-0.258*						
	(0.156)						
Proportion in the village of							
salaried, primary sector	-1.149***						
	(0.365)						
salaried, other sectors	4.631***						
	(0.787)						
plots registered	-2.153***						
	(0.348)						
plots inherited	0.199 ***						
	(0.074)						
N	4966	4966	4966	2550	2416	2550	2416

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported. The traditional religion, Merina (ethnic group 4) and "little land" are respectively the religion, ethnic and land group omitted variables. The proportion of plots inherited and shared were computed using the plots surface and not the number. Registered plots referred to plots that have been either titled or registered on a cadaster. Productive asset is an index of productive assets at the household level, animals are excluded.Regime 1 corresponds to the regime with access to the land and labor markets. Regime 2 corresponds to the regime where access is limited or inexistant.

directly child labor. Therefore it also becomes a potential instrument. The issue here is that both wage and land are used as predictors for expenditures when only one of them (depending on the regime) should be used. Panel C instruments and estimates each regime again, according to the regime separation obtained from the switch. The main results are the same: land clearly increases the amount of child labor (that is, there is a substitution effect) for households with limited access (regime 2) and local wage increases child labor when there is full market access. Mother education decreases child labor. The results for the switch are a bit less interesting in the sense that more or less all variables are predictors of the switch[32] so we will not try to give an interpretation here, results are still shown in the last column.

Table 11 shows the results of the switching regimes with the land variable in three categories (to allow for a non linear effect), the first category for households with little land (which is the omitted category in the estimation). Results are practically the same: the regimes seem to be identified (land group is positive and significant for households with no access and wage is indeed positive and significant for households with access). For that reason, we focus on differences to see what the non linearity brings forward. Focusing on the switching equation, it seems that land decreases the likelihood for a household to face a restricted access. Knowing that we are controlling for wealth, that effect does not reflect higher wealth. It is possible that land comes with a certain prestige that cannot be captured by the expenditure variable and which is clearly not linearly increasing with land itself. Only the highest category of land (group 3 in the estimations) displays this effect. Though not strong, The ratio of girls seems to lower the probability to be deprived access to the market as this ratio increases. However, looking at the effect of this ratio on the number of hours worked, we find that it is negative and strongly significant when the household

[32]It is not exactly a perfect prediction because each observation has a non zero probability to be in the first regime instead of being assigned with probability one to one of the two regimes.

Table 11: Results from the switching regression model and post estimation, with land in groups

	Panel A: switching regression			Panel B: Post estimations		
	regime 1	regime 2	switch	regime 1	regime 2	switch
constant	0.361 **	-0.179	-0.686 *	0.301**	-1.278	-1.997***
	(0.147)	(1.033)	(0.360)	(0.151)	(0.777)	(0.220)
land, group 2		1.512 ***	0.060		1.235***	0.226***
		(0.502)	(0.135)		(0.307)	(0.085)
land, group 3		0.899*	-0.343 **		0.560	-0.754***
	.	(0.536)	(0.150)		(0.373)	(0.095)
local wage (primary)	0.220 ***		0.310***	0.302***		0.855***
	(0.033)	.	(0.085)	(0.036)		(0.050)
expenditures	-1.037***	-1.522 **	-0.302	-1.248***	-0.694	-0.786***
	(0.135)	(0.772)	(0.313)	(0.139)	(0.569)	(0.156)
average age	0.481 ***	1.522 ***	0.105***	0.582***	1.414***	0.259***
	(0.039)	(0.283)	(0.071)	(0.039)	(0.123)	(0.036)
education, mother	-0.118 ***	-2.171***	-0.074	-0.158***	-1.815***	-0.274***
	(0.040)	(0.462)	(0.085)	(0.039)	(0.239)	(0.046)
education, father	-0.039	-0.626 ***	-0.040	-0.055	-0.505***	-0.068
	(0.037)	(0.251)	(0.073)	(0.037)	(0.175)	(0.042)
nb. Children 6-15	-0.016	-0.146	-1.258 ***	-0.090**	-0.751**	-3.254***
	(0.042)	(0.404)	(0.145)	(0.043)	(0.302)	(0.114)
nb. Individuals	0.220 ***	-0.236***	0.233 **	0.289***	-0.504**	0.546***
	(0.047)	(0.315)	(0.113)	(0.049)	(0.227)	(0.063)
Catholics	-0.271***	-1.101	0.014	-0.396***	-0.877*	0.020
	(0.096)	(0.690)	(0.204)	(0.100)	(0.457)	(0.117)
Protestants	-0.342 ***	0.482 ***	-0.150	-0.469***	0.321	-0.455***
	(0.091)	(0.641)	(0.206)	(0.098)	(0.443)	(0.119)
Other religions	-0.179 *	0.556	0.289	-0.195*	0.553	0.874***
	(0.102)	(0.577)	(0.193)	(0.101)	(0.385)	(0.119)
ethnic group 1	-0.372 ***	0.989 ***	0.727 ***	-0.522***	1.340**	1.960***
	(0.115)	(0.773)	(0.271)	(0.121)	(0.589)	(0.159)
ethnic group 2	-0.287***	-0.790	0.409	-0.364***	-0.275	1.136***
	(0.108)	(0.857)	(0.256)	(0.111)	(0.653)	(0.149)
ethnic group 3	-0.577 ***	-1.598 ***	0.771***	-0.719***	-0.809	2.085***
	(0.109)	(0.847)	(0.243)	(0.110)	(0.599)	(0.146)
ethnic group 5	-0.184 *	1.110	1.006 ***	-0.159	1.395**	2.750***
	(0.111)	(0.774)	(0.243)	(0.109)	(0.566)	(0.154)
ethnic group 6	0.190 **	0.112 ***	0.416 *	0.196*	0.381	1.098***
	(0.096)	(0.828)	(0.250)	(0.101)	(0.607)	(0.150)
female headed hh.	-0.030	-1.121 *	-0.498 ***	-0.066	-1.159***	-1.344***
	(0.082)	(0.650)	(0.186)	(0.086)	(0.428)	(0.110)
rural area	-0.067 **	0.509**	0.189 ***	-0.095***	0.421***	0.494***
	(0.033)	(0.231)	(0.072)	(0.033)	(0.151)	(0.043)
ratio of girls	0.055	-1.670 ***	-0.295*	0.052	-1.316***	-0.585***
	(0.091)	(0.544)	(0.165)	(0.092)	(0.303)	(0.090)
non farm. Entreprise	0.034	-1.393	-1.042***	-0.000	-1.573***	-2.701***
	(0.069)	(0.890)	(0.194)	(0.072)	(0.579)	(0.124)
productive asset		-0.087	-0.022***		-0.019	-0.002
		(0.199)	(0.056)		(0.131)	(0.035)
ethnic minority			0.273*			0.747***
			(0.155)			(0.103)
religion minority			0.216 *			0.522***
			(0.115)			(0.076)
Bank present			0.134			0.279***
			(0.131)			(0.086)
moneylender present			-0.226			-0.453***
			(0.145)			(0.092)
Proportion in the village of						
salaried, primary			-1.117 ***			-2.818***
			(0.341)			(0.218)
salaried, other sectors			4.617***			11.675***
			(0.721)			(0.513)
registered plots			-1.939			-5.177***
			(0.312)			(0.228)
inherited plots			0.187***			0.486***
			(0.069)			(0.045)
N	4966	4966	4966	2803	2163	4966

