Parental Income and School Attendance in a Low-Income Country: a semi-parametric analysis[*]

Denis Cogneau (DIAL-IRD)   Eric Maurin (CREST-INSEE)

January 2001

Abstract

Using data covering three successive generations of people from Madagascar, we construct a new semi-parametric estimator of the effect of parental income on the decision to send children to school. We propose new tests for detecting the simultaneity and hereditary biases that affect the usual estimates of this effect. We demonstrate the importance of the first type of bias. Our results show that the existing literature underestimates the real effect of family resources on decisions as to whether or not to send children to school by not paying enough attention to the simultaneity of income formation processes and schooling decisions.

1. Introduction

Despite being historically and culturally quite school-orientated, Madagascar's level of schooling is currently one of the lowest in the world. One Malagasy child in five never goes to school[1]. The majority start school one to two years later than the normal age. Many children then leave school early and rare are those who attend secondary school. Madagascar is not an isolated case. The same phenomenon can be seen in most of the sub-Saharan African and South Asian countries[2]. Children in poor countries today generally still receive little

[1]See Cogneau et al. (2000).

[2]See, for example, Filmer and Pritchett (1999).

or no schooling with the result that one of the basic conditions for economic development is left unfulfilled.

Obviously, it is very important to understand the underlying reasons for the extremely low level of schooling among children in low-income countries. From a strictly economic point of view, the reasons can be found in both the lack of school infrastructures (insufficient supply) and the lack of family resources (insufficient demand). When parents have no or virtually no means of borrowing to finance their children's studies, the lower the family income the lower the level of the children's schooling. This paper concentrates mainly on this second aspect of the problem. To what extent does the lack of household resources explain the extremely low level of schooling in low-income countries? How sensitive is a family's educational demand to its income? We have a wealth of Malagasy data at our disposal to identify the real impact of parental income on schooling decisions. These data show the rates of school attendance across three successive generations (children, parents and grandparents). Based on these data, we make a two-part contribution:

(a) In terms of method, we construct a new semi-parametric estimator of the effect of parental income on the decision as to whether to send children to school. We use the new estimator for qualitative response model with endogenous regressors recently introduced by Lewbel (2000). We propose new tests of simultaneity biases and hereditary biases affecting the usual estimates of this income effect.

(b) In terms of findings, we reveal the importance of the simultaneity biases affecting usual estimates of the income effect on schooling decisions. We find that the existing literature greatly underestimates the real effect of family resources on schooling decisions by not paying enough attention to the simultaneity of the income formation processes and schooling decisions. In the Malagasy case, our analysis provides evidence that the low level of schooling is a direct consequence of the poverty and inequalities that became rooted in the country with the economic collapse of the early 1980s. Despite the country's having a long-standing tradition of schooling children and despite the large investments made to improve and modernise school infrastructures in the pre-recession years, the lack of individual resources has place school out of the reach of a considerable proportion of Malagasy families.

2

1.1. Simultaneity biases and hereditary biases

A large body of research has already addressed the impact of income on schooling decisions in the developing countries. However, we believe that this question has still not been perfectly answered. The main problem is that the correlations observed by the surveys between parental income and the level of children's schooling are merely an indirect and potentially biased reflection of the real impact of income on schooling. There are at least two reasons for this.

The first reason is the simultaneity of schooling decisions and production decisions in the family. Some factors simultaneously determine family income and whether or not it is worth schooling the children[3]. In the presence of such factors, the gross correlations between income and schooling are affected by a bias that can be called a simultaneity bias.

A second kind of problem has to do with family resources and skills passed on from generation to generation. Where these factors are received by the parents, they determine the family's productivity and current income. Where they are passed on to the children, they affect schooling decisions either by raising the anticipated return on schooling or by raising the anticipated return of training the children up on the family trade. In the presence of such factors, the correlations between income and schooling could be as much the manifestation of unobserved resources transmitted from parents to children as the manifestation of a real income effect on schooling.

In general, the existing literature essentially attemps to reduce biases due to heredity, i.e. the biases caused by resources transmitted from generation to generation and unmeasured by the surveys. One of the most frequently used econometric strategies is to analyse schooling and school performance differences between descendents of a same lineage[4]. Berhman and Wolfe (1987) use data collected in Nicaragua to analyse the differences in the number of years of schooling between pairs of cousins (whose mothers are sisters) as a function of observed differences in income and education between their parents. They find no significant relation between the observed schooling and resource differences. They conclude from this that the generally observed correlations between parental income and

---

[3] Let us assume that the occupational skills acquired by parents over their life are at least partially passed on to the children. When these skills are substitutes for educational skills, the most skilled parents in their trade are both those with the highest incomes and those with the least interest in sending their children to school.

[4] Pioneering studies in this area use American data: see Chamberlain and Griliches (1975) and Corcoran, Jenks and Olneck (1976).

children's schooling derive from unobserved abilities and resources passed on from generation to generation.

Another method consists of simultaneously analysing the parents' schooling and their children's schooling based on data covering a number of generations. Lillard and Willis (1992) use Malaysian data covering four generations to simultaneously estimate a schooling transition model for parents and children without excluding the possibility of a correlation between the unobserved determinants of the parents' transitions and the children's transitions. Assuming normal residuals and exogenous parental income, they conclude that the parents' education influences the children's education. Yet they do not identify any significant income effect[5].

This paper takes a new look at the impact of income on the level of schooling by attempting to model the causes for simultaneity and how heredity affects decisions and then to correct for the resulting biases. We simultaneously analyse the parents and children's education, but also the parents' income based on data covering three generations. We center our analysis on a single transition[6], but develop semi-parametric estimation techniques without any restrictive assumption about the residuals.

The article is structured as follows. First, we explain the theoretical framework in which our econometric work can be situated (Section 2). We then detail our econometric specifications and the semi-parametric estimators used (Section 3). Next, we describe the results of our estimates (Section 4). The datasets are described in detail in Appendix.

2. The theoretical framework

Generally speaking, our theoretical framework is part of the family of models with an imperfect credit market pioneered by Becker and Tomes (see, for example, Becker and Tomes, 1984). We consider dynasties (indexed by i) that are each made up of an infinite succession of generations (indexed by t). Each generation

---

[5] As the authors themselves say, the effect of parental income is nevertheless hard to interpret in the Lillard and Willis analysis (1992): a certain number of variables potentially linked to income (such as the quality of the dwelling) are also used as control variables. For a description of the problems posed by the joint estimation of the income effect and variables potentially linked to income, see, for example, Blau (1999).

[6] In so doing, we avoid the basic problems of non-parametric identification of transition models in the education system, as detailed by Cameron and Heckman (1998).

experiences two periods: in the first one, it lives with its parents' generation and, in the second one, with its children's generation. In each period, the parents determine the allocation of income between current consumption and educational investment in order to maximise the discounted welfare of the succession of generations. A key point in this is the limited access to credit: the parents cannot contract debts on behalf of their children[7]. Within this framework, if we denote $U(x)$ as the utility function, then the objective of dynasty i's generation t is written:

$$\max E \sum_{k=0}^{\infty} \beta^{t+k} U(C_{it+k}) \tag{2.1}$$

subject to: $C_{it+k} + cS_{it+k} = Y_{it+k}$ and $Y_{it+k} = F(S_{it+k-1}, u_{it+k-1}, \varepsilon_{it+k})$,

where Y_{it} is the income of the parents of generation t, C_{it} their level of consumption and S_{it} a dummy variable that takes the value 1 if the parents send their children to school. The random variable u_{it-1} measures the quality of non-school resources passed on by generation $t-1$ to generation t and ε_{it} measures generation t's intrinsic productive capacity. Given that we can always redefine ε_{it} as the residual of its projection on u_{it-1}, ε_{it} can be assumed to be orthogonal to u_{it-1}. Function F is a production function. It describes how income is produced from formal and informal human capital. Parameter β is a discounting coefficient while parameter c measures the cost of sending one's children to school. It corresponds to the schooling costs (e.g. transport and clothing). It also corresponds to the loss of earnings implied by the decrease in the children's contribution to the family's work. [8].

In each period, the optimal schooling choice S_{it-1}^* depends on the formal (S_{it-1}) and informal (u_{it-1}) resources obtained from the previous generation and the current productivity parameter ε_{it}. We can therefore write $S_{it}^* = S^*(S_{it-1}, u_{it-1}, \varepsilon_{it})$.

