The Wealth Paradox for Whom? Child Labor and the Identification of Households Excluded from the Land and the Labor Markets in Madagascar

Samia Badji*

Abstract

The paper investigates child farm labor in Madagascar. To identify households having more child labor than what would prevail under full access to the land and the labor markets, the paper relies on the farm household model. The simple theoretical model highlights that land should not influence the number of hours of child work when the household can fully access the land or the labor markets. When access is limited, land may impact child work whereas the external wage should not. Using a switching regression model with unknown sample separation, this paper shows that not belonging to the largest ethnic or religious group at the local level significantly decreases market access.

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

JEL classification: D13, J13, J24, J43, O12, O13

*Univ Lyon, CNRS, GATE L-SE UMR 5824, F-69131 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 consists 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 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 their offspring will be more likely to attend school and the cycle goes on. In that perspective, 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 necessary. Nevertheless, the amount of worked hours 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) excluding the most vulnerable households. Improving the access is critical to avoid suboptimal outcomes that generally lead to more poverty and more child labor in particular (Bhalotra and Heady, 2003; Dumas, 2013). The literature has 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 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, they should work less as their wage increases. She finds a negative elasticity though only for boys. As the number of observations is limited in the study and because 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, such borrowing is threatened and only land-rich households can provide the collateral[1]. Casting doubt on the previous poverty explanations, the empirical literature has highlighted that some countries display the "wealth paradox": child labor is 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), 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 (positive) 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,

[1]Of course, in the ultimate case of no credit market, richer households would obviously be less likely to be in the subsistence situation described in the poverty hypothesis just before.

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. When used as collateral, more land leads to lower interest rates so that credit market imperfections induce less child labor. Another channel however is through credit use. If 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).

If households have more to farm than adult on-farm-labor supply can manage, they may need the labor market to hire workers. If the market fails, the household might rely on his/her own labor supply to face this extra work. A household with more land to farm than it can take care of, could sell the extra plots on the land market. If not properly functioning, the household might be unable to sell and rely on its own labor, including child labor. In both cases, the opportunity cost of not farming the land increases with land (as the marginal productivity of labor increases with land) and land will be positively correlated with child labor. Obviously, if the land market solely fails, plots that cannot be sold, can be farmed by external workers. Conversely, when the labor market alone fails, plots can simply be sold. However, when both markets fail, households with large plots have a stronger incentive to rely on child labor, inducing a positive relationship between land and child farm work, once the wealth effect is taken into account.

Of course, there is always an incentive to have children work on the farm given that they cannot work outside and that they are cheaper than external workers. The incentive to rely on family labor is clearly an increasing function of the local wage (when households have access to the labor market). Conditional on household wealth however, this incentive is unrelated to land quantity as the theoretical model will highlight[2].

[2]This holds true for farmers that are not directly influencing the local wage such as very large farm owners

Undoubtedly, as the amount of land increases, some households might reach a level of wealth for which they can afford not to farm the plots for which they could not find work (even though the opportunity cost is high). The decreasing marginal utility of consumption might reach a point where family leisure becomes more valuable, a point which might be reached particularly quickly if child leisure is a luxury good[3].

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 and therefore need labor at the same time (simultaneity issue). Without asymmetry of information, hired-in labor and family labor should both be remunerated at their marginal productivity. However farm's output does not solely depend on effort and as effort cannot be observed (moral hazard), people prefer family labor which 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 its market value (the present value of the future profits generated). The uncertainty regarding land ownership can be worsened by new land reforms leading former and new property systems to coexist and sometimes conflict. Besides land has collateral value that is only available to the owner. Land may have some value to the seller who might keep it for his children. Buyers might also think of land as being an investment for their children but the difference lies in the seller's potential attachment (the land may have belonged to his/her family for generations). The mismatch between buyers and sellers' value hampers the sales.

To circumvent these issues, the household might decide to rent the plot which is less prone

or farmers that would be the only (or almost) households with land in the village (in which case the fact that there is any market locally seems unrealistic unless that household has lots of land again, which is very rare in Madagascar and if to be investigated, should not be done together with the rest of the farmers).

[3]Of course as a household becomes really rich, it is difficult to think of the household as being constrained in its access to say, the labor market, in the sense that such a household would be in a position to hire by offering a more substantial wage than the market wage and attract, say, far away workers. Nevertheless such a strategy pays only as long as this higher wage is not greater than the marginal productivity of labor on the land. If the amount to attract far away workers is greater than the productivity, even richer households can be limited in their access to the labor market.

to market imperfections. Still, asymmetries of information on 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 help as it consists in sharing the output between the owner and the farmer. Yet, if the owner cannot observe the amount produced, and therefore the due production, the incentives for sharecropping will be reduced.

To identify adapted policies to improve land and labor market access, one must first identify the households that lack access to these two markets. The wealth paradox is an average effect on all households and 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 (or rationed) 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) in the sense that its own supply does not match its own demand.