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported. The traditional religion, Merina (ethnic group 4) and "little land" are respectively the religion, ethnic and land group omitted variables. The proportion of plots inherited and shared were computed using the plots surface and not the number. Registered plots referred to plots that have been either titled or registered on a cadaster. Productive asset is an index of productive assets at the household level, animals are excluded.Regime 1 corresponds to the regime with access to the land and labor markets. Regime 2 corresponds to the regime where access is limited or inexistant.

is constrained. That is, having (proportionally) more girls decreases the probability to be constrained, but if the household lacks access, then having more girls decreases the number of hours worked when hours of work are most needed. One explanation is that girls will free up time from other household members (typically the mother) that will allow a specialization of the tasks: girls would not work at all on the farm and their mother might go and spend more time on the farm.

Results for the switching estimations with land in level were very similar[33] and only the reestimations (that is, the estimation obtained once the households have been classified into the two regimes by the switching equation) on each regime are shown in the appendix, table 13. The first column shows the switching equation (reestimated according to the regime membership given by the switching model). The second and third columns refer to the unconstrained and constrained regimes respectively. The last two columns show the first and respectively second regimes once the instrumentation has been run again (that is, once the switching equation has indicated to which regime households belong) excluding the variables for land and productive asset in the first regime and excluding wage for the second regime. One can see that results barely change after the reinstrumentation. Land has a strong, significant and positive impact on child labor in the first regime. In the second one, the wage offered has a strong and positive impact on child hours of work. The effect of expenditure is again stronger (in absolute value) in the regime with market access suggesting an increasing marginal productivity of labor[34] only consistent with small amount of land. Mother education still has a stronger effect when market access is not granted. The ownership of a non farm enterprise has no impact on child hours of work but significantly

[33]Without logarithm (that is, in level), the production function would write as $g(h,k) = \frac{ce^{\gamma_k k}}{\gamma_h}(1-e^{-\gamma_h h})$ and the shadow wage as $\tilde{w}^c = p\delta ce^{-\gamma_h h^c + \gamma_k k})$ which is not realistic since zero amount of land produces a strictly positive amount of output. To the extent that only households with land are examined such a case would never occur so that it is still worth investigating what happens when land is in level.

[34]This can be seen from examining the equations for the parametrization. The parameters are computed at the end of the section.

reduces the probability to be constrained though it could clearly be endogenous[35]. Female headed households seem to be less likely to be constrained on the markets. Globally regardless of the specification, it seems that mother education significantly reduces child hours of work especially when market access fails.

Table 12: Coefficient estimates from table 10

Coefficient on	Regime 1		Regime 2	
	Notation	Value	Notation	Value
Constant	A_1	0.272	A_2	(0.960)
Expenditures	B_1	-1.017	B_2	-1.953
Wage	C_1	0.212		
Land			C_2	0.449

To improve our understanding of the estimates, results from the parametrization subsection are used. We do the exercise for the estimates in table 10 which evaluates the effect of land in logarithm, corresponding to the functional form in equation (17). Table 12 recalls the notations used and values obtained. Accordingly, we obtain from the first regime β : $1/0.212 = 4.717$ and α: $-(-1.017) * 4.717 = 4.797$. The ratio δ/b is (with a set to one): $e^{4.717*0.272} * 4.717/4.797 = 3.625$. These estimations simply concur with the assumptions made in our theoretical model, namely that more consumption increases utility but the extra utility from higher consumption is increasing with consumption. In the same way, more work leads to a higher diutility and the marginal disutility of work is increasing with hours of work. Finally, using information on both regimes one obtains: $\gamma_h = (-4.797/-1.953) - 4.717 = -2.26$ and $\gamma_k = 0.449(-2.26 + 4.717) = 1.1$. That is, the marginal productivity of labor increases with land (γ_k is positive) and the marginal productivity of labor increases with labor. The later goes against the assumption made in the theoretical model (where all parameters are positive). However because the amount

[35]The results still hold when dropping this variable from the switch and the regimes

of workers on a farm (and farm size) is typically small, it is possible that the marginal productivity of labor increases with labor for small amount of labor as there might be a better division of the tasks (it then eventually decreases but the plots are too small to allow for the observation of this feature). Along with our assumptions, the marginal productivity of labor increasing with land is supported by the data.