For each period, we can also define the optimal level of consumption C_{it}^* as a function of S_{it-1}, u_{it-1} and ε_{it}, where,
$C_{it}^* = C^*(S_{it-1}, u_{it-1}, \varepsilon_{it}) = F(S_{it-1}, u_{it-1}, \varepsilon_{it}) - cS^*(S_{it-1}, u_{it-1}, \varepsilon_{it})$.

In this model, sending children to school involves a loss of current consumption and hence a decrease in current welfare,

---

[7]This could be interpreted as the consequence of a social norm whereby it is not possible to force an adult to reimburse a debt contracted by his or her parents.

[8]Children at school contribute less to the family's work. See Section 4.3 on this subject.

$$L(Y_{it}) = U(Y_{it}) - U(Y_{it} - c) \geq 0 \qquad (2.2)$$

Assuming that U is concave, the lower current income Y_{it} the greater the loss $L(Y_{it})$. The poorer the family, the more sensitive it is to the immediate cost of schooling. The decision to school also entails a welfare gain for generation $t + 1$, whose discounted value is written,

$$G_{it} = \beta E(U(C^*(1, u_{it}, \varepsilon_{it+1})) - U(C^*(0, u_{it}, \varepsilon_{it+1})) \geq 0.$$

Insofar as the ε_{it+1} shock is independent of past history, the anticipated gain G_{it} depends on the current period only via the intermediary of u_{it} and can be rewritten $G(u_{it})$. Children are sent to school if the immediate losses do not exceed the anticipated gains, i.e. $L(Y_{it}) \leq G(u_{it})$. Given that L decreases, this first-order condition is rewritten,

$$S_{it} = 1 \Leftrightarrow Y_{it} \geq L^{-1}(G(u_{it})) = Z(u_{it}) = z_{it} \qquad (2.3)$$

For any given u_{it}, condition (2.3) means that only the families with a high enough current income will see a point in sending their children to school. In this model, any exogenous increase in poor families' current incomes reduces the proportion of families under their schooling threshold z_{it}. The aim of this paper is specifically to identify this income effect and test the extent to which redistribution aimed at the poorest families might increase the level of education of future generations.

If u_{it} could be deemed orthogonal to u_{it-1} and ε_{it}, then z_{it} would be orthogonal to Y_{it} and there would be no real identification problem. The differences in the schooling rates of families with different incomes would be representative of the real impact of income on schooling[9].

The difficulty with the identification problem is due to the fact that we cannot a priori exclude the existence of correlations between u_{it-1} and u_{it} or between ε_{it} and u_{it}[10]. When such correlations are present, incomes Y_{it} are correlated with

---

[9]For each level of income Y, the proportion of schooled children among families with income Y would directly give an estimator of $F_z(Y)$, where F_z is the distribution function of thresholds z_{it}. Knowing the distribution of thresholds z_{it} in the population and knowing that this distribution is independent of income, there would be no problem estimating the impact on schooling of any change in the distribution of income.

[10]Variables u_{it-1} and u_{it} potentially represent two successive forms of the same technical and/or cultural assets. Insofar as the culture and techniques passed on from one generation to

thresholds z_{it} and the income effect can no longer be identified from the observation of just S_{it} and Y_{it}. Once variables Y_{it} and z_{it} are intercorrelated, schooling differences between different income groups reflect both the income effect and the fact that different income groups have different schooling thresholds z_{it}. In such circumstances, identifying the real effect of current income requires more than a simple analysis of the observed correlations between income and schooling. Before moving onto our proposed solution to this identification problem, we will develop two possible extensions of our basic model to broaden the scope of empirical applications and possible interpretations of the effect of income on schooling.

2.1. Extensions

The model developed in the previous section simplifies the schooling question to the extreme, since there are only two possible options: school the child or not. An immediate extension is to assume not just 2, but K possible degrees of investment in schooling. Variable S_{it} takes its values not in $\{0,1\}$, but in $\{0,...,K\}$. Denoting c_k as the cost of an investment $S_{it} = k$ and setting down $c_0 = 0$, we can define for all $k \geq 1$:

$$L_k(Y_{it}) = U(Y_{it} - c_{k-1}) - U(Y_{it} - c_k), \qquad (2.4)$$

the marginal loss of current utility associated with the decision $S_{it} = k$ and:

$$G_{ikt} = \beta E(U(C^*(k, u_{it}, \varepsilon_{it+1})) - U(C^*(k-1, u_{it}, \varepsilon_{it+1})), \qquad (2.5)$$

the marginal gain expected from this investment, where $C^*(k, u_{it}, \varepsilon_{it+1})$ represents the optimal consumption in $t+1$, conditional on an investment $S_{it} = k$ in t.

Assuming that $U(x)$ is concave in x, c_k is convex in k and Y_{it} is concave in S_{it-1}, we verify that the current marginal loss L_{ikt} increases with k while the anticipated marginal return G_{ikt} decreases with the schooling investment. In this context, each family chooses either not to invest in school (if $G_{i1t} - L_{i1t} \leq 0$) or the highest schooling investment among those whose net return $G_{ikt} - L_{ikt}$ is

---

the next only change slowly over time, u_{it-1} and u_{it} have a good chance of being intercorrelated. Variables ε_{it} and u_{it} characterise the same adults (i.e. the parents of generation t). They are determined in the same context and in the same period of time. There is also good reason to think that these two variables are intercorrelated, for example, because parents with specific informal skills (strong ε_{it}) can pass their trade onto their children at home without sending them to school (low u_{it}).

positive. Denoting z_{ikt} as the threshold corresponding to $L_k^{-1}(G_{ikt})$, and setting down by convention $z_{i0t} = 0$ and $z_{iK+1t} = \infty$, we verify that the series of Z_{ikt} increases and that the decision $S_{it} = k$ is equivalent to $(z_{ikt} \leq Y_{it} < z_{ik+1t})$. The extent of investment in schooling can therefore be analysed very simply using a multinomial model.

In the basic model, educational investment costs are completely exogenous. Another possible extension of the basic model is to assume that schooling difficulties can vary from one family to the next depending on exogenous characteristics such as the child's gender and his or her actual age when schooling becomes compulsory (i.e. as an indicator of his or her maturity). Testing this type of assumption is simply a question of testing whether the schooling threshold varies from one family to the next in accordance with the gender and age of the children.

It is also possible to assume that the returns to education vary with the parents' level of schooling S_{it-1}. The idea here is that the parents with basic education might be in a better position to help their children benefit from school. Adopting this type of assumption is tantamount to assuming that Y_{it} depends not only on S_{it-1}, u_{it-1} and ε_{it}, but also on S_{it-2}. In this context, we can easily check that the anticipated gains and schooling thresholds vary both in line with the resources u_{it} transmitted from generation to generation and with the parents' education S_{it-1}. In the econometric application, we propose a number of tests of this assumption that the parents' educational capital has a direct effect on the probability of schooling the children. In general, we reject it and conclude that the effect of the parents' schooling on decisions to school the children is in itself weak.

3. Econometric specifications

To specify our empirical models, we assume that production function F combines schooling capital (S_{it}) with the other forms of productive resources (u_{it}) in line with a technology with a constant elasticity of substitution. We also assume that utility U is concave and homogenous (i.e. $U(x) = x^\alpha$ with $\alpha \in [0, 1]$) and that the costs of schooling remain low against income. In this framework, the first-order approximation of $L(Y_{it})$ is $\beta\rho Y_0^\alpha \exp(\phi u_{it})$, while G_{it} approximates $\beta\rho Y_0^\alpha \exp(\phi u_{it})$, where ρ represents the return to education and ϕ a constant whose sign and value depend on the elasticity of substitution between the schooling capital and the other forms of productive resources (see Appendix A). Based on these assumptions, the schooling decision can be written,

$$S_{it} = I(a \ln Y_{it} + bX_{it} + \phi u_{it}), \hspace{3cm} (3.1)$$

where $I(x)$ is a dummy function that takes the value 1 whenever x is positive and where we have assumed that the relative costs of schooling, i.e. $\ln(\frac{\beta \rho}{c})$, vary exogenously from one family to the next and can be written as a linear combination of the variables X_{it} observed in the surveys. The problem is how to identify a, given that we also have,

$$\ln Y_{it} = cS_{it-1} + u_{it-1} + \varepsilon_{it}. \hspace{3cm} (3.2)$$

3.1. The identification problem

Before describing the different estimators used, we will briefly outline our identification strategy and the procedures used to test the validity of this strategy. For the sake of simplicity, we temporarily drop index i and temporarily assume that S_t can be treated as a linear function of $\ln Y_t$ and X_t. In the following section, we explain the conditions that make this assumption valid. Based on the above hypotheses, relations (3.1) and (3.2) give rise to a system of linear relations between education and income,

$$
\begin{aligned}
S_t &= aY_t + bX_t + \phi u_t, \\
Y_t &= cS_{t-1} + u_{t-1} + \varepsilon_t, \\
S_{t-1} &= aY_{t-1} + bX_{t-1} + \phi u_{t-1}, \\
Y_{t-1} &= cS_{t-2} + u_{t-2} + \varepsilon_{t-1}...
\end{aligned}
$$

where (to reduce the notations) Y_t now represents the logarithm of income. The problem posed is how to estimate parameter a. Given the dynamic structure of the links between education and income, the "right" strategy for estimating a clearly depends on the variance-covariance structure of residuals u_t and ε_t. A number of cases can a priori be envisaged.