This paper relies on the relationship between land and child labor to identify households constrained on the land and labor markets. Indeed, we cannot assess beforehand which households are constrained, we can only evaluate the probability to be constrained by looking at market imperfection measures. The model allows not only to identify households excluded from the formal markets but also from the informal markets. The labor-land relationship holds as long as one has no other alternatives, including alternatives on the informal sector. 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 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 ratio of girls decreases the probability to be constrained whereas it has a negative effect on the household average supply of child work suggesting that girls in the household might serve more as an adjustment variable through domestic work.

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 54,995 individuals belonging to 11,781 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-five-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[4]. In the following subsections, a description of the credit market and more particularly the land and the labor markets is undertaken.

2.1 Credit Market

The credit market in Madagascar is largely underdeveloped. According to the survey, 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. Among households who borrowed money in the 12 months preceding the interview, half inquired neighboring households, a quarter relatives and 6.6% a moneylender. Less than a fifth had recourse to the formal sector. The credit market seems to suffer from a lack of information with 95% of households not knowing 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 in the year before the interview by the 532 households, 40% were contracted in order to buy food and a quarter for professional activity including 13% for agriculture. This shows not only the high vulnerability of households in Madagascar but also that only a few loans are agriculture-related.

[4]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 recent migrations. Ethnic groups and religions had to be aggregated and their distribution is given in the appendix in tables 13 to 16.

2.2 Labor Market

In Madagascar, the labor market is quite limited despite a particularly high participation rates of 82% for women and 90% for men as described in table 1[5]. Most work consists in farm work especially for the 6 to 15 years old[6]. More than 97% of children in the agricultural sector work in their families which goes against the assumption of having a child farm wage. Not surprisingly, 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% which sinks to less than 5% for the primary sector which may indicate a poorly functioning labor market. A third of the Malagasy rely on a second activity where they are mostly independent (35%) or salaried (22% of those having a secondary activity) workers. For the latter case, 83% are in a temporary activity. More generally, 84% of the main activity is permanent 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.

All participation statistics are based on the seven days prior to the survey and one may worry that they do not reflect the reality as agricultural is typically seasonal. Using the EPM 2004 survey which allowed for the measure of seasonal employment, Stifel et al. (2007) found that it accounted for only 1% of total employment. This does not mean that there is no seasonality in the intensity of work and table 2 gives the number of hours worked in a typical week (not the last seven days) for participating individuals. Adults work on average 45 hours a week, 40 for adults in the agricultural sector. They spend almost the same amount of time in domestic chores. About one child out of five is working, 90% of which do in the agricultural sector indicating that child labor is clearly dominated by child farm labor. For

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

[6]14.5% of 6 to 10 years old work and 13.5% do farm work which means that 13.5/14.5=93% are in farm work.

participating children aged 11 to 15 (6-10), they work on average five (four) days a week and 6.5 (5.5) hours a day which seems more like a full time activity rather than simply helping the parents.

2.3 Land Market

The different uses of land are portrayed in table 3. 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 some of the market failures. 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. Less than 15 % of plots were bought and there were actually more plots acquired through clearing than buying[7]. Among the plots that were bought, table 5 indicates that half the plots bought from relatives (a fifth from a close relative and a third from a friend or a distant relation). This suggests potential issues of trust and information asymetries on the land market.

It is possible that inheritance schemes are adjusted so that the households are left with the amount of land they need. However table 6 suggests that adjusting land quantity is challenging. Most land is ancestral (57%), 12% of owned plots are actually state land, a fifth are bound by a contract (12% 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 while a fifth are under the name of a deceased person. On average, it takes five years and a half to title a plot, although it took less than two months for half of the plots suggesting strong

[7]The description is similar if we focus on households rather than plots (i.e. figures are not driven by large land owners).

ownership issues on the land market.

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 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 trade the quantity needed on the market. Dumas' model (2013) allows for four regimes depending on where the marginal productivity of labor of the household's farm is located with respect to two wages; the wage at which adult household members can be hired and the wage at which the household can hire labor. The setting presented here cannot be straightforwardly associated with the regimes in Dumas' (2013) paper as the price band is not made explicit: we are assuming that there is a continuum of wages. The continuum comes from a continuum in productivity rather than actual wages (payment at piece rate, widespread in agriculture, corresponds to this situation). In the model, 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 assume no land market which means that when a household cannot access the labor market, it will be unable to access both the land and the labor markets[9].

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

[9]As we are interested in the relationship between land and child labor, it is clear that when both are variables we cannot show their relationship (we could show how they relate when another parameter changes but our interest lies in the change in child hours when land varies, not another variable), one of them at least should be a parameter. Therefore we assume no land market since regardless of market imperfections child

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. Adults can choose between spending their time $\bar{E}_a$ 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 devote their time $\bar{E}_c$ either working on the family farm denoted as h^c or not work l^c[10]. We assume the subutilities are additively separable[11] and the utility function to maximize is the following:

$$u(x) - v_a(h^a + t^a) - v_c(h^c) \tag{1}$$

with $v_c' > 0$, $v_a' > 0$,$v_a'' > 0$, $v_c'' > 0$, $u' > 0$, $u'' < 0$, that is, the marginal disutility of work increases with the amount of work and the marginal utility of consumption decreases with consumption. 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. All types of labor are perfect substitutes and 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 at its marginal productivity for an interior solution. This assumption is not strong as 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 marginal productivity of labor (g_i denotes the marginal productivity of factor i and g_{ij} the second or cross derivative for $i \neq j$) which decreases with labor and increases with land:

labor is always an option (a variable).