Comparing the coefficients in the two regimes, we find that the effect of the variables for mother education, household composition or even religion are systematically stronger when households lack access to the markets (it is not the case for ethnicity but to the extent that the ethnic group usually predicts well the area we prefer to be cautious regarding the interpretation of their potential difference). It can be that any effect is more visible when household lack access. The first order conditions in the theoretical model do indicate that the preferences have a more preponderant role when there is no access.

6 Discussion and conclusion

This paper has used a simple theoretical model combined with a switching regression strategy to identify households limited in their access to the land and the labor markets in Madagascar. The work undertaken went beyond finding the wealth paradox and investigated for whom the substitution effect on child non working time appeared. The results suggest that around 45% of households cannot fully access the land and the labor markets. These families have their children work more than they would otherwise do had they had access even when they do not display the wealth paradox. To the extent that this effect only appears when households need access to the market, it can be seen as a lower bound. Another result is that numerically dominated ethnic or religious groups at the local level are systematically more likely to belong to the regime with a restricted access. This seems

to reflect agreements on the local informal labor market that would need to be improved.

The present work is limited by the use of household size as an exogenous variable. The number of household members might be endogenous in the sense that households could adjust their workforce depending on how much land they have to farm. Excluding parents whose only aim would be to have children for the sake of having workforce (raising a child being costlier than the benefits from farm child work, the exclusion is reasonable), arrangement within the extended family or other individuals in terms of household composition can still take place. If a household is indeed adjusting its size according to the amount of land there is to farm, then such adjustment should only lower the positive effect of land on child labor. At least such effect cannot make the substitution effect stronger so that even in the presence of strategic household size, the household identified as lacking access should still be considered as indeed lacking access.

The results highlight other negative consequences of market imperfections on the land and the labor markets. It is therefore convenient to argue that improving access for the land and the labor markets has more benefits than usually recognized and child labor can be tackled through new policy tools. However two points should be made. First, it is not clear whether the number of extra hours worked by children impede a more productive use of their time such as schooling. To that extent it is difficult to assess the benefits of a reduction in child hours on the farm and to compare them with the cost of improving land tenure (which according to Jacoby and Minten (2007) should not be undertaken[36]). Second, our empirical analysis looks at peasant households. That is, the results hold for current peasant households and the representativity of our sample might not extend to households that would be new owners of land if the land market was to be improved. For these two reasons, improving labor market access seems like the better option.

[36]Their benefit analysis comprised land productivity, land investment and land value of titled plots compared to untitled plots.

References

Accelerating action against child labour. Report of the Director-General, International Labour Conference, 99th session, 2010. Report, May 2010.

K. Basu, S. Das, and B. Dutta. Child labor and household wealth: Theory and empirical evidence of an inverted-u. *Journal of Development Economics*, 91(1):8–14, Jan. 2010.

S. Bhalotra. Is child work necessary?*. *Oxford Bulletin of Economics and Statistics*, 69(1):29–55, 2007.

S. Bhalotra and C. Heady. Child farm labor: The wealth paradox. *The World Bank Economic Review*, 17(2):197 –227, Dec. 2003.

M. Browning. *Modelling commodity demands and labour supply with m-demands*. Institute of Economics, University of Copenhagen, 1999.

S. B. Caudill. A partially adaptive estimator for the censored regression model based on a mixture of normal distributions. *Statistical Methods & Applications*, 21(2):121–137, June 2012.

H. Congdon Fors. The determinants of rural child labor: an application to India. *rapport nr.: Working Papers in Economics 256*, 2007.

A. de Janvry, M. Fafchamps, and E. Sadoulet. Peasant household behaviour with missing markets: Some paradoxes explained. *Economic Journal*, 101(409):1400–417, 1991.

A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.

C. Dumas. Why do parents make their children work? A test of the poverty hypothesis in rural areas of Burkina Faso. *Oxford Economic Papers*, 59(2), Apr. 2007.

C. Dumas. Market imperfections and child labor. *World Development*, 42:127–142, 2013.

D. Filmer and L. H. Pritchett. Estimating wealth effects without expenditure data–or tears: an application to educational enrollments in states of india. *Demography*, 38(1):115–132, Feb. 2001.

H. C. Fors. Child labour: A review of recent theory and evidence with policy implications. *Journal of Economic Surveys*, 26(4):570–593, Sept. 2012.

P. C. Godfrey. Toward a theory of the informal economy. *The Academy of Management Annals*, 5(1):231–277, 2011.

C. Grootaert and R. Kanbur. Child labor: a review. *World Bank Policy Research Working Paper*, (1454), 1995.

M. Gupta and Y. Chen. *Theory and Use of the EM Algorithm.* Now Pub, 2011.

H. Hartley. Maximum likelihood estimation from incomplete data. *Biometrics*, 14(2):174–194, 1958.

M. J. Hartley. Estimating mixtures of normal distributions and switching regressions: Comment. *Journal of the American Statistical Association*, 73(364), Dec. 1978.