(i) In the first case, u_t is orthogonal to ε_t and u_{t-1} and all their past realisations (i.e. $E(u_t \varepsilon_{t-k}) = E(u_t u_{t-1-k}) = 0$, for all $k \geq 0$). This is the simplest case where what is passed on from parents to children (u_{t-1}) is correlated neither with what is passed on from grandparents to parents (u_{t-2}) nor with the parents' intrinsic productive capacities (ε_t). Based on these assumptions, $E(Y_t u_t) = 0$ and there is no particular identification problem. A simple regression can be used to estimate

9

the income effect. The correlations between Y_t and S_t give an accurate idea of the real impact of Y_t on S_t.

(ii) In the second case, u_t is orthogonal to u_{t-1} and ε_{t-1} (and their past realisations), but not to ε_t. In this case, there is a link between the parents' intrinsic productive ability (i.e. skills acquired during life) and what they pass onto their children, in particular by contributing to their occupational training. In this case, Y_t is no longer necessarily orthogonal to u_t and the correlation between Y_t and S_t is a potentially biased estimator of the effect of income on schooling u_{t-1}.

In this case, however, S_{t-1} (parents' education) has the dual property of being correlated with Y_t without being correlated with u_t (i.e. $E(S_{t-1}u_t) = E(u_tu_{t-1}) = 0$). The income effect can therefore be identified using S_{t-1} as an instrumental variable and all the lineage's performances previous to S_{t-1}, i.e. Y_{t-1}, S_{t-2}, etc.

(iii) If u_t is orthogonal to u_{t-2} and ε_{t-1} (and all their past realisations), but not u_{t-1} or ε_t, then S_{t-1} becomes potentially correlated with ut and is no longer an instrumental variable usable to identify the impact of Y_t. However, Y_{t-1} (grandparents' income) remains a valid instrument as do all the previous performances, especially S_{t-2}.

(iv) If u_t is orthogonal to u_{t-2} and ε_{t-2} (and all their past realisations), but not u_{t-1} or ε_{t-1}, then Y_{t-1} is no longer a good instrumental variable, but S_{t-2} (grandparents' education) still is as are all the previous performances.

Our problem in choosing between the different estimators and instrumental variables is obviously that we do not know the real variance-covariance structure of the non-school determinants of income (i.e. u_t and ε_t). The main asset we have to help us solve this problem is our wealth of data. These data can be used to measure the current educational investment (S_t), the parents' income and education (S_{t-1} and Y_t), and the grandparents' education (S_{t-2}). They can also be used to measure the grandparents' professional status (farmer, employee in the formal sector or worker in the informal sector), which we can consider to form an indirect measurement of the grandparents' income level (Y_{t-1}).

We can use the information on the education and income of three successive generations to estimate the impact of Y_t on S_t in four different ways: without using an instrumental variable procedure, using the parents' education as an instrument, using the grandparents' income as an instrument, and using the grandparents' education as an instrument.

If the structure of the residuals corresponds to case (i), then the four estimates clearly have to give identical results. If the structure of the residuals corresponds to case (ii), only the three instrumental regressions deliver identical results. The

10

overidentification tests should not reject the hypothesis of the consistency of instruments S_{t-1}, Y_{t-1} and S_{t-2}. If the structure of the residuals corresponds to case (iii), then only instruments Y_{t-1} and S_{t-2} should be consistent, while in case (iv), Y_{t-1} and S_{t-2} are no longer necessarily consistent. In other words, by comparing the results obtained using the four different instruments available, we can test the extent to which certain conditions necessary for the absence of a hereditary bias or a simultaneity bias may or may not be satisfied.

In general, these conditions are necessary, but not sufficient. The structure of the possible residuals need to be specified in more detail in order to express sufficient conditions for the absence of simultaneity and hereditary biases. In Appendix B, we detail the case in which the residuals are combined in line with a composed error model $u_{it} = u_i + \theta\varepsilon_{it}$. This specification is simple enough to be used to construct tests and general enough for the different forms of bias to be combined. In this framework, testing for the absence of a simultaneity bias amounts to testing whether $\theta = 0$ and testing for the absence of a hereditary bias comes down to testing whether $\sigma_u^2 = 0$, where σ_u^2 represents the variance of the fixed dynasty effects. In Appendix B, we show that a necessary and sufficient condition for the existence of a simultaneity bias and the absence of a hereditary bias is that the estimates obtained using S_{t-1}, Y_{t-1} and S_{t-2}, and even $(S_{t-1} - S_{t-2})$ as instrumental variables produce the same results and that these results are different from those obtained by the ordinary least squares technique.

3.2. A semi-parametric estimator

In this paper, we look at the decision as to whether or not to school children. The dependent variable is a dummy variable and the model considered is not a linear model, but a binary choice model. In this framework, the problem of identifying the income effect is more complicated than suggested by the previous discussion.

In the linear case, identification calls simply for an observation of an instrumental variable Z_t in the usual sense such that $E(Z_t u_t) = 0$ and $E(Z_t Y_t) \neq 0$. These conditions are no longer sufficient in the case of the binary choice model. A fair amount of literature has recently been developed to explore the different complementary hypotheses whereby the identification of the effect of an endogenous explanatory variable becomes possible again in a non-linear model (see the Blundell and Powell survey, 2000). This paper draws on Lewbel's recent contribution, which we feel to be particularly well suited to our problem and the data

at our disposal[11]. Lewbel (2000) quite generally shows how to identify the effect of an endogenous explanatory variable Y_t in a binary choice model of the form $S_t = I(aY_t + bX_t + u_t)$, where X_t is a set of exogenous variables. The method used requires to observe (a) an instrumental variable Z_t (i.e. such as $E(Z_t u_t) = 0$ et $E(Z_t Y_t) \neq 0$) and (b) an explanatory variable x_{0t} in X_t that is continuous[12] and such that the distribution of u_t conditionally on Y_t and X_t is independent of x_{0t}.

In our case, the problem of identifying the effect of family income on the decision to school is the same as an identification problem in a linear model when an exogenous and continuous determinant of the decision to school children is observed.

Lewbel (2000) moreover establishes that, in the case whereby an exogenous and continuous x_{0t} variable is observed, the effect of the endogenous variable is identified by applying the usual instrumental variables method to the linearised model $LS_t = aY_t + bX_t + \phi u_t$, where LS_t corresponds to $S_t - I(x_{0t} > 0)$ divided by the density of x_{0t} conditionally on (X_{1t}, Z_t), where X_{1t} corresponds to X_t minus x_{0t}.

In other words, once we can observe a continuous and exogenous determinant of the decision to send children to school, the problems of identifying and estimating parameter a are exactly the same as those analysed in the previous sub-section by replacing S_t with its linearisation LS_t.

In our specific case, there are at least two possible candidates for x_{0t}. The first is the child's date of birth in the year. This variable can reasonably be assumed to be exogenous. It determines the child's level of maturity on the date on which he or she can start school. For a given age group (in the school institution sense), the later in the year the child is born, the younger he or she is and the smaller his or her chances of being schooled. Of the children aged 6 to 8 when the survey was taken, 41% of those born in the first half of the year started school when they were supposed to as opposed to only 32% of those born in the second half of the year (Table 1). Obviously, the older the children considered, the less noticeable the effect of the date of birth on schooling. Of the teenagers aged 15 to 17 when the survey was taken, 21% of those born in the first half of the year had never been to school as opposed to 18% of those born in the second half of the year. Given the size of the sample, the deviation is significant, but small.

