CEPREMAP

CENTRE POUR LA RECHERCHE ECONOMIQUE ET SES APPLICATIONS

Patterns of Intergenerational Transmission of Education: the case of Senegal

Christelle Dumas et Sylvie Lambert

Décembre 2005

Docweb n° 0520

Les modes de transmission intergénérationnelle de l'éducation : le cas du Sénégal

Résumé: Cet article étudie l'effet de l'origine sociale des enfants sur leur éducation. L'analyse économétrique exploite la richesse d'une enquête originale conduite au Sénégal en 2003. Celle-ci fournit en effet des instruments permettant de corriger l'endogénéité des variables d'origine. L'effet estimé de l'éducation du père double une fois instrumenté et, de façon inattendue mais cohérente avec le contexte sénégalais, devient très supérieur à l'impact de l'éducation maternelle. Nous présentons aussi des résultats qui suggèrent que les variables d'origine ont autant d'influence après l'entrée à l'école qu'avant et que l'éducation des parents affecte la scolarisation des enfants à travers un changement des préférences parentales pour l'éducation.

Mots clefs : mobilité scolaire, demande d'éducation
Classification JEL : D12, I21, O12

Patterns of Intergenerational Transmission of Education: the case of Senegal

Abstract : This paper addresses the relationship between schooling and family background characteristics. The econometric analysis uses an original survey conducted in 2003 in Senegal that, uniquely, provides instruments to deal with the endogeneity of background variables. The estimated effect of father's education more than doubles when its endogeneity is accounted for and, unexpectedly, becomes much bigger than the impact of mother's education. We also present results suggesting that family background continues to have as much impact after entry to school as it does at younger ages, and that parental education affects children's schooling through its contribution to parental preferences.

Keywords : schooling mobility, education demand.
JEL Classification : D12, I21, O12

Patterns of Intergenerational Transmission of Education: the case of Senegal *

Christelle Dumas[†] Sylvie Lambert[‡]

19th December 2005

JEL Classification: D12, I21, O12.
Key-words: schooling mobility, education demand.

Abstract

This paper addresses the relationship between schooling and family background characteristics. The econometric analysis uses an original survey conducted in 2003 in Senegal that, uniquely, provides instruments to deal with the endogeneity of background variables. The estimated effect of father's education more than doubles when its endogeneity is accounted for and, unexpectedly, becomes much bigger than the impact of mother's education. We also present results suggesting that family background continues to have as much impact after entry to school as it does at younger ages, and that parental education affects children's schooling through its contribution to parental preferences.

*Acknowledgement: We are grateful to David Blau, Eve Caroli, Denis Cogneau, Geeta Kingdon, Eric Maurin, Jean-Marc Robin and Muriel Roger for their very useful comments on this and earlier versions of this paper.

[†]LSE-DESTIN. Houghton Street, London WC2 2AE, UK. Tel: (44-0) 207 955 6224. Fax: (44-0) 207 955 6844. E-mail: c.dumas@lse.ac.uk

[‡]LEA-INRA 48, bd Jourdan, 75014 Paris, France. Tel: (33-0) 1.43.13.63.60. Fax: (33-0) 1.43.13.63.62. E-mail: sylvie.lambert@ens.fr

Introduction

Inequality among households is a hotly topic debated in Latin American countries. In Africa, although the usual indicators point at inequality levels that are very comparable to those observed in Latin America, inequality is seldom studied. This paper aims to contribute to the analysis of income inequality and social mobility in an African country, Senegal, by scrutinizing the issue of intergenerational schooling mobility.

In fact, in terms of social dynamics, what matters is not only current inequality, but also social mobility. Schooling is usually thought of as one of the main mechanisms at play in intergenerational social mobility processes (Behrman, Birdsall, and Szekely, 1999; Behrman, 1999). Hence, the focus of this paper is on the relationship between children's schooling and family background. While most studies on education demand in Africa and elsewhere find that background variables have a strong impact on children's education, and particularly when it comes to maternal education, only a few of them convincingly demonstrate causal impacts. There is nonetheless an important body of articles and techniques addresses the question of endogeneity of parental background (see Blow, Goodman, Kaplan, Walker, and Windmeijer (2005) for a review of econometric approaches). Three approaches can be distinguished. The first one relies on the use of data that includes information on siblings and on their respective children so that, by differencing, family fixed effects can be removed (Blau, 1999; Behrman and Rosenzweig, 2002; Behrman and Wolfe, 1987). This type of estimate generally imposes fairly severe constraints on unobservables, since it assumes that differencing between siblings (or twins in the case of Behrman and Rosenzweig (2002)) can indeed remove a household effect on the outcomes of their respective children and that such a household fixed effect is the only source of endogeneity. The second way to deal with endogeneity of parental background is to take advantage of exogenous variations (natural or random experiments) in order to instrument it (Dahl and Lochner, 2005; Morris, Duncan, and Rodrigues, 2004). In particular, several studies use changes in the length of compulsory schooling to instrument parental education level (Chevalier, 2004; Oreopoulos, Page, and Stevens, 2003; Black, Devereux, and Salvanes, 2005). Unfortunately, this technique only allows estimates for the sub-sample directly affected by the reform. The last of the three possible approaches relies on the use of instruments specific to the household. This approach, which is the one we apply here, is the least common, in particular because of the difficulty of finding appropriate instruments. Furthermore, to our knowledge, no existing study deals with both the endogeneity of parental education and that of parental income. For example, **?** tackle the issue of the simultaneous determination of parents' and children's education, while Cogneau and Maurin (2001) use grandparents' education to instrument for wealth. Nevertheless, neglecting the endogeneity of one of the background variables may lead to biased estimates for the others, since they are very likely to be correlated. In our paper, we carry out an econometric analysis of the impact of background characteristics on schooling achievement in Senegal and deal with the issue of endogeneity of both parental education and wealth. We are in the unique position of being able to do this convincingly thanks to the use of original instruments for parental education and wealth, that are rarely available in household surveys. Notably, we have information on the environment in which parents lived when they were aged 10, and on their birth order among their siblings. We are therefore able to disentangle both effects and provide reliable causal estimates for each of them.