G. J. McLachlan and T. Krishnan. *The EM algorithm and extensions.* John Wiley & Sons, Mar. 2008.

E. Mueller. The value and allocation of time in rural botswana. *Journal of Development Economics*, 15(1-3):329–360, June 1984.

D. Ray. *Development economics.* Princeton University Press, 1998.

I. Singh, L. Squire, and J. Strauss. A survey of agricultural household models: Recent findings and policy implications. *The World Bank Economic Review*, 1(1):149–179, 1986.

D. Stifel, F. H. Rakotomanana, and E. Celada. *Assessing labor market conditions in Madagascar, 2001-2005.* World bank, 2007.

K. Train. *Discrete choice methods with simulation.* Cambridge University Press, 2003.

7 Appendix

Table 13: Switching regression model output with land in level

	r1: full access		r2: constrained		switch	
constant	0.267	*	0.724		-0.321	
	(0.137)		(0.589)		(0.427)	
land			0.837	***	-0.174	
			(0.243)		(0.145)	
local wage (primary)	0.208	***			0.343	***
	(0.031)				(0.103)	
expenditures	-1.013	***	-1.079	**	-0.321	
	(0.126)		(0.428)		(0.389)	
age	0.458	***	1.061	***	0.270	***
	(0.038)		(0.146)		(0.094)	
education, mother	-0.114	***	-1.054	***	-0.076	
	(0.037)		(0.193)		(0.102)	
education, father	-0.028		-0.490	***	-0.187	*
	(0.034)		(0.146)		(0.097)	
nb. children 6-15	0.027		0.165		-1.575	***
	(0.040)		(0.231)		(0.208)	
household size	0.206	***	-0.053		0.354	**
	(0.045)		(0.173)		(0.142)	
Catholics	-0.236	***	-0.938	**	-0.097	
	(0.089)		(0.393)		(0.241)	
Protestants	-0.300	***	-0.021		-0.188	
	(0.086)		(0.357)		(0.237)	
other religions	-0.187	**	0.194		0.225	
	(0.094)		(0.325)		(0.240)	
ethnic group 1	-0.368	***	0.327		0.778	**
	(0.106)		(0.453)		(0.335)	
ethnic group 2	-0.312	***	-0.490		0.484	
	(0.099)		(0.500)		(0.309)	
ethnic group 3	-0.614	***	-1.072	**	0.862	***
	(0.099)		(0.481)		(0.297)	
ethnic group 5	-0.251	**	0.554		1.272	***
	(0.107)		(0.450)		(0.310)	
ethnic group 6	0.175	**	0.125		0.498	*
	(0.088)		(0.471)		(0.293)	
female headed hh.	0.002		-0.546		-0.552	**
	(0.075)		(0.344)		(0.229)	
rural	-0.072	**	0.289	**	0.226	**
	(0.030)		(0.132)		(0.090)	
ratio of girls	0.083		-0.947	***	-0.395	*
	(0.085)		(0.285)		(0.206)	
non farm. enterprise	0.075		-0.277		-1.225	***
	(0.063)		(0.429)		(0.246)	
productive assets	0.000		-0.090		-0.128	*
	.		(0.109)		(0.076)	
ethnic minority					0.548	***
					(0.213)	
religion minority					0.371	**
					(0.153)	
Bank present					0.074	
					(0.164)	
moneylender present					-0.261	
					(0.189)	
Proportion in the village of						
salaried, primary					-1.672	***
					(0.468)	
salaried, other sectors					5.286	***
					(0.986)	
registered plots					-2.493	***
					(0.438)	
inherited plots					0.231	***
					(0.087)	
sigma	1.048	***	3.047	***		
	(0.024)		(0.251)			

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported. The traditional religion, Merina and "little land" are respectively the religion, ethnic and land group omitted variables. The proportion of plots inherited and shared were computed using the plots surface and not the number. Registered plots referred to plots that have been either titled or registered on a cadaster. Productive asset is an index of productive assets at the household level, animals are excluded.

Table 14: Results from the switching regression model and post estimation, with land and squared land

	Switching regression					
	regime 1		regime 2		switch	
constant	0,335	**	1,254		-0,610	
	(0,155)		(0,921)		(0,371)	
land			1,323	***	-0,136	
			(0,492)		(0,156)	***
land, squared			-0,837	*	0,020	
			(0,442)		(0,092)	
local wage (primary)	0,206	***	0,000		0,273	***
	(0,035)				(0,087)	
expenditures	-0,972	***	-1,848	***	-0,322	
	(0,140)		(0,670)		(0,319)	***
age	0,479	***	1,396	***	0,102	
	(0,042)		(0,235)		(0,075)	
education, motherUM	-0,158	***	-1,981	***	-0,062	
	(0,042)		(0,382)		(0,092)	***
education, father	-0,070	*	-0,600	***	-0,005	
	(0,039)		(0,214)		(0,078)	
nb. Children	-0,032		-0,026		-1,303	***
	(0,045)		(0,328)	***	(0,160)	
household size	0,194	***	0,004		0,286	**
	(0,050)		(0,251)		(0,121)	***
Catholics	-0,255	**	-1,113	*	0,026	
	(0,101)		(0,579)		(0,213)	**
Protestants	-0,325	***	0,378		-0,098	
	(0,097)		(0,528)		(0,212)	
Other religions	-0,161		0,377		0,257	
	(0,107)		(0,481)		(0,211)	
ethnic group 1	-0,381	***	0,626		0,674	**
	(0,120)		(0,628)		(0,290)	
ethnic group 2	-0,280	**	-0,979		0,361	
	(0,113)		(0,713)		(0,265)	
ethnic group 3	-0,597	***	-1,647	**	0,725	***
	(0,115)		(0,699)		(0,252)	
ethnic group 5	-0,198	*	0,812		0,927	***
	(0,117)		(0,633)	***	(0,253)	
ethnic group 6	0,176	*	0,059		0,426	*
	(0,101)		(0,672)		(0,259)	
female headed	-0,031		-0,857	*	-0,466	**
	(0,084)		(0,518)		(0,196)	
rural	-0,072	**	0,465	**	0,176	**
	(0,035)		(0,196)		(0,077)	
ratio of girls	0,060		-1,553	***	-0,320	*
	(0,098)	***	(0,459)		(0,176)	
non farm enterprise	0,027		-1,022		-1,057	***
	(0,072)		(0,721)		(0,201)	
productive assets			-0,095		-0,025	
			(0,165)		(0,061)	