---

[11]Maurin (1999) applies Lewbel's estimator to the analysis of repeating years of primary school based on French data covering several generations with a structure close to ours.

[12]The interval of variation of x_{0t} also has to be broad and (even if it means redefining the variable) contain 0.

In other words, the date of birth in the year provides a definitely pertinent yardstick for the semi-parametric analysis of the effect of income on the age at which children start school. However, it is less suited to providing a gauge of the effect of income on the total length or total absence of schooling. A second candidate for x_{0t} is the quality and density of school infrastructures in the region in which the child lives. Our survey can be used to reconstitute the number of primary schools per child aged 6 to 15 in the child's commune ("fokontany"). This variable is continuous and a priori represents a schooling factor[13]. The problem is that parents may decide to move to another region to be closer to school infrastructures. In other words, it is not certain that the infrastructure density available to the children is completely exogenous to the schooling process studied. However, the survey we use contains the data needed to determine the region in which parents have lived before their migration[14]. We hence construct an indicator (DP_{it}) that measures the density of primary schools for the place in which the parents spent their childhood (if they have not moved) or before any migration (if they have moved). This variable is determined before the schooling process and the formation of current income. It can therefore be considered to be an exogenous measurement of the way in which the school supply quantity has influenced the schooling of the Malagasy children studied. Over 30% of the children aged 15 to 16 in families with a DP_{it} below the median have never been schooled. Only 9% of the 15 to 17 year old children of families with a DP_{it} above the median have never been schooled (Table 1).

3.3. The results and their interpretation

In this section, we present an econometric analysis of the impact of income on the decision to school children, applying the identification and estimation strategies described in sub-sections 3.1 and 3.2. The construction of the samples and variables used for this analysis are detailed in Appendix C. We have estimated four series of models, each corresponding to a particular choice for variable S_t or for auxiliary variable x_{0t} required for the construction of the semi-parametric estimator. The first series of models concentrates on the schooling of children aged 6 to 8 at the time of the survey. Variable S_t takes the value 1 if these children started

---

[13]The effects of school infrastructures have often been underscored, especially by Lavy (1996) and Glewwe and Jacoby (1994).

[14]Appendix C details the construction of this variable and the other variables used in the analysis.

school at 6 years old and 0 if not[15]. In 1993, over 60% of Malagasy children aged 6 to 8 did not start school when they became old enough to do so (Table 1). The second series of models analyses the schooling of teenagers aged 15 to 17 when the survey was taken. Variable S_t takes the value 1 if they attend school or have already been schooled, and 0 if not. This variable identifies those who will never go to school or, in any case, never under normal circumstances. Approximately 20% of the children had never been schooled even though they were old enough to be at secondary school. They will therefore never really be schooled. The third series of models studied analyses these same teenagers aged 15 to 17. This time, however, the models are trichotomic. Variable S_t takes three values: 0 if they have never been schooled, 1 if they were schooled for 1 to 4 years (i.e. shortened primary schooling) and 2 if they have already been schooled for more than 4 years.

In the first three series of estimates, the only explanatory variables considered are the child's gender, parental income (in logarithm form) measured by the household's consumption expenditures[16], and the child's date of birth. This last variable is used as a reference auxiliary variable for the construction of the semi-parametric estimator[17]. In the fourth series of regressions, we take the measurement of the density of the school infrastructure (of the region in which the father grew up) as a reference auxiliary variable and compare the results obtained with this reference to those obtained using the date of birth (Table 5).[18]

For each of the four different types of model, we provide (a) an ordinary least squares (OLS) estimate of the income effect, (b) a generalized method of moments' estimate (GMM) using the father's schooling as an instrument, (c) a GMM estimate using both the father and grandfather's schooling as instruments and testing

---

[15]A great deal of recent work has more specifically focused on analysing not exactly the total number of years spent at school, but the moment at which the decision is made to start school and then the moment at which the decision is made to attend school less and less regularly. Glewwe and Jacoby (1995) study the role of malnutrition in decisions to delay starting school. Jacoby (1994) studies the importance of indebtedness constraints on the timing of starting and gradually leaving school. More recently, DeVreyer et al (1999) and Bommier and Lambert (1999) have emphasised the role of informal skills that can be acquired in the family before starting school.

[16]The models have also all been estimated taking the income reported when the survey was taken as a measure of family income, i.e. a more direct measurement of family income, but also less well measured and more approximate than spending. The findings (not reported) obtained from this measurement are very similar to those obtained using spending.

[17]The estimated effects should be considered relative to the effect of the date of birth.

[18]Here, the estimated effects should be appreciated relative to the effect of this infrastructure density measurement.

the extent of these two instruments' consistency using Hausman and Sargan tests, (d) a GMM estimate using the *difference* between the father and grandfather's schooling as instrument. We also provide the results of the regressions (using the least squares technique and the instrumental variables technique) introducing the amount of the father's schooling as an additional regressor and using the grandfather's occupational status (taken as an indicator of his income level) as an additional instrumental variable.

3.3.1. Parental income and delayed schooling

Model (1) in Table 2 corresponds to the ordinary least squares estimate of the effect of income on the probability of being schooled at 6 years old. This estimate confirms what the statistics suggest: there is a significant link between parental income and the probability of starting school on time. This OLS estimator is only valid to the extent that both hereditary and simultaneity biases can be disregarded. Model (2) corresponds to the instrumental variables technique (GMM) re-estimation of parameter a using the father's extent of schooling as an instrumental variable. This estimator remains exposed to the hereditary bias, but escapes the simultaneity biases. The effect of parental income re-estimated in this way is three times higher ($\hat{a}_{iv1} = 0.62$) than the effect estimated by the ordinary least squares technique. In addition, the difference between the two estimators is significantly different to zero[19]. This finding suggests that the usual estimators obtained by the OLS technique are affected by relatively large simultaneity biases and tend to underestimate the income effect. Some unmeasured variables positively affect income and negatively affect the probability of being schooled.

To evaluate the magnitude of the hereditary biases, we have re-estimated the income effect by introducing the grandfather's length of schooling as an instrument in addition to the father's schooling (Model 3). We have also instrumented income using the difference between these two variables (Model 4). Each of these two estimators provides significantly higher results than the initial OLS estimator. Moreover, the results are neither significantly different from one another nor significantly different from the estimator obtained by taking the father's length of schooling as the only instrument. Furthermore, overidentification tests do not

---

[19]This estimate shows that a 50% parental income difference is equivalent to an age difference of 0.6 x 0.5 = 0.3 years at the time when schooling becomes compulsory. Given that the deviation in schooling probabilities between the youngest and oldest children in their age group is 25 points, a 50% increase in parental income implies a 7.5 point increase in the probability of the children being schooled at the normal age.

reject the hypothesis of the consistency of the two father and grandfather's education instruments.

In general, these findings are consistent with the assumption that the hereditary biases can be disregarded. If we accept that residuals u_{it} can be represented as the combination of a fixed effect (dynastic u_i) and a purely generational effect, we can unambiguously conclude that there are no hereditary biases (see the analysis developed in Appendix B). The only potentially large source of biases therefore looks to be the simultaneity of schooling decisions and income formation processes.

In addition to the diagnosis on the relative importance of the different forms of endogeneity biases, these preliminary estimates also show that the father's length of schooling (like the grandfather's) can be considered to be exogenous to the process of schooling the children. To complete this first series of models, we have therefore re-estimated our models by introducing the father's length of schooling as an additional regressor so that its effect can be compared with that of parental income. The ordinary least squares analysis suggests that the difference in the probability of schooling between children whose father went to secondary school and the other children (i.e. 0.35) is significantly greater than the difference in the probability of schooling caused by a doubling of income (i.e. 0.13). Remembering that parental income is endogenous and potentially negatively correlated with residual u_{it}, this least squares estimator of the effect of the father's schooling is therefore potentially biased. The bias a priori takes an opposite sign to that affecting the estimator of the parental income effect[20].