[10] Assuming $\bar{E}_c = \bar{E}_a$ does not change the results.

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

$$q = g(\delta_a h^a + \delta_c h^c + \delta_i h^i; k) \text{ with } g_k > 0, g_h > 0, g_{hh} < 0, g_{hk} > 0 \tag{2}$$

Fixing the price of output at p, denoting y exogenous transfers, and w_a the off-farm family wage, the budget constraint writes:

$$y + w_a \bar{E}_a + pq = x + w_i h^i + w_a h^a + w_a l^a \tag{3}$$

x can be seen as expenditures on goods or the amount of goods purchased if we normalize the respective price to one. Total wealth, that is, exogenous transfers, the value of adult time[12] 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 (1) to (3), the function to maximize becomes:

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

The first order conditions are: $u' w_a = v'_a$; $u' p g_h \delta_a = v'_a$; $u' p g_h \delta_c = v'_c$ and $p g_h \delta_i = w_i$.

With $w_c \equiv \delta_c w$, the child labor function can be rewritten as

$$u'/1 = \quad v'_c/w_c \Leftrightarrow \tag{4}$$

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 hired-in work, nor do we observe profit. We can however derive the m-demand (Browning, 1999) of child labor from the first order condition.

[12]There is no value for child time as they cannot work outside the farm. They may indirectly increase wealth through farm output, freeing up adult time or hiring less.

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[13]:

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

The partial derivatives are:

$$\frac{\partial h^c}{\partial x} \equiv h_x^c = \frac{u''}{v''}\delta_c w < 0, \qquad \frac{\partial h^c}{\partial w} \equiv h_w^c = \frac{u'}{v''}\delta_c > 0. \tag{5}$$

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. Clearly, 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.

The previous results hold for an interior solution. 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[14] 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 fixed, then it will not vary with respect to land[15].

[13] δ_c is already included in the function f

[14] Adult labor refers to either adult off-farm labor t^a, adult on-farm labor h^a or (adult) hired-in labor h^i.

[15] In the production function, child labor is necessarily positive. However, we do not ignore than land may

If all adult variables are zero or if $t^a = h^i = 0$, then no adult is 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. Since $h^i = 0$, then $pg_h < w$, so that if there is market access, we must 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. Adults have no incentive to work their own farm as they can earn more by working off the farm. 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, less than 2% of households with land would not work on their farm while their children do ($h^c > 0$ and $h^a = 0$). Therefore the case for which in addition $h^i = 0$ is even more rare[16]. 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

In this situation, the household cannot rely on the market to satisfy its needs regarding work: $h^i = \bar{h^i}$ and $t^a = \bar{t^a}$ (note that $h^i = t^a = 0$ is a particular case of the model). The function

have an impact on the extensive margin. At the extensive margin, only the substitution effect may appear which, as for the intensive margin should only appear in case of limited access to the land and the labor markets.

[16] As mentionned information on hired-in labor is lacking so that the corresponding percentage cannot be given.

to maximize becomes:

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

and the FOCs are $u'pg_h\delta_a = v'_a$ and $u'pg_h\delta_c = v'_c$

Here $pg_h = w$ no longer holds and 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:

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

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 impacts child labor quantity through the shadow wage.

$$\frac{\partial h^c}{\partial k} \equiv h^c_k = f_{\widetilde{w}}\frac{\partial \widetilde{w}}{\partial k} = f_{\widetilde{w}}p\delta_c(g_{hh}\delta_c h^c_k + g_{hk}) \Leftrightarrow h^c_k = \frac{f_{\widetilde{w}}\delta_c g_{hk}p}{1 - f_{\widetilde{w}}pg_{hh}\delta_c^2} > 0 \tag{6}$$

$h^c_k > 0$[17] indicates that when the household is constrained by the land and the labor markets[18], 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

[17]The denominator is positive since all terms are positive except for g_{hh}.

[18]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.

to the market, land quantity will not affect child labor and the only corner solution left is such that $h^a = 0$ and $h^c > 0$. The m-demand then writes as $h^c = f(pg_h, x)$ and 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 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, equation (4) translates into $u'w\delta_c = v'$ and becomes: $\alpha a e^{-\alpha x} w\delta_c = \beta b e^{\beta h^c}$. In the regression we will estimate an additive form so that we take the logarithm of the previous equation and obtain for h_1^c (which stands for h^c in regime 1, i.e. with market access) :

$$h_1^c = \underbrace{\frac{1}{\beta}\ln(\frac{\alpha a \delta_c}{\beta b})}_{A_1} - \underbrace{\frac{\alpha}{\beta}}_{B_1} x + \underbrace{\frac{1}{\beta}}_{C_1} \ln w \tag{7}$$

A_1, B_1 and C_1 are the coefficients from the regression. Setting $a = 1$, we can recover all

parameters for regime 1 except, unsurprisingly, child productivity.