7.1 Identification results

With an exponential form for u and v such as:

$$\begin{aligned} u &= -a\exp(-\alpha x) \\ v_c &= b\exp(\beta h^c) \end{aligned}$$

we obtain the following derivatives:

$$\begin{aligned} u' &= \alpha a\exp(-\alpha x) > 0 \\ v' &= \beta b\exp(\beta h^c) > 0 \\ u'' &= -\alpha^2 a\exp(-\alpha x) < 0 \\ v'' &= \beta^2 b\exp(\beta h^c) > 0 \end{aligned}$$

and therefore the FOC on child labor in the non constrained regime, $u'w_c = v'$, becomes:

$$\begin{aligned} & \alpha a e^{-\alpha x} w_c = \beta b e^{\beta h^c} \\ \Leftrightarrow \quad & \ln(\alpha a e^{-\alpha x} w_c) = \ln(\beta b e^{\beta h^c}) \\ \Leftrightarrow \quad & \ln\alpha a + \ln e^{-\alpha x} + \ln w_c = \ln\beta b + \ln e^{\beta h^c} \\ \Leftrightarrow \quad & h^c = \tfrac{1}{\beta}\ln\left(\tfrac{\alpha a}{\beta b}\right) + \tfrac{1}{\beta}\ln w_c - \tfrac{\alpha}{\beta}x \end{aligned}$$

since we have $w_c = \delta w$, the later equation rewrites as:

$$\begin{aligned} h^c &= \frac{1}{\beta}\ln\left(\frac{\alpha a}{\beta b}\right) + \frac{1}{\beta}\ln \delta_c w - \frac{\alpha}{\beta}x \\ h^c &= \frac{1}{\beta}\ln\left(\frac{\alpha a}{\beta b}\right) + \frac{1}{\beta}\ln w + \frac{\delta_c}{\beta} - \frac{\alpha}{\beta}x \end{aligned}$$

$$\boxed{h^c = \frac{1}{\beta}\ln\left(\frac{\alpha a \delta_c}{\beta b}\right) + \frac{1}{\beta}\ln w - \frac{\alpha}{\beta}x}$$

Now for the constrained regime, the FOC is $u'\widetilde{w_c} = v'$. Because we have assumed the following form for the endogenous wage: $p\delta_c c k^{\gamma_k} e^{-\gamma_h h^c}$; the FOC can be expressed in the following way:

$$\begin{aligned} \overbrace{\alpha a e^{-\alpha x}}^{u'}\overbrace{p\delta_c k^{\gamma_k} e^{-\gamma_h h^c}}^{\tilde{w}_c} &= \overbrace{\beta b e^{\beta h^c}}^{v'} \\ \Leftrightarrow \frac{a\alpha p\delta_c c}{b\beta} e^{-\alpha x} e^{-\gamma_h h^c - \beta h^c} k^{\gamma_k} &= 1 \\ \Leftrightarrow \ln\frac{a\alpha p\delta_c c}{b\beta} - \alpha x - h^c(\gamma_h + \beta) + \gamma_k \ln k &= 0 \\ \Leftrightarrow \boxed{h^c = \frac{\ln\frac{a\alpha p\delta_c c}{b\beta}}{\gamma_h + \beta} - \frac{\alpha x}{\gamma_h + \beta} + \frac{\gamma_k \ln k}{\gamma_h + \beta} = 0} \end{aligned}$$

which is the reduced-form equation found in the theoretical section.

To sum up, there are two-reduced form that we will label as follows:

$$\begin{aligned} h_1^c &= \underbrace{\frac{1}{\beta}\ln(\frac{\alpha a \delta_c c}{\beta b})}_{A_1} - \underbrace{\frac{\alpha}{\beta}}_{B_1} x + \underbrace{\frac{1}{\beta}}_{C_1}\ln w \\ h_2^c &= \underbrace{\frac{\ln\frac{a\alpha p\delta_c c}{b\beta}}{\gamma_h + \beta}}_{A_2} - \underbrace{\frac{\alpha}{\gamma_h + \beta}}_{B_2} x + \underbrace{\frac{\gamma_k}{\gamma_h + \beta}}_{C_2}\ln k \end{aligned}$$

As explained in the theoretical part, we can recover α, β, γ_h, and γ_k from our estimation in the following order:

$$\begin{aligned}
\beta &= 1/C_1 \text{ since } C_1 = 1/\beta \\
\alpha &= -B_1 * \beta \text{ since } B_1 = -\alpha/\beta \\
\frac{a\delta_c}{b} &= e^{\beta A_1}\frac{\beta}{\alpha} \text{ since } A_1 = \frac{1}{\beta}\ln(\frac{\alpha a \delta_c}{\beta b})
\end{aligned}$$