The instrumental variables technique analysis (i.e. using the grandfather's length of schooling and the grandfather's industry as instruments of the effect of parental income) provides quite a different finding to the least squares analysis. It confirms that the real effect of parental income is significantly higher than suggested by the naïve OLS analysis, but also confirms that the effect of the father's length of schooling is overestimated by the OLS estimator. Once the question of simultaneity is taken into account, the father's length of schooling has

---

[20]If ϕ represents the direct effect of S_t on S_{t+1} in Model (5) and $\widehat{\phi}_{mco}$ its OLS estimate, we verify that:

$B_s = E(\widehat{\phi}_{mco} - \phi) = \frac{(R_t' S_t)(R_t' u_t)}{(S_t' S_t)^2 - (R_t' S_t)^2} = B_R \frac{\gamma_0}{1-\gamma_0}$

where B_R is the bias affecting the OLS estimator of the parental income effect and γ_0 is the OLS estimator of the return on education. The B_s bias affecting the OLS estimator of the effect of the father's schooling hence definitely takes an opposite sign to that affecting parental income.

no significant effect as such on the children's length of schooling. Note that this finding is in keeping with the theoretical framework developed in the previous sections, but is at odds with the literature that posits that children's schooling is first and foremost a cultural problem.

3.4. Parental income and non-schooling

The second series of models (Table 3) analyses the probability of never having been to school by the age of 15. The third series (Table 4) corresponds to trichotomic models analysing the breakdown of children aged 15 to 17 between those who have never been to school, those who have not attended school enough to attain normal primary schooling, and those who have spent more than four years at school. These two series of models hence more explicitly address the evaluation of the effect of income on the process of non-schooling. They provide generally similar results consistent with those obtained regarding delayed entry into school:

   - The ordinary least squares regressions find a significant statistical link between the children's length of schooling and parental income;

   - The instrumental variables technique regressions provide mutually consistent estimates that are higher again than the estimates obtained using ordinary least squares. They reject the hereditary bias hypothesis, but suggest that simultaneity biases may exist biasing the naïve estimates towards zero;

   - The addition of the father's length of schooling as an extra regressor does not change the finding regarding the effect of parental income on the children's schooling. In general, the effect of a doubling in parental income (i.e. $\Delta \ln(R) \approx 1$) is significantly higher than the advantage of having a father who went to secondary school, since this advantage is not significantly different from zero (see models (12) and (18)).

   The fourth series of regressions draws on a new auxiliary variable: the density of primary schools in the district of origin (see Section 3.2 above for the exact definition). This variable captures the way in which the quantity and distribution of school infrastructures across the territory may have influenced the schooling processes. Table 4bis presents the findings of the semi-parametric estimates using the density of primary schools instead of the date of birth as the auxiliary variable. This series of regressions provides qualitatively similar findings to the previous series[21]. The effect of parental income appears to be wholly significant in both

---

[21]In terms of point estimations, the findings suggest that a 50% increase in an average household's income would raise the probability of schooling by an equivalent amount to that which

17

the least squares regressions and the instrumental variables technique regressions. The hereditary biases again appear negligible. Even if the estimates are not as accurate, the simultaneity biases still seem to prevail.

To wind up, we re-estimated the models presented in tables 4 and 4bis based on the assumption of residual normality (Table 5). In this parametric framework, our model has a Tobit model structure. A first equation describes parental income as a function of the instrumental variables used (typically, the grandfather's schooling). A second equation describes the schooling process as a function of parental income and the residuals of these two equations are assumed to comply with a bivariate normal law. The estimators correspond to maximum likelihood. In general, this parametric approach provides results consistent with those previously obtained using semi-parametric estimators. Compared to the coefficients of the other variables, especially the auxiliary variables, the parental income effect more than doubles when the father and grandfather's schooling are used as instruments[22]. This finding still holds when the father's schooling is included as an additional regressor and when the grandfather's schooling and industry are used as sole instruments[23]. The correlation coefficient ρ estimated for the bivariate normal distribution takes a negative sign, confirming the nature of the simultaneity biases affecting the naïve estimation of the income effect.

3.5. Extension: the effect of child labour

The model developed in Section 2 does not explicitly introduce child labour as a possible alternative to schooling. In fact, including child labour changes nothing in the theoretical analysis except that (a) the schooling costs are taken to include the losses of current earnings due to sending the children to school rather than to work, (b) parental income (i.e. variable Y_t) is taken to be the parents' income excluding the children's contribution. Based on the assumption that child labour is not a marginal phenomenon and given that we only observe total income in the surveys, our main explanatory variable is potentially affected by a measurement error. Rather than using Y_t as a regressor, we actually use $Y_t + h_t$, where h_t represents the children's contribution. In this framework, the ordinary least squares estimate

---

might be prompted by an increase in the school supply of 0.25 x 0.5 = 0.125 primary schools per child, i.e. approximately 15 points.

[22] In this non-linear framework, using the grandfather's education as the instrument means using the grandfather's education as an explanatory variable in the equation explaining parental income, without using it in the equation explaining schooling.

[23] As in model 6 and tables 4 and 6.

of the effect of income on schooling potentially suffers from a measurement error bias, whose effects are similar to a simultaneity bias tending towards zero. In other words, the simultaneity biases shown by our previous analyses are potentially interpreted as measurement error biases. To test this interpretation, we have used the information available on child labour in our survey by introducing, as an additional regressor, a dummy variable taking the value 1 if the child works and zero if not. In 1993, 15% of children aged 7 to 8 had worked in Madagascar over the year. In general, the ordinary least squares regressions confirm that child labour is negatively correlated with a child's length of schooling, even though the effect is poorly estimated and not significantly different from zero (Table 6, models 30 and 33). Its effect disappears totally when this variable is instrumented along with parental income. However, the re-estimation of the parental income effect does not change.

4. Conclusion

In this paper, we develop a fairly general detection and rectification methodology for the bias that could affect the estimation of the effect of parental income on the decision to send children to school. Our estimates show that the main source of bias comes from the simultaneity of the decision to school children and the parental income formation process. Certain factors not measured by the surveys increase the parents' productivity and income and simultaneously undermine the sense of sending children to school. These factors are typically the skills that can be informally passed from one generation to the next without the need for schooling. The greater these skills, the more productive the parents and the less point they will see in investing in formal education for their children. Given the existence of these factors, the gross correlation between family income and the probability of schooling children tends to underestimate the causal effect of income on schooling. Neutralising the effect of these factors means that we have to considerably re-estimate up the income effect.

In the Malagasy case, the "real" effect of income seems to be quite considerable overall. A 50% drop in parental income results in an approximately 20-point decrease in the probability of sending children to school. A good deal of the drop in primary schooling in Madagascar since the early 1980s can therefore be explained not so much by the dilapidation or lack of school infrastructures, but more by the degeneration in household incomes and the spread of poverty due to the economic slump in the early 1980s.

In the planet's poorest countries, especially in sub-Saharan Africa, universal primary schooling for children is still a remote goal at the end of a long, hard road. In many countries, the schooling situation is now stagnating and sometimes even deteriorating. At first glance, the children in these countries are more excluded from the education system when they come from poor families. Our study's findings suggest that this problem could be partially solved by a reasonable increase in the redistribution effort to help the poorest families and a reduction in the most extreme forms of inequality. The question nonetheless remains open as to how exactly this redistribution effort could be put into practice. There are many potential ways in which income affects the decision to school children. More income means better nutrition, better housing, better health and the ability to get around more easily. It is the extent to which these fundamental problems are better solved that dictates whether the family can envisage investing more in the formal education of their children.Future research should help identify which of the basic goods are politically most likely to be redistributed and economically most efficient to redistribute in order to promote a long-term development of education and standards of living.

References

[1] Becker G. S. and N. Tomes, 1984, Human Capital and the Rise and Fall of Families, *Journal of Labor Economics*, 4, S1-S39.

[2] Behrman J.R. and B. L. Wolfe, 1987, Investments in Schooling in Two Generations in Pre-revolutionary Nicaragua, *Journal of Development Economics*, 27, 395-419.

[3] Blau D.M., 1999, The Effect of Income on Child Development, *The Review of Economics and Statistics*, Vol LXXXI, n°2,261-277.

[4] Blundell R. and J. Powell, 2000, Endogeneity in Nonparametric and Semi-parametric Regression Models, Unpublished Manuscript, University College London, University of California at Berkeley.

[5] Bommier A., S. Lambert, 2000, Education Demand and Age at School Enrollment in Tanzania, *Journal of Human Resources*, vol. XXXV, n°1.

[6] Cameron S.V and J.J. Heckman, 1998, Life Cycle Schooling Decisions, *Journal of Political Economy*,

[7] Chamberlain, G. and Z. Griliches, 1975, Unobservables with a variance Component Structure: Ability, Schooling and the Economic Success of Brothers, *International Economic Review*, 16, 2, 442-449.