When the household is constrained (regime 2) we have:

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

Assuming that the production function is of the following form :

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

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 and increases with land. Then $\tilde{w}_c = p\delta_c c k^{\gamma_k} e^{-\gamma_h h^c}$ and we obtain h_2^c (which stands for h^c in regime 2, the regime with limited/no access):

$$h_2^c = \underbrace{\frac{1}{\gamma_h + \beta} \ln \frac{a\alpha p \delta_c c}{b\beta}}_{A_2} - \underbrace{\frac{\alpha}{\gamma_h + \beta}}_{B_2} x + \underbrace{\frac{\gamma_k}{\gamma_h + \beta}}_{C_2} \ln k \tag{8}$$

Since α and β are already identified, we get γ_h from B_2 and then γ_k from C_2. Setting $a = 1$ 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. Both h_1^c and h_2^c 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. This strategy is adopted as it is not possible to identify constrained households ex ante from the data. There are measures for market imperfection and we can see whether

households trade nevertheless it does not mean that non-trading households did not need the market nor does it imply that trading households fully met their needs. A switching regression model with unknown regime separation will help us in the assessment of the two regimes. One must simultaneously estimate the following system:

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

The first equation determines regime membership and the two others depict child labor. h_i^{c*} (observed when positive) is the average number of hours worked for all children of the household in regime i and C^* (unobserved) can be interpreted as the amount of constraints. The higher C^*, the higher the probability that the household will be constrained[19]. 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[20].

Regime membership is a function of different variables including market imperfection measures and household head characteristics. Some of these variables are common to all equations and are represented by vector $\mathbf{Z}$. $\mathbf{Z}$ includes ethnicity, religion and sex of the household head, parent's education, family composition, total expenditure and a dummy for living in a rural area. ln w is the logarithm of the wage (included in the switching equation and the first regime) and ln k is the logarithm of owned land (included in the switching equation and the second regime). $\mathbf{G}$ are all the variables that are likely to influence regime membership but not the quantity of hours worked. $\mathbf{G}$ includes a dummy indicating whether the household is part of the local ethnic minority at the community level, a similar dummy for religion, the percentage of titled plots at the community level, the percentage of plots

[19]By construction whenever $C^* \geq \bar{C}$ (set to zero without loss of generality), people are assigned to the constrained regime 2, otherwise to regime 1)

[20]Setting $\sigma = 1$ has no implication in the structural model as it only appears in the switch. For the switch we mainly want to see which variables significantly predict the regime and their sign therefore $\sigma = 1$ does not affect the interpretation.

registered on a cadaster at the community level, the rate of salaried workers in the primary sector and the rate of salaried workers in the non primary sector. Access to credit is also included through the availability (dummy variables) of a bank (formal) or a moneylender (informal sector).

Variables included in the switch that do not directly impact child labor help the identification of the regimes. Credit infrastructure unavailability might lead to exclusion and through market exclusion influence the actual number of hours but a *direct* impact is doubtful. Indeed the unavailability of a bank can lead to constraints regarding the types of input that can be bought but once we are in the constrained (or non-constrained) regime, this should have no impact. Similarly, the proportion of salaried workers might impact the level of constraints C^* yet not the quantity of child work. It might be argued that it could have an impact for people with access through more or less pressure on the wages but since we do control for the local wage in the regime with market access, the proportion of salaried workers should no longer influence child work quantity. Minority dummy variables might induce a form of discrimination and therefore a form of constraint but they will not explain the amount of hours worked (especially since ethnicity is already controlled for), that is why they are excluded from both regimes. They are not *per se* exclusion variables as if there is no discrimination, they will not be significantly different from zero which is what we are investigating. At least some exclusion variables should be non zero to improve the identification which is what the results show in the next section.

4.2 The regimes

The exclusion variables for the regimes are land and wage. Wage is defined as the wage prevalent where the household lives or local wage (at locality level). 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 worked 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. The probability to observe the wage is likely not random and the Heckman procedure is implemented. We use land to help identification as it should influence the probability to observe the wage (or the decision to work outside the farm) but not the wage itself. Indeed, 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[21]. 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[22].

Table 7 displays the estimation for the wage with the Heckman procedure. Regional dummies were introduced to account for 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 productivity 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

[21]Local wage is the average income received by all workers in the same locality excluding workers from the household.

[22]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.

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 earn significantly more than the 16 to 20 year old.

The theoretical model showed that for an individual i the wage is $w_i = \delta_i p g_h$. Hence the regional dummies and the area (rural or urban) can be interpreted as proxies for the price p of the output. g_h, the marginal productivity of labor, is independent of 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 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.

As shown in the descriptive statistics, one child out of five is working. That is, we do not observe hours of child 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 and the dependent variable becomes the average number of hours worked by children (aged 6 to 15) per household. As we classify households and not children,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[23]. 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

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

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[24].