A further question of interest with respect to the role of economic and social background on school achievement is to try to assess the timing of its impact. More specifically, the question is to know whether the background plays a determining role on the schooling trajectory through the impact it has on the cognitive level of the child when s/he enters school, or whether it continues to affect the trajectory through its complementarity to school inputs. Since before the child enters school, his/her main influences are family ones, it is to be expected that divergence in trajectories due to differing backgrounds will happen then. Whether this divergence continues to happen during school years is an open question[1]. We will be able to contribute to answering it thanks to the fact that for a sub-sample of children, we have data on test scores at a very early stage. We can therefore look at the way the impact of background variables on schooling attainment changes when we control for initial cognitive level. This question of the timing of the influence of family background points to the role of educational institutions in facilitating mobility. In fact, institutions could aim to be able to evaluate children's skills as they enter school and push them, by providing adequate resources, to an educational level that reflects their initial promise, regardless of their social background.

Finally, Boudon (1973) argues that the impact of parental education on schooling outcomes is due to two mechanisms. On the one hand, the probability of reaching the required cognitive level to change grades is affected by parental background through its impact on the learning process of the child. On the other hand, at each transition in the education system, the choice made to continue on the track of longer general education rather than dropping out or starting professional training is determined by parental background through its impact on preferences and expectations[2]. In the last part of our paper, we try to assess whether a direct impact of preferences on achievement exists and whether we can distinguish this from the "productivity" effect. In order to gain some insight into this issue, we look at the impact of family background on the repetition of classes during the course of primary schooling. Since it seems reasonable to assume that a higher number of classes repeated is not a matter of preferences towards slow progress, the results of this exercise will be suggestive about the direct role of preferences in explaining the impact of unfavourable background

[1] Fryer and Levitt (2004) find that the gap between whites and blacks increases over time in the United States.

[2] Boudon also suggests that the relative weight of these two mechanisms changes during the course of progress through school levels. In fact, due to the selection process, those students who are handicapped by an unfavourable background in terms of their learning capacities are progressively eliminated and for the remaining ones, the role of parental preferences becomes more markedly important. We will not be able to explore this question with the data we have.

on lower schooling attainment.

In this paper, we deal with these three questions (impact of parental background, timing of this impact and existence of a preference effect). Our main results show that father's education is a more important determinant of schooling outcomes than mother's, once background variables are properly instrumented for. In fact, the estimated effect of the father's education more than doubles when its endogeneity is accounted for, while that of the mother's declines. Although this result might be contrary to a fairly large literature that tends to emphasize the role of the mother's education, it is in accordance with the results put forward by Behrman and Rosenzweig (2002). Wealth has a significant impact as well, which is attenuated when instrumented. The second set of results indicate that family background still matters as much when cognitive level at school entry is controlled for. This result suggests that educational institutions fail to facilitate schooling mobility in Senegal. Finally, we conclude from our last set of results that parental education affects attainment partly through shaping parental preferences while wealth seems to affect it mainly through an increase in productivity in the human capital production function.

Section 1 describes the data used in this study and presents some descriptive statistics. The model and the econometric method used for the ensuing analysis is developed in section 2, while section 3 presents and discusses the estimates obtained from a simple model of schooling mobility. Section 4 discusses the issues of timing and preferences mentioned above and the last section concludes.

1 Description of the data

1.1 The survey

The data used in this paper comes from an original survey entitled Éducation et Bien-Être des Ménages au Sénégal (EBMS)[3] conducted between April and June 2003. This survey covers a national sample of 1800 households. It provides information on household composition, household asset ownership and housing characteristics. At the individual level, information was collected on education, health, employment status and activities of every household member. The sample includes households of children who participated in an earlier survey conducted by the PASEC[4] from 1995 to 2000. In 1995, 120 schools were selected throughout the country. In each school, a class of second graders (CP) was randomly selected and 20 randomly chosen children in each of these classes were given cognitive tests at the beginning and end of the school year. Those children were aged 7 to 10. They were then surveyed every year throughout their

[3]This survey was designed by a team composed of Peter Glick, David Sahn, and Léopold Sarr (Cornell University, USA), and Christelle Dumas and Sylvie Lambert (LEA-INRA, France), and implemented in association with the Centre de Recherche en Économie Appliquée (Dakar, Senegal).

[4]Programme d'analyse des systèmes éducatifs des pays membres de la CONFEMEN. We would like to thank the PASEC and in particular Jean-Marc Bernard for his help in setting and conducting the EBMS survey.

primary schooling. The EBMS survey was designed so as to complement the PASEC survey, in order to obtain more information on the households, to gain some data on siblings and to include children who were not attending school. The cluster structure of the original PASEC survey was therefore maintained. We recovered on average 13 children per cluster (