[8] Cogneau D., J.-C. Dumont. P. Glick, M. Razafindrakoto, J. Razafindravonona , I. Randretsa, F. Roubaud (2001), Pauvreté, dépenses d'éducation et de santé, le cas de Madagascar, OECD Development Centre, forthcoming.

[9] Corcoran M., C. Jencks and M.R. Olneck, 1976, The Effects of Family Background on Earnings, *American Economic Review*, 66, 430-435.

[10] De Vreyer Ph., S. Lambert, Th. Magnac, 1999, Educating Children: A Look at Household Behavior in Côte d'Ivoire, mimeo CREST & INRA-LEA, 50 pp.

[11] Filmer D., and L. Pritchett, 1999, The Effect of Household Wealth on Educational Attainment: Evidence from 35 countries, *Population and Development Review*, 25, 1, 85-120.

[12] Glewwe P. and H. Jacoby, 1994, Student Achievement and Schooling Choice in Low-Income Countries, Evidence from Ghana, *Journal of Human Resources*, vol. XXIX, n°3.

[13] Glewwe P. and H. Jacoby, 1995, Delayed Primary School Enrollment and Childhood Malnutrition in Ghana, An Economic Analysis, LSMS Working Paper n°98.

[14] Jacoby H., 1994, Borrowing Constraints and Progress Through School: Evidence from Peru, *Review of Economics and Statistics*, vol. LXXVI, n°1,151-161.

[15] Lavy V., 1996, School Supply Constraints and Children's Educational Outcome in Rural Ghana, *Journal of Development Economics*, 51, 2

[16] Lewbel A., 2000, Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables, *Journal of Econometrics*, 145-177.

[17] Lillard L.A. and R.J. Willis, 1998, Intergenerational Educational Mobility, Effects of Family and State in Malaysia, *Journal of Human Resources*, XXIX, 4, 1126-1166.

[18] Maurin E., 1999, The Impact of Parental Income on Early Schooling Transitions: A Reevaluation using Data on Three Generations, CREST Working Paper n°99-53.

Appendix A: The schooling condition

In this appendix, we analyze the schooling condition when the income production function is with constant elasticity of substitution.

Let $H_t = \exp \rho S_{t-1}$ be the measurement of the human capital generated by the schooling decision S_{t-1} (where ρ represents the return to education) and let $U_t = \exp u_t$ be the measurement of non-school productive resources. We assume that income can be written $Y_t = (1 + \varepsilon_{t+1})\varphi(H_t, U_t)$, where φ corresponds to a technology with a constant elasticity of substitution $\sigma = \frac{1}{1+\theta}, \theta \in [-1, \infty]$. We note $Y_{0t} = \varphi(1, U_t)$ the average income in the absence of schooling. With these notations, the expected effect of a decision to school is,

$$G_{it} = \beta E(U((1+\varepsilon_{t+1})\varphi(\exp \rho, U_t) - cS^*(1, u_{it}, \varepsilon_{it+1})) - U((1+\varepsilon_{t+1})Y_{0t} - cS^*(0, u_{it}, \varepsilon_{it+1}))$$

Assuming that ρ and c are small compared with Y_{0t}, the first-order approximation of G_{it} is,

$$G_{it} \approx \beta \rho U'(Y_{0t}) \frac{\partial \varphi}{\partial H}(1, U_t)) \approx \beta \rho \alpha (1 + \lambda(\theta + \alpha)u_{it})$$

The immediate loss associated with a decision to school can be written,

$$L(Y_{it}) = U(Y_{it}) - U(Y_{it} - c) \approx cU'(Y_{it}) = cY_{it}^{\alpha-1}$$

and the schooling condition can be written,

$$(1 - \alpha)\ln(Y_{it}) \geq \ln(\frac{\beta \rho \alpha}{c}) + \lambda(\theta + \alpha)u_{it},$$

which corresponds to the type of schooling condition studied in the article. Interestingly, the sign of $\phi = \lambda(\theta + \alpha)$ (and hence the sign of the correlation between the residuals of the schooling and income equations) depends on (a) the elasticity of substitution between the two forms of productive resources and (b) the elasticity of welfare with respect to consumption. If the two productive factors are sufficient substitutes (θ close to -1 and $(\theta + \alpha)$ negative), a strong non-school capital tends to reduce the probability of schooling. In general, the more complementary the income production factors (i.e. the higher θ), the more the households rich in non-school capital also tend to send their children to school.

Appendix B: Tests for the existence of simultaneity and hereditary biases

In this appendix, we show how the comparison of the different income effect estimators can be used to test for the existence of hereditary and simultaneity biases. Let the system of equations be,

$$S_{it} = aY_{it} + \phi u_{it}, \tag{4.1}$$

$$Y_{it} = cS_{it-1} + u_{it-1} + \varepsilon_{it}. \tag{4.2}$$

with $u_{it} = u_i + \theta\varepsilon_{it}$,, where the ε_{it} are randoms verifying $E(\varepsilon_{it}\varepsilon_{i't'}) = \sigma_\varepsilon^2$ if $i = i'$ and $t = t'$ and $E(\varepsilon_{it}\varepsilon_{i't'}) = 0$ if not, whilst the u_i verify that $E(u_iu_{i'}) = \sigma_u^2$ if $i = i'$ and $E(u_iu_{i'}) = 0$ if not.

It can be rewritten,

$$S_{it} = acS_{it-1} + \phi u_{it} + a(u_{it-1} + \varepsilon_{it}), \tag{4.3}$$

$$Y_{it} = acY_{it-1} + (1 + c\phi)u_{it-1} + \varepsilon_{it}, \tag{4.4}$$

which implies,

$$S_{it} = \frac{\phi + a}{1 - ac}u_i + \sum_{k=0}^{\infty} (ac)^k(a\theta\varepsilon_{it-1} + (\theta\phi + a)\varepsilon_{it}), \tag{4.5}$$

$$Y_{it} = \frac{1 + c\phi}{1 - ac}u_i + \sum_{k=0}^{\infty} (ac)^k(\theta(1 + c\phi)\varepsilon_{it-1} + \varepsilon_{it}). \tag{4.6}$$

These equations clearly show that the lagged values of income (Y_{it-k}, $k > 0$) and schooling ($S_{it-k}, k > 0$) are correlated with current income Y_{it}, even in the absence of simultaneity ($\theta = 0$) and heredity ($\sigma_u^2 = 0$). These lagged values are hence potential instruments for identifying the current income effect. Regarding their correlation with residual u_{it}, we verify that,

$$E(Y_{it-1}u_{it}) = \frac{1 + c\phi}{1 - ac}\sigma_u^2, E(S_{it-1}u_{it}) = E(S_{it-2}u_{it}) = \frac{\phi + a}{1 - ac}\sigma_u^2, \tag{4.7}$$

while,

$$E(Y_{it}u_{it}) = \frac{1 + c\phi}{1 - ac}\sigma_u^2 + \theta\sigma_\varepsilon^2. \tag{4.8}$$

24

When $\sigma_u^2 \neq 0$ and $\sigma_\varepsilon^2 \neq 0$, the ordinary least squares estimator is exposed to both simultaneity and hereditary biases, while the estimators using the income or schooling lags as instrumental variables are exposed essentially to hereditary biases. The estimations using the father's or grandfather's education as instruments are potentially affected by a hereditary bias, but the education differential between father and grandfather $(S_{it-1} - S_{it-2})$ makes it possible to construct an unbiased instrumental variable estimator. Testing for the absence of a hereditary bias (i.e., $\sigma_u^2 = 0$) is therefore simply a question of testing whether the estimator of a obtained using $(S_{it-1} - S_{it-2})$ as an instrument is different to that obtained using S_{it-2}. Based on the assumptions adopted concerning the residuals, equality between the two estimators is a necessary and sufficient condition for the absence of any hereditary bias. Another possible test is to compare the estimate of a obtained with S_{it-1} as an instrument with that obtained using S_{it-2}. The father's education is indeed better correlated with income than the grandfather's education: $0 < E(Y_{it}S_{it-2}) < E(Y_{it}S_{it-2})$. Since elsewhere,

$$E(S_{it-1}u_{it}) = E(S_{it-2}u_{it}) = \frac{\phi + a}{1 - ac}\sigma_u^2, \qquad (4.9)$$

we hence have,

$$\frac{E(S_{it-2}u_{it})}{E(Y_{it}S_{it-2})} = \frac{E(S_{it-1}u_{it})}{E(Y_{it}S_{it-1})} \iff \sigma_u^2 = 0. \qquad (4.10)$$

The absence of any significant difference between the estimate obtained using the father's schooling as an instrument and the estimate obtained using the grandfather's schooling is therefore another necessary and sufficient condition for the absence of any hereditary bias. Lastly, conditionally on the absence of any hereditary bias, a necessary and sufficient condition for the absence of any simultaneity bias (i.e., $\theta = 0$) is that there is no difference between the estimators obtained by the ordinary least squares technique and those obtained by the instrumental variables method.