Indeed we have that:

$$\begin{aligned}
A_1 &= \frac{1}{\beta}\ln(\frac{\alpha a\delta_c}{\beta b}) \Leftrightarrow \beta A_1 = \ln(\frac{\alpha a\delta_c}{\beta b}) \\
\Leftrightarrow \beta A_1 &= \ln\frac{\alpha}{\beta} + \ln\frac{a\delta_c}{b} \Leftrightarrow \beta A_1 - \ln\frac{\alpha}{\beta} = \ln\frac{a\delta_c}{b} \\
\frac{a\delta_c}{b} &= e^{\beta A_1 - \ln\frac{\alpha}{\beta}} = e^{\beta A_1}\frac{\beta}{\alpha}
\end{aligned}$$

Obviously, setting $a = 1$, we have:

$$\frac{\delta_c}{b} = e^{\beta A_1}\frac{\beta}{\alpha}$$

Note that there is no loss of generality because no parameters was set for any of the function in the utility. Therefore setting a to one does not change the preferences represented by the utility function. Finally, we show that we can identify the parameters for the endogenous child wage function (itself a function of the production function):

$$\gamma_h = \frac{-\alpha}{B_2} - \beta$$

Indeed we have that:

$$B_2 = -\frac{\alpha}{\gamma_h + \beta} \Leftrightarrow \gamma_h + \beta = -\frac{\alpha}{B_2} \Leftrightarrow \gamma_h = -\frac{\alpha}{B_2} - \beta$$

and

$$\gamma_k = C_2(\gamma_h + \beta) \text{ since } C_2 = \frac{\gamma_k}{\gamma_h + \beta}$$

7.2 Corner solutions for the case with market access

The previous results apply for interior solutions only. The first order conditions discussed allow for the definition of an exogenous wage which itself induces a zero relationship between land and child labor after controlling for x. Here we show in greater details what happens for corner solutions. To summarize, if an exogenous wage cannot be defined, land and child labor will be positively related even after controlling for x.

Corner solutions such that $h^c = 0$ will not be examined because their case is trivial. Land will not have an impact on child labor if child labor is fixed. So we now turn to corner solutions with $h^c > 0$ and only one of the "adult" variables (that is adult off farm labor t^a, adult on farm labor h^a or (adult) hired-in labor h^i) is equal to zero.

- if only $h^a = 0$, the maximization program becomes

$$\max_{h^i, h^c, t^a} u(y + w_a t^a + pg(\delta_i h^i + \delta_c h^c, k) - w_i h^i) - v_a(t^a) - v_c(h^c) \tag{18}$$

and the first order conditions are:

$$\frac{\partial \mathcal{P}}{\partial t^a} = 0 \Leftrightarrow u' w_a = v'_a \tag{19}$$

$$\frac{\partial \mathcal{P}}{\partial h^c} = 0 \Leftrightarrow u' p g_h \delta_c = v'_c \tag{20}$$

$$\frac{\partial \mathcal{P}}{\partial h^i} = 0 \Leftrightarrow u' p g_h \delta_i = u' w_i \Leftrightarrow p g_h \delta_i = w_i \tag{21}$$

Clearly an exogenous wage can still be defined. Combining the last two equations leads to $u' w \delta_c = v'_c$ as in the general case for an interior solution (one of the"adult" FOC in the general case is redundant). Obviously, the same m-demand is obtained and land does not impact child labor when controlling for x.

- If we look at what happens for $h_i = 0$ only, the program to maximize is:

$$\max_{h^a, h^c, t^a} u(y + w_a t^a + p g(\delta_a h^a + \delta_c h^c, k)) - v_a(h^a + t^a) - v_c(h^c) \tag{22}$$

and the FOC are:

$$\frac{\partial \mathcal{P}}{\partial t^a} = 0 \Leftrightarrow u' w_a = v'_a \Leftrightarrow \frac{v'_a}{u'} = w_a \tag{23}$$

$$\frac{\partial \mathcal{P}}{\partial h^a} = 0 \Leftrightarrow u' p g_h \delta_a = v'_a \Leftrightarrow \frac{v'_a}{u'} = p g_h \delta_a \tag{24}$$

$$\frac{\partial \mathcal{P}}{\partial h^c} = 0 \Leftrightarrow u' p g_h \delta_c = v'_c \tag{25}$$

$$\tag{26}$$

Combining the first two equations, we obtain $w_a = p g_h \delta_a$ and we can write $p g_h = w$. The first order conditions for child labor can be rewritten into $u' w \delta_c = v'_c$ which amounts to the general case.

- If $t_a = 0$ the program to maximize becomes:

$$\max_{h^i, h^a, h^c} u(y + pg(\delta_a h^a + \delta_i h^i + \delta_c h^c, k) - w_i h^i) - v_a(h^a) - v_c(h^c) \tag{27}$$

and the first order conditions are:

$$\frac{\partial \mathcal{P}}{\partial h^a} = 0 \Leftrightarrow u' p g_h \delta_a = v'_a \tag{28}$$

$$\frac{\partial \mathcal{P}}{\partial h^c} = 0 \Leftrightarrow u' p g_h \delta_c = v'_c \tag{29}$$

$$\frac{\partial \mathcal{P}}{\partial h^i} = 0 \Leftrightarrow u' p g_h \delta_i = u' w_i \Leftrightarrow p g_h \delta_i = w_i \tag{30}$$

Again the last equation allows to fix pg_h to some constant value w and we can rewrite the equation for child labor as $h^c = f(w, x)$.