When h^c equals zero, it is the result of a corner solution for the latent variable "h^{c*}, average number of hours wished by each household". For this reason, we use a Tobit model and regress h^c on $\mathbf{Z}$. The variables in Z were already enumerated and here we justify their use. Household characteristics such as the parents' 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). Since child labor might be gender specific and since information was aggregated at the household level, a variable indicating the proportion of girls among the 6-15 year old children is added. An 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 it might be endogenous as child labor may indrectly bring revenue. It may avoid hiring external labor or may free up adults' time allowing them to bring more money (which may increase expenditure) and richer households might be more able to forgo child labor. Consequently expenditure is instrumented with permanent wealth (using an asset index[25]). Table 8 first column displays the results of the regression of hours of child work before the instrumentation.

[24]Total hours of work could be used but would make the interpretation on household composition more difficult as it would be directly linked the number of children which we control for.

[25]See Vyas and Kumaranayake (2006), Filmer and Pritchett (2001) and Montgomery et al. (2000). 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.) and housing information such aswalls / floor/ ceiling material, source of lighting, type of toilet, whether there is electricity, the source for the water, how the liter is getting rid of, square meters are included. No productive asset was included (tractors, plows, carts, harrows, etc.)

The second column shows the results after instrumentation and the last column shows the first stage. 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[26]. 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 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, have 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. With land quantity classified in three groups we do observe the wealth paradox (coefficient of land is positive) as displayed in table 9. The logarithmic transformation might not completely capture the non linear effect of land (or a different one) and the following section provides consistent results with different measures of land.

5 Results

Table 10 provides the results for the estimation of the model presented in the previous section. In panel A, we draw on Hartley's work (1978) to use the EM or Expectation Maximization algorithm[27] as the likelihood function to estimate is a sum of probability distributions (weighted by the probability to fall into each regime) and the log of a sum is notoriously known to be troublesome to maximize. Panel B (discussed afterwards) uses the regime division in Panel A to reestimate the two regimes.

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

[27]The EM algorithm first introduced by Hartley (1958) and popularized by Dempster, Laird and Rubin (1977) consists in maximizing the expectation would we observe censored data, conditional on what we actually observe. The parameters obtained from the maximization are then introduced into the expectation which is again maximized and the process continues until convergence is achieved. 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. 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) applies the algorithm to a mixture of two censored normal distributions.

Panel A seems to show 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 a higher probability 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. 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[28]. In that sense having more individuals is more constraining.

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. Higher proportion of salaried workers is an indication of the existence of a labor market and therefore it is easier to find workforce. The proportion of salaried workers in the other sectors increases the likelihood to be constrained[29]. 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

[28]In the sense that more "able" members are much more likely to leave the household and start their own household. 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.

[29]If a person is neither salaried in the primary sector, nor in other sectors then s/he can be not working or not working but not salaried.

on a cadaster at the local level increases market access. 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 access the markets.

Looking at the regimes, we find that wealth has a negative impact on child labor. Controlling for wealth, land increases child hours of work when people lack access to the markets consistent with the theoretical model[30] 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. Unsuprinsingly, in both regimes, working hours increase as children age. 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 enterprise does not seem to have an effect on child working hours regardless of whether the household has access to the markets but it significantly reduces the probability to be constrained though it could be endogenous[31].

Table 10 panel B shows the estimation taking into account the results from the switch. That is, we use households for which $C^* < 0$ to estimate the tobit equation of regime 1 and the rest of the sample for the tobit equation of regime 2 (as opposed to using the whole population with weights as done in panel A). In addition to assessing the robustness of the results, the reestimation 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[32].

One of the main difference is also in the instrumentation as it was undertaken on the whole

[30]By construction, land has no impact in regime 1, the point however is that land could also have had no impact when access is limited. It was even more possible as with the logarithmic specification, the pooled tobit showed no impact of land.

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

[32]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 impact on child amount of work. Wage was added when there is a limited access and it did not have a significant impact either. As it is not a strong empirical test for the absence of impact we should highlight that it is a result of the theoretical model which was based on a limited number of assumptions. Moreover, 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. For that reason it is necessary to have restrictions to identify the two regimes.

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. 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[33]. 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 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. Reinstrumenting and estimating each regime again leads to similar results[34].

Table 11 shows the results of the switching regimes with the land variable in three categories, the first category for households with little land (the omitted category in the estimation). Results are practically the same and the regimes seem to be identified (land is positive in the first regime while wage is positive for households with access). For this reason, we focus on differences to see what the (different) 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. Land may come 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. Its effect on hours of work in the constrained regime is negative and strongly significant. 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) allowing for task specialization: girls would not work at all

[33]As we cannot know ahead the regime to which households belong, we are forced to instrument 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.

[34]Results are available upon request.

on the farm and their mother might go and spend more time on the farm. Results for the switching estimations with land and land squared or land in level are very similar[35] as shown in the appendix table 18 and 17 respectively. The effect of expenditure is again stronger (in absolute value) in the regime with market access suggesting an increasing marginal productivity of labor (it becomes clear when examining the equations for the parametrization hereafter) only consistent with small amount of land.