Appendix C: Presentation of the data and construction of the variables

From April 1993 to April 1994, the EPM93 survey was taken of a sample of 4,508 Malagasy households stratified by six major "regions" (faritany) and three categories of "residence" (large urban centres/small towns/rural area). Sampling was by area: 320 "districts" (fokontany) were drawn from each stratum and some fifteen whole households (all household members) were interviewed in each selected fokontany. The urban households (approximately 20% of the population) were over-represented. Alongside the household survey, a "community" questionnaire was put to a local contact for the 220 fokontany drawn from outside the major urban centres. This contact supplied, in particular, information on the state of the available school infrastructures such as the number of primary schools. The household survey covered a very wide range of fields: education and health, employment, income and production conditions, migrations, consumption, etc. It was used from the educational point of view to reconstitute each respondent's schooling path: number of years spent at school, number of times kept back a year and the level reached. The respondents also reported on their parents' level of education. Where children lived with their parents, information was obtained on both the details of their own and their parents' schooling as well as their grandparents' level of education.

The survey provided information on the household's standard of living in the form of expenditure and the households' production for own consumption. This information was used to reconstitute a "current annual consumption" variable including for own food and non-food consumption valued at market prices, but excluding expenditure on durables. The survey was also used to reconstitute an "annual income" variable covering income from agricultural and non-agricultural family production (including for own consumption production), wages received and transfer income. Last but not least, the data were used to class the parents and grandparents into three major types of employment: agricultural/non-agricultural self-employed/wage earners.

Two sub-samples were constituted corresponding to two different analyses of primary school conditions.

The sample of children aged 6 to 8 years

The first sub-sample covers all children aged six to eight living in their parents' household or with their grandparents or uncle (these last two cases form a very small minority). The size of this sample is N = 1511. The variable analysed for this sample is a dummy variable taking the value 1 if the child has attended

school since the age of six. The survey directly provided the information as to whether six year olds were at school or not. For the seven and eight year olds, information on the class currently attended and the number of times the child had been kept back a year was used to reconstitute whether they were already at school at the age of six. The quality of this reconstitution can been appreciated from the fact that the estimated rate of starting school at six years old is the same (approximately 37%) for the children aged seven and eight when the survey was taken as for the children aged six when the survey was taken. The survey also provided the exact date of birth of each child, i.e. the day and month of the year in which they were born. It also detailed whether the children aged 7 and 8 when the survey was taken had been put to work during the survey's reference year.

The sample of adolescents aged 15 to 17

The second sample covers all adolescents aged 15 to 17 living in their parents' households or with their grandparents or uncle (both last cases form a very small minority). The size of this sample is N = 1063. The first variable analysed for this sample is a dummy variable taking the value 1 if the child has already attended school and 0 if not. A child who has not attended primary school by the age of 15 no longer has any chance of doing so. A full 20% of the children aged 15 to 17 were in this situation. The remaining 80% had either left primary school or, in the case of a minority, were still at school. We reconstituted the number of years these adolescents spent at primary school. We counted one to four years of attendance as "incomplete schooling", found for 27% of the population. We counted more than five years' attendance at primary school as "complete schooling" (i.e. the normal complete length of time without being kept back a year). Some 53% of the population aged 15 to 17 were found to have "complete schooling". This classification gave us a second, this time trichotomic, analytic variable taking the value 0 if the child had never been to school, 1 if he or she had incomplete primary schooling and 2 otherwise. Given the wealth of data, we were able to construct a similar variable (trichotomic) for the parents and grandparents of these adolescents. This variable identified the parents (or grandparents) who had never or virtually never attended primary school, those who had reached a level of primary education, and those who had attended secondary school and even university.

School density

The "community" questionnaire put at district (fokontany) level and the household survey's "migration" section were used together to create two variables describing the densities of school infrastructures liable to influence the schooling

process.

The number of primary schools in each district was available for most of the fokontany. When compared with the fokontany's population of 6 to 15 year olds, this indicator forms a measurement of the density of the primary school supply in the place where the children live. This variable is available for 220 fokontany out of the 320 surveyed. The majority of the fokontany for which the variable is unknown are located in the major urban centres. We ascribed them the maximum number of primary schools recorded in the other fokontany, i.e. three schools. To the other fokontany whose number of primary schools was unknown and which are located outside of an urban centre, we allocated the average school density variable corresponding to their survey stratum.

We then had a first school density variable for the place in which the children were currently living, after any migration by their parents. To correct the effect of parental migration, we constructed a second school density variable based on the first in the following manner. We firstly divided the households into two categories: those whose head had never left the current fokontany and the others. The second households account for 40% of the sample for the 15-17 year olds. The survey's "migration section" provided the head of household's previous residential stratum. We then attributed the average school density for this stratum to the 40% of households that had migrated. The second variable constructed in this way corresponds to the school density before parental migration.

Table 1: The extent of schooling of Malagasy children and adolescents in 1993: some descriptive statistics.

	6-8 years old (a)	15-17 years old		
	Schooling at 6 years old	Number of years in primary school		
		0	1-4	5 and over
Parental income(b)				
< median	25.8	29.4	36.0	34.6
> median	47.9	9.6	17.7	72.7
Density of school supply				
< median	30.7	29.8	32.1	38.1
> median	43.1	9.1	21.5	69.4
Semester of birth				
First	41.2	17.7	27.1	55.2
Second	32.0	21.3	26.5	52.2
Total	36.9	19.5	26.8	53.7
N	1511	1063		

Source: EPM 1993 survey.
(a) The statistics are very similar for the sub-sample of 7-8 year olds, covering 985 individuals, used in Table 6.
(b) Parental income is measured here by the total volume of household expenditure.

Table 2: Effect of parental income on the probability of starting school at six years old.

Independent variable	1 OLS	2 FE	3 FE&GFE	4 FE-GFE	5 OLS	6 GFE&GFS
Boy	+0.002	-0.074	+0.008	-0.028	+0.003	-0.008
	(0.041)	(0.051)	(0.047)	(0.052)	(0.050)	(0.050)
Income [ln]	+0.221	+0.617	+0.721	+0.663	+0.129	+0.469
	(0.036)	(0.091)	(0.109)	(0.152)	(0.040)	(0.314)
Father≥secondary					+0.348	+0.100
					(0.064)	(0.227)
Hausman stat.			0.247			1.629
Test level			0.970			0.804
Sargan stat.			0.244			1.609
Test level			0.621			0.447

Source: EPM 1993 survey. Field: All children aged six to eight years. Reading: The dependent variable is a dummy taking the value 1 if the child started school at six years old and 0 if not. It is linearised using the Lewbel technique (2000) by taking the day of birth in the year as an auxiliary variable. Models 1 and 5 are estimated by the ordinary least squares technique. Models 2, 3, 4 and 6 are estimated by the generalised method of moments (GMM). In Model 2, parental income is instrumented by the father's education (i.e. by FE, dummy variable taking the value 1 if the father went to secondary school). In Model 3 (resp. 4), parental income is instrumented by the father's education and the grandfather's education (resp. by the difference between the father's education and the grandfather's education). In Model 6, parental income is instrumented by the grandfather's education and activity sector (i.e. a set of three dummy variables indicating whether the grandfather was a farmer, a non-agricultural self-employed worker or other). The coefficient's standard deviation is given in brackets. Parental income is measured by the household's total expenditure.

Table 3: Effect of parental income on the probability of never being schooled.