All preceding cases led to the same m-demands. Now we look at what happens if two of the adult variables are equal to zero.

- If $h^a = t^a = 0$. The program to maximize becomes:

$$\max_{h^i, h^c} u(y + pg(\delta_i h^i + \delta_c h^c, k) - w_i h^i) - v_c(h^c) \tag{31}$$

The first order conditions are then:

$$\frac{\partial \mathcal{P}}{\partial h^i} = 0 \Leftrightarrow u' p g_h \delta_i = u' w_i \Leftrightarrow p g_h \delta_i = w_i \tag{32}$$

$$\frac{\partial \mathcal{P}}{\partial h^c} = 0 \Leftrightarrow u' p g_h \delta_c = v'_c \tag{33}$$

We obtain $pg_h \delta_i = w_i$ and $u' w \delta_c = v'_c$ with $w = pg_h$. That is, we are still able to have an exogenous wage. We can then write that $h^c = f(w, x)$ and so controlling for

x and w, land k will not impact child labor.

- When $h^i = t^a = 0$ the program and the FOC will look exactly as the ones for no market access (following section) and land quantity will influence the amount of child labor. Actually this corner solution corresponds exactly to a case with no market access as with a full market access this corner solution is not possible. Indeed, if $h^i = 0$ then $pg_h\delta_i < w_i$, that is, $pg_h < w$. If $t^a = 0$ then $pg_h\delta_a > w_a$, that is, $pg_h > w$. There is an obvious contradiction. Note that it would be possible to look at such a case by expliciting transaction costs. The household, or actually its farm's productivity (proxied by the farm size) would be exactly inside the price band (seeing transaction costs as marking the difference between the wage the household would have by working outside and the wage it would have to pay external workers to work on the farm) formed by these transaction costs. Being inside the price band, corresponds to not having access so that in the end it comes back to the same model.

- Now if $h^i = h^a = 0$, the program writes as:

$$\max_{h^c, t^a} u(y + w_a t^a + pg(\delta_c h^c, k)) - v_a(t^a) - v_c(h^c) \tag{34}$$

and the first order conditions are:

$$\frac{\partial \mathcal{P}}{\partial t^a} = 0 \Leftrightarrow u' w_a = v'_a \tag{35}$$

$$\frac{\partial \mathcal{P}}{\partial h^c} = 0 \Leftrightarrow u' pg_h \delta_c = v'_c \tag{36}$$

Clearly the case differs from the preceding ones as only children are working on the farm. $h^i = 0$ means that external workers are too expensive in the sense that paying external laborers brings less value than the value obtained from their work.

Mathematically, $pg_h\delta^i < w_i$ or, by definition of w_i,$pg_h < w$. In other words, the marginal productivity of an hour of work on the farm is lower than the hour wage that must be paid to outside workers. Obviously adult members of the household have no incentive to farm the plot either: $pg_h < w \Leftrightarrow pg_h\delta^a < w^a$ so that adults would be losing money by spending time on the farm. Regarding child labor, from the second first order condition, we can write $h^c = f(pg_h, x)$. In this case, as the amount of land increases, the marginal productivity of labor g_h increases and therefore h^c increases.

- Finally the case where all adult variables are zero, that is, no adult is working: $h^a = h^i = t^a = 0$. In that case, the household maximizes:

$$\max_{h^c} u(y + pg(\delta_c h^c, k)) - v_c(h^c) \tag{37}$$

and the first order condition is $pg_h\delta_c = v_c'$. There is clearly a positive relationship between land quantity and child labor. Assuming such households can fully access the market is not realistic. As mentioned when examining the case where $h^i = t^a = 0$, a contradiction occurs as we have both $pg_h < w$ and $pg_h > w$. This is not possible unless we assume transaction costs and that the household is inside the price band formed by the transaction costs. Assuming such a situation means that we are actually assuming that the household has no access to the market.

In the end, we have the intuition proposed in the introduction, namely that households with market access have a zero relationship between land quantity and child labor once we control for wage and expenditure. Only one corner solution violates the assumption of a zero relationship between land and child labor, namely: $h^i = h^a = 0$ but the case is almost never found empirically.

7.3 Distribution of ethnic groups and religions in Madagascar

Table 15: Distribution of ethnicities in Madagascar before aggregation

	Number	Percent
Antakarana	111,342	0.59
Antambahoaka	34,548	0.18
Antandroy	1,049,444	5.57
Antanosy	529,273	2.81
Antefasy	151,520	0.80
Antemoro	534,397	2.84
Antesaka	593,713	3.15
Arab	5,989	0.03
Bara	452,219	2.40
Betsileo	3,195,925	16.96
Betsimisaraka	2,479,850	13.16
Bezanozano	110,282	0.59
French	2,038	0.01
Karana	8,371	0.04
Comoros	33,466	0.18
Mahafaly	361,028	1.92
Merina	5,757,952	30.55
Sakalava	678,620	3.60
Sihanaka	630,348	3.34
Chinese	1,437	0.01
Tanala	647,301	3.43
Tsimihety	1,248,928	6.63
Vezo	87,900	0.47
Other Ethnicities	140,336	0.74
Missing	584	0.00
Total	18,846,812	100

Table 16: Distribution of ethnicities in Madagascar after aggregation

	Number	Percent
1: Other	1,552,492	8.24
2: South East	1,961,479	10.41
3: East	3,220,480	17.09
4: Highlands	5,757,952	30.55
5: Center	3,195,925	16.96
6: South	2,391,963	12.69
7: West	766,520	4.07
Total	18,846,812	100