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 (8). Table 12 recalls the notations used and values obtained. From the first regime, we obtain 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 decreasing with consumption. In the same way, more work leads to a higher disutility 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 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). Consistent 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

[35] In level, the production function writes $g(h,k) = \frac{ce^{\gamma_k k}}{\gamma_h}(1 - e^{-\gamma_h h})$ and the shadow wage $\tilde{w}^c = p\delta ce^{-\gamma_h h^c + \gamma_k k}$). This is not realistic since no land produces a strictly positive output. Since only households with land are examined such a case would never occur and it is still worth investigating what happens when land is in level.

for mother education, household composition or even religion are systematically stronger when households lack access to the markets[36]. It can be that any effect is more visible when households lack access which is consistent with the first order conditions where preferences have a more preponderant role when access is limited.

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 if they had access to the markets even when they do not display the wealth paradox. To the extent that this effect only appears when households need access to the markets, it can be seen as a lower bound. Numerically dominated ethnic and 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 needs 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, arrangements 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 adjustments 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

[36]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.

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[37]). 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 and the paper has identified the population of greater concern.

[37]Their cost-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. *Working Papers in Economics*, (256), 2007.

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

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.

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.

H. G. Jacoby and B. Minten. Is land titling in Sub-Saharan Africa cost-effective? Evidence from Madagascar. *The World Bank Economic Review*, 21(3):461–485, 2007.

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

M. R. Montgomery, M. Gragnolati, K. A. Burke, and E. Paredes. Measuring living standards with proxy variables. *Demography*, 37(2):155–174, May 2000.

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.

S. Vyas and L. Kumaranayake. Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning*, 21(6):459 –468, Nov. 2006.

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.

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.

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 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	

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

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)
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. Dummies for region are included. The traditional religion, Merina and the 15-20 years old are the omitted categories of their respective group.

Table 8: Pooled estimation 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 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.

Table 10: Switching regression estimation, land in logarithm

		Panel A		Panel B	
	switch	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*	.	0.449***		0.480**
	(0.066)	.	(0.169)		(0.220)
expenditures	-0.346	-1.017***	-1.953***	-1.362***	-0.896**
	(0.333)	(0.135)	(0.610)	(0.140)	(0.409)
wage, primary	0.267***	0.212***	.	0.293***	
	(0.087)	(0.034)	.	(0.036)	
productive asset	-0.042		-0.090		-0.099
	(0.061)	.	(0.148)		(0.094)
average age	0.105	0.467***	1.296***	0.575***	1.181***
	(0.075)	(0.039)	(0.209)	(0.041)	(0.093)
education, mother	-0.107	-0.117***	-1.655 ***	-0.120***	-1.126***
	(0.091)	(0.040)	(0.327)	(0.039)	(0.158)
education, father	-0.068	-0.024	-0.556 ***	-0.036	-0.511***
	(0.079)	(0.037)	(0.200)	(0.036)	(0.133)
nb. children 6-15	-1.301***	0.000	0.055	-0.025	-0.065
	(0.160)	(0.043)	(0.294)	(0.045)	(0.201)
nb. individuals	0.300**	0.203***	0.127	0.288***	-0.056
	(0.122)	(0.049)	(0.225)	(0.050)	(0.161)
Catholics	0.015	-0.236**	-1.083**	-0.354***	-1.084***
	(0.212)	(0.097)	(0.522)	(0.100)	(0.341)
Protestants	-0.097	-0.319***	0.222	-0.422***	-0.176
	(0.212)	(0.093)	(0.475)	(0.099)	(0.325)
other religions	0.303	-0.150	0.265	-0.263***	0.234
	(0.210)	(0.104)	(0.429)	(0.101)	(0.292)
ethnic group 1	0.560*	-0.401***	0.482	-0.447***	0.205
	(0.287)	(0.116)	(0.567)	(0.120)	(0.419)
ethnic group 2	0.251	-0.315***	-0.901	-0.359***	-0.702
	(0.270)	(0.106)	(0.651)	(0.110)	(0.451)
ethnic group 3	0.700 ***	-0.588***	-1.624**	-0.775***	-1.173***
	(0.251)	(0.111)	(0.638)	(0.108)	(0.422)
ethnic group 5	0.871***	-0.215*	0.594	-0.236**	0.498
	(0.253)	(0.113)	(0.565)	(0.111)	(0.386)
ethnic group 6	0.314	0.164*	0.054	0.235**	0.012
	(0.253)	(0.097)	(0.601)	(0.100)	(0.423)
household head female	-0.446**	-0.016	-0.776	-0.019	-0.555*
	(0.193)	(0.081)	(0.473)	(0.086)	(0.304)
rural area	0.204***	-0.072**	0.421**	-0.109***	0.294***

	(0.077)	(0.033)	(0.177)	(0.033)	(0.111)
ratio of girls	-0.273	0.106	-1.343***	0.157*	-1.117***
	(0.175)	(0.092)	(0.407)	(0.095)	(0.235)
non farm entreprise	-1.086 ***	0.043	-0.676	0.094	-0.401
	(0.205)	(0.068)	(0.636)	(0.071)	(0.354)
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

* 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.

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 titled or registered on a cadaster. Productive asset is an index of productive assets at the household level, animals are excluded. Regime 1 is the regime with access to the land and labor markets. Regime 2 corresponds to the regime with limited or inexistant access.