Independent variable	7 OLS	8 FE	9 FE&GFE	10 FE-GFE	11 OLS	12 GFE&GFS
Boy	+0.027	+0.041	+0.009	-0.009	+0.029	+0.066
	(0.053)	(0.054)	(0.055)	(0.062)	(0.053)	(0.061)
Income [ln]	+0.150	+0.385	+0.368	+0.439	+0.101	+0.689
	(0.036)	(0.087)	(0.077)	(0.200)	(0.040)	(0.236)
Father≥secondary					+0.193	-0.244
					(0.069)	(0.184)
Hausman stat.			0.046			5.661
Test level			0.997			0.226
Sargan stat.			0.046			5.655
Test level			0.831			0.059

Source: EPM 1993 survey. Field: All adolescents aged 15 to 17 years. Reading: The dependent variable is a dummy taking the value 1 if the adolescent has already attended school and 0 if not. It is linearised using the Lewbel technique (2000) by taking the day of birth in the year as an auxiliary variable. Models 7 and 11 are estimated by the ordinary least squares technique. Models 8, 9, 10 and 12 are estimated by the generalised method of moments. In Model 8, parental income is instrumented by the father's education (i.e. by FE, dummy variable taking the value 1 if the father went to secondary school). In Model 9 (resp. 10), parental income is instrumented by the father's education and the grandfather's education (resp. by the difference between the father's education and the grandfather's education). In Model 12, parental income is instrumented by the grandfather's education and activity sector (i.e. a set of three dummy variables indicating whether the grandfather was a farmer, a non-agricultural self-employed worker or other). The coefficient's standard deviation is given in brackets. Parental income is measured by the household's total expenditure.

Table 4: Effect of parental income on the extent of schooling of adolescents aged 15 to 17: a semi-parametric estimate (auxiliary variable: day of the year in which the child is born).

Independent variable	13 OLS	14 FE	15 FE&GFE	16 FE-GFE	17 OLS	18 GFE&GFS
Boy	+0.033	-0.008	+0.057	+0.099	+0.036	+0.067
	(0.051)	(0.050)	(0.052)	(0.056)	(0.050)	(0.053)
Income [ln]	+0.204	+0.499	+0.556	+0.605	+0.118	+0.506
	(0.034)	(0.072)	(0.075)	(0.161)	(0.038)	(0.174)
Father primary ed.					+0.239	+0.153
					(0.061)	(0.076)
Father$\geq$secondary					+0.419	+0.065
					(0.079)	(0.174)
Hausman stat.		2.457	6.865	1.754		3.130
Test level		0.467	0.076	0.625		0.680
Sargan stat.		2.557	6.820	1.744		3.099
Test level		0.110	0.078	0.187		0.377

Source: EPM 1993 survey. Field: All adolescents aged 15 to 17 years. Reading: The dependent variable takes the value 1 if the child has never been to school, 1 if he or she has only had incomplete primary schooling and 2 otherwise. It is linearised using the Lewbel technique (2000) by taking the day of birth in the year as an auxiliary variable. Models 13 and 17 are estimated by the ordinary least squares technique. Models 14, 15, 16 and 18 are estimated by the generalised method of moments. In Model 14, parental income is instrumented by the father's education (i.e. FE, two dummy variables, the first taking the value 1 if the father only had primary schooling and the second taking the value 1 if the father reached secondary school). In Model 15 (resp. 16), parental income is instrumented by the father's education and the grandfather's education (resp. by the difference between the father's education and the grandfather's education). In Model 18, parental income is instrumented by the grandfather's education and activity sector (i.e. a set of three dummy variables indicating whether the grandfather was a farmer, a non-agricultural self-employed worker or other). The coefficient's standard deviation is given in brackets. Parental income is measured by the household's total expenditure.

Table 4b: Effect of parental income on the extent of schooling of adolescents aged 15 to 17 years: a semi-parametric estimate (auxiliary variable: primary school density).

Independent variable	19 OLS	20 FE	21 FE&GFE	22 FE-GFE	23 OLS	24 GFE&GFS
Boy	-0.021	+0.078	+0.060	+0.071	-0.018	-0.005
	(0.034)	(0.051)	(0.048)	(0.044)	(0.034)	(0.035)
Revenu [ln]	+0.078	+0.265	+0.338	+0.231	+0.014	+0.184
	(0.023)	(0.074)	(0.086)	(0.147)	(0.026)	(0.118)
Father primary ed.					+0.051	+0.009
					(0.041)	(0.041)
Father≥secondary					+0.266	+0.094
					(0.058)	(0.106)
Hausman stat.		1.114	0.457	2.392		0.534
Test level		0.774	0.928	0.495		0.991
Sargan stat.		1.113	0.458	2.452		0.564
Test level		0.291	0.928	0.117		0.905

Source: EPM 1993 survey. Field: All adolescents aged 15 to 17 years. Reading: The dependent variable takes the value 1 if the child has never been to school, 1 if he or she has only had incomplete primary schooling and 2 otherwise. It is linearised using the Lewbel technique (2000) by taking the density of primary schools as an auxiliary variable. Models 19 and 23 are estimated by the ordinary least squares technique. Models 20, 21, 22 and 24 are estimated by the generalised method of moments. In Model 20, parental income is instrumented by the father's education (i.e. FE, two dummy variables, the first taking the value 1 if the father only had primary schooling and the second taking the value 1 if the father reached secondary school). In Model 21 (resp. 22), parental income is instrumented by the father's education and the grandfather's education (resp. by the difference between the father's education and the grandfather's education). In Model 24, parental income is instrumented by the grandfather's education and activity sector (i.e. a set of three dummy variables indicating whether the grandfather was a farmer, a non-agricultural self-employed worker or other). The coefficient's standard deviation is given in brackets. Parental income is measured by the household's total expenditure.

Table 5: Effect of parental income on the extent of schooling of adolescents aged 15 to 17 years: a parametric estimate (ordered Tobit).

Independent variable	25 -	26 FE&GFE	27 -	28 GFE&GFS
ρ	0	-0.697	0	-0.348
	-	(0.039)	-	(0.102)
Boy	-0.050	-0.027	-0.046	-0.035
	(0.076)	(0.056)	(0.078)	(0.074)
Date of birth	-0.075	-0.055	-0.078	-0.073
	(0.133)	(0.097)	(0.135)	(0.127)
Income [ln]	+0.685	+1.447	+0.514	+0.974
	(0.060)	(0.057)	(0.062)	(0.137)
School density	+4.157	+2.185	+3.206	+2.774
	(0.315)	(0.302)	(0.369)	(0.399)
Father primary ed.			+0.539	+0.455
			(0.087)	(0.089)
Father$\geq$secondary			+1.481	+1.259
			(0.136)	(0.154)

Source: EPM 1993 survey. Field: All adolescents aged 15 to 17 years. Reading: The dependent variable takes the value 1 if the child has never been to school, 1 if he or she has only had incomplete primary schooling and 2 otherwise. The models are estimated by the maximum likelihood technique. Models 25 and 27 are ordered probits. Models 26 and 28 are ordered tobits. In models 26 and 28, the parental income equation, estimated together with the schooling equation, is not presented. In Model 26, parental income is explained by the father's education and the grandfather's education. In Model 28, parental income is explained by the grandfather's education and activity sector (i.e. a set of three dummy variables indicating whether the grandfather was a farmer, a non-agricultural self-employed worker or other). The coefficient's standard deviation is given in brackets. Parental income is measured by the household's total expenditure.

Table 6: The link between child labour and schooling.

Independent variable	29 OLS	30 OLS	31 IV3	32 OLS	33 OLS	34 IV3
Boy	+0.059 (0.041)	+0.061 (0.069)	+0.009 (0.070)	+0.081 (0.069)	+0.083 (0.069)	+0.076 (0.081)
Income 1 [ln]	+0.217 (0.049)	+0.211 (0.050)	+0.791 (0.161)			
Income 2 [ln]				+0.129 (0.034)	+0.126 (0.034)	+0.693 (0.153)
Work [1=yes]		-0.090 (0.100)	+0.500 (0.605)		-0.116 (0.100)	+1.200 (0.744)
Hausman stat.			0.233			0.950
Test level			0.994			0.917
Sargan stat.			0.229			0.938
Test level			0.892			0.626

Source: EPM 1993 survey. Coverage: All adolescents aged 7 to 8 years. Reading: The dependent variable is a dummy taking the value 1 if the child started school at six years old and 0 if not. It is linearised using the Lewbel technique (2000) by taking the day of birth in the year as an auxiliary variable. Models 29, 30, 32 and 33 are estimated by the ordinary least squares technique. Models 31 and 34 are estimated by the generalised method of moments, taking the father's education, the grandfather's education and the grandfather's socio-economic status as instruments. The coefficient's standard deviation is given in brackets. Parental income is measured by the household's total expenditure (income 1) or by the household's stated income (income 2).