Table 17: Distribution of religions in Madagascar before aggregation

	Number	Percent
Traditional	2,737,813	14.59
Catholics	6,766,713	36.07
Anglicans	448,526.48	2.39
FJKM Protestants	3,971,491	21.17
Lutherans	1,790,772	9.54
Adventists	153,813.41	0.82
Jesosy Mamonjy	144,107.66	0.77
Jehovah's Witnesses	82,754.53	0.44
Other Protestants	389,624.89	2.08
Muslims	132,496.75	0.71
Hindus	2,062.72	0.01
No religion	1,582,701	8.44
Atheists	301,206.99	1.61
Other	257,809.55	1.37
Total	18,761,892	100

Table 18: Distribution of religions in Madagascar after aggregation

	Number	Percent
Traditional	2,737,813	14.53
Catholics	6,766,713	35.90
Protestants	6,981,090	37.04
Other	2,361,196	12.53
Total	18,846,812	100

7.4 Exogeneity of household land on local wage

	ln(wage)
Land	-0.061
	(0.045)
Education	0.099***
	(0.011)
Female	-0.124***
	(0.021)
Rural area	-0.105***
	(0.023)
Age groups	
20-25	0.074*
	(0.039)
25-30	0.058
	(0.038)
30-35	0.115***
	(0.040)
35-40	0.103**
	(0.042)
40-45	0.192***
	(0.047)
45-50	0.169***
	(0.049)
50-55	0.160***
	(0.052)
55-60	0.125*
	(0.071)
Religions	
Catholic	0.143***
	(0.047)
Protestants	0.094*
	(0.048)
Other religions	0.005
	(0.054)
Ethnic groups	
Ethnic group 1	0.033
	(0.074)
Ethnic group 2	0.177***
	(0.066)
Ethnic group 3	0.036

	(0.062)
Ethnic group 5	0.097**
	(0.045)
Ethnic group 6	-0.094
	(0.091)
Regions	
Diana	0.289*
	(0.165)
Sava	0.683***
	(0.143)
Itasy	-0.293***
	(0.077)
Vakinankaratra	-0.549***
	(0.078)
Bongolava	0.020
	(0.076)
Sofia	-0.281*
	(0.143)
Boeny	0.075
	(0.121)
Betsiboka	0.283***
	(0.086)
Melaky	0.217*
	(0.120)
Alaotra-Mangoro	0.217**
	(0.102)
Atsinanana	-0.536***
	(0.099)
Analanjirofo	-0.183*
	(0.109)
Amoron'i Mania	-0.483***
	(0.084)
Mahatsiatra Ambony	-0.667***
	(0.094)
Vatovavy- Fitovinany	-0.699***
	(0.095)
Atsimo-Atsinanana	-0.797***
	(0.125)
Ihorombe	0.053
	(0.169)
Menabe	-0.218

	(0.133)
Atsimo-Andrefana	-0.567***
	(0.155)
Androy	0.303
	(0.196)
Anosy	-0.034
	(0.118)
Constant	7.060***
	(0.092)
Observations	2904

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported. Analamanga (region of the capital), the traditional religion, Merina and the 15-20 years old are the omitted categories of their respective group.

Regression of the wage in the primary sector for an individual on his/her amount of owned land]
[

7.5 Descriptive Statistics for the main variables

Table 20: Descriptive statistics for the variables that have been standardized, after outliers withdrawal

	N	Mean	sd	p25	Median	p75
child labor hours	4966	-0.003	0.982	-0.523	-0.523	0.192
land	4966	-0.068	0.548	-0.432	-0.173	0.080
wage, primary sector	4966	-0.011	0.993	-0.562	-0.006	0.682
expenditure	4966	-0.040	0.788	-0.548	-0.244	0.225
education, mother	4966	2.172	1.112	1	2	3
education, father	4966	2.322	1.191	1	2	3
age, years	4966	-0.009	0.994	-0.872	-0.010	0.636
rural area	4966	0.008	0.992	-1.135	0.789	0.789
nb. children 6-15	4966	0.005	1.001	-1.038	-0.216	0.605
household size	4966	0.003	1.001	-0.919	0.009	0.472
ratio of girls	4966	0.500	0.384	0	0.5	1
catholics	4966	0.322	0.467	0	0	1
protestants	4966	0.328	0.470	0	0	1
other religions	4966	0.175	0.380	0	0	0
ethnic group 1	4966	0.144	0.351	0	0	0
ethnic group 2	4966	0.122	0.328	0	0	0
ethnic group 3	4966	0.175	0.380	0	0	0
ethnic group 5	4966	0.168	0.374	0	0	0
ethnic group 6	4966	0.193	0.394	0	0	0
farm assets	4966	0.153	1.056	-0.876	0.581	0.581
female household head	4966	0.151	0.359	0	0	0
ethnic minority	4966	0.168	0.374	0	0	0
religious minority	4966	0.354	0.478	0	0	1
bank	4966	0.292	0.455	0	0	1
microfinance institution	4966	0.632	0.482	0	1	1
moneylender	4966	0.816	0.388	1	1	1
% salaried, primary sector	4966	0.150	0.207	0	0.043	0.245
% salaried, other sectors	4966	0.093	0.133	0	0.048	0.111
% rented plots	4966	0.045	0.080	0	0	0.063
% inherited plots	4966	0.001	0.998	-0.805	0.028	0.832