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

7 Appendix

Table 13: 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 14: 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 15: 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 16: 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

Table 17: Switching regression model output with land in level

	Panel A: switching regression			Panel B: Post estimations		
	regime 1	regime 2	switch	regime 1	regime 2	switch
constant	0.249*	0.825	-0.249	0.267 *	0.724	-0.321
	(0.143)	(0.530)	(0.473)	(0.137)	(0.589)	(0.427)
land	.	0.886***	-0.177		0.837 ***	-0.174
	.	(0.224)	(0.157)		(0.243)	(0.145)
local wage (primary)	0.210***	.	0.349***	0.208 ***		0.343 ***
	(0.033)	.	(0.111)	(0.031)		(0.103)
expenditures	-1.000***	-1.100***	-0.375	-1013 ***	-1079 **	-0.321
	(0.128)	(0.381)	(0.429)	(0.126)	(0.428)	(0.389)
age	0.460***	1.027***	0.254**	0.458 ***	1061 ***	0.270 ***
	(0.041)	(0.130)	(0.098)	(0.038)	(0.146)	(0.094)
education, mother	-0.155***	-1.034***	-0.027	-0.114 ***	-1054 ***	-0.076
	(0.038)	(0.170)	(0.114)	(0.037)	(0.193)	(0.102)
education, father	-0.079**	-0.479***	-0.138	-0.028	-0.490 ***	-0.187 *
	(0.035)	(0.128)	(0.107)	(0.034)	(0.146)	(0.097)
nb. children 6-15	0.005	0.169	-1.517***	0.027	0.165	-1575 ***
	(0.043)	(0.209)	(0.226)	(0.040)	(0.231)	(0.208)
household size	0.187***	-0.016	0.362**	0.206 ***	-0.053	0.354 **
	(0.047)	(0.158)	(0.156)	(0.045)	(0.173)	(0.142)
Catholics	-0.236**	-0.911**	-0.075	-0.236 ***	-0.938 **	-0.097
	(0.094)	(0.360)	(0.259)	(0.089)	(0.393)	(0.241)
Protestants	-0.302***	-0.030	-0.167	-0.300 ***	-0.021	-0.188
	(0.091)	(0.327)	(0.257)	(0.086)	(0.357)	(0.237)
other religions	-0.184*	0.164	0.199	-0.187 **	0.194	0.225
	(0.099)	(0.301)	(0.264)	(0.094)	(0.325)	(0.240)
ethnic group 1	-0.367***	0.247	0.709**	-0.368 ***	0.327	0.778 **
	(0.110)	(0.405)	(0.359)	(0.106)	(0.453)	(0.335)
ethnic group 2	-0.295***	-0.557	0.482	-0.312 ***	-0.490	0.484
	(0.104)	(0.445)	(0.332)	(0.099)	(0.500)	(0.309)
ethnic group 3	-0.604***	-1.127***	0.801**	-0.614 ***	-1072 **	0.862 ***
	(0.105)	(0.423)	(0.316)	(0.099)	(0.481)	(0.297)
ethnic group 5	-0.250**	0.449	1.154***	-0.251 **	0.554	1272 ***
	(0.110)	(0.394)	(0.329)	(0.107)	(0.450)	(0.310)
ethnic group 6	0.180*	0.046	0.478	0.175 **	0.125	0.498 *
	(0.093)	(0.417)	(0.313)	(0.088)	(0.471)	(0.293)
female headed hh.	0.005	-0.514*	-0.506**	0.002	-0.546	-0.552 **
	(0.078)	(0.312)	(0.247)	(0.075)	(0.344)	(0.229)
rural	-0.075**	0.280**	0.226**	-0.072 **	0.289 **	0.226 **
	(0.032)	(0.121)	(0.098)	(0.030)	(0.132)	(0.090)
ratio of girls	0.071	-0.914***	-0.376*	0.083	-0.947 ***	-0.395 *
	(0.091)	(0.258)	(0.227)	(0.085)	(0.285)	(0.206)
non farm. enterprise	0.065	-0.206	-1.120***	0.075	-0.277	-1225 ***
	(0.065)	(0.360)	(0.251)	(0.063)	(0.429)	(0.246)
productive assets	.	-0.081	-0.118	.	-0.090	-0.128 *
	.	(0.098)	(0.084)	.	(0.109)	(0.076)
ethnic minority			0.540**			0.548 ***
			(0.245)			(0.213)
religion minority			0.357**			0.371 **
			(0.171)			(0.153)
Bank present			0.093			0.074
			(0.181)			(0.164)
moneylender present			-0.270			-0.261
			(0.208)			(0.189)
Proportion in the village of						
salaried, primary			-1.613***			-1.672 ***
			(0.504)			(0.468)
salaried, other sectors			5.044***			5.286 ***
			(1.123)			(0.986)
registered plots			-2.407***			-2.493 ***
			(0.486)			(0.438)
inherited plots			0.229**			0.231 ***
			(0.096)			(0.087)
sigma	1.058***	2.986***		1.048 ***	3.047 ***	
	(0.027)	(0.224)		(0.024)	(0.251)	

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported. The traditional religion and Merina are respectively the religion and ethnic 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 18: Switching regression model output with land and land square

	Panel A					Panel B		
	regime 1		regime 2		switch	regime 1	regime 2	pooled
constant	0.335**		1.254		-0.610	0.223	-0.183	-1.385***
	(0.155)		(0.921)		(0.371)	(0.220)	(0.453)	(0.360)
land			1.323***		-0.136		0.429*	0.319***
			(0.492)		(0.156)		(0.227)	(0.099)
land, squared			-0.837	*	0.020		-0.335*	-0.215***
			(0.442)		(0.092)		(0.192)	(0.063)
local wage (primary)	0.206***		0.000		0.273 ***	0.248***		0.148***
	(0.035)				(0.087)	(0.049)		(0.047)
expenditures	-0.972***		-1.848	***	-0.322	-1.720***	-1.130***	-1.467***
	(0.140)		(0.670)		(0.319)	(0.187)	(0.328)	(0.169)
age	0.479***		1.396	***	0.102	0.630***	0.845***	0.576***
	(0.042)		(0.235)		(0.075)	(0.060)	(0.076)	(0.040)
education, mother	-0.158***		-1.981	***	-0.062	-0.140**	-1.010***	-0.270***
	(0.042)		(0.382)		(0.092)	(0.059)	(0.140)	(0.049)
education, father	-0.070*		-0.600	***	-0.005	-0.076	-0.333***	-0.083*
	(0.039)		(0.214)		(0.078)	(0.054)	(0.105)	(0.045)
nb. children 6-15	-0.032		-0.026		-1.303 ***	0.007	-0.397**	0.464***
	(0.045)		(0.328)	***	(0.160)	(0.068)	(0.174)	(0.052)
household size	0.194***		0.004		0.286 **	0.239***	0.040	0.207***
	(0.050)		(0.251)		(0.121)	(0.067)	(0.133)	(0.062)
Catholics	-0.255**		-1.113	*	0.026	-0.596***	-0.611**	-0.408***
	(0.101)		(0.579)		(0.213) **	(0.151)	(0.276)	(0.124)
Protestants	-0.325***		0.378		-0.098	-0.507***	0.048	-0.257**
	(0.097)		(0.528)		(0.212)	(0.150)	(0.268)	(0.122)
other religions	-0.161		0.377		0.257	-0.233	0.308	-0.212*
	(0.107)		(0.481)		(0.211)	(0.154)	(0.238)	(0.120)
ethnic group 1	-0.381	***	0.626		0.674 **	-0.551***	0.543	-0.294*
	(0.120)		(0.628)		(0.290)	(0.181)	(0.353)	(0.152)
ethnic group 2	-0.280	**	-0.979		0.361	-0.515***	-0.463	-0.547***
	(0.113)		(0.713)		(0.265)	(0.169)	(0.379)	(0.150)
ethnic group 3	-0.597	***	-1.647	**	0.725 ***	-0.904***	-0.677*	-0.843***
	(0.115)		(0.699)		(0.252)	(0.164)	(0.347)	(0.152)
ethnic group 5	-0.198	*	0.812		0.927 ***	-0.170	0.627*	-0.232*
	(0.117)		(0.633)	***	(0.253)	(0.167)	(0.328)	(0.139)
ethnic group 6	0.176	*	0.059		0.426 *	0.224	0.203	-0.043
	(0.101)		(0.672)		(0.259)	(0.154)	(0.355)	(0.138)
female headed hh.	-0.031		-0.857	*	-0.466 **	-0.111	-0.554**	0.057
	(0.084)		(0.518)		(0.196)	(0.131)	(0.255)	(0.107)
rural	-0.072	**	0.465	**	0.176 **	-0.134***	0.221**	-0.067*
	(0.035)		(0.196)		(0.077)	(0.050)	(0.089)	(0.040)
ratio of girls	0.060		-1.553	***	-0.320 *	0.049	-0.842***	-0.067
	(0.098)	***	(0.459)		(0.176)	(0.142)	(0.187)	(0.100)
non farm. enterprise	0.027		-1.022		-1.057 ***	0.112	-0.856**	0.576***
	(0.072)	***	(0.721)	***	(0.201)	(0.105)	(0.345)	(0.094)
productive assets	0.000		-0.095		-0.025		-0.080	-0.049
			(0.165)		(0.061)		(0.078)	(0.037)
ethnic minority					0.394 **			
					(0.172)			
religion minority					0.260 **			
					(0.122)			
Bank present					0.117			
					(0.138)			
moneylender present					-0.214			
					(0.155)			
Proportion in the village of								
salaried, primary					-1.163***			
					(0.356)			
salaried, other sectors					4.584***			
					(0.762)			
registered plots					-2.074***			
					(0.337)			
inherited plots					0.192***			
					(0.073)			
sigma	1.116***		3.836***					2.045***
	(0.028)		(0.422)					(0.041)
N	4966		4966			2717	2249	4966

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported. The traditional religion and Merina are respectively the religion and ethnic 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 19: Regression of the wage in the primary sector for an individual on his/her amount of owned land

	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)
Observations	2904	

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are reported in parentheses. Region dummies are included. The traditional religion, Merina and the 15-20 years old are the omitted categories of their respective group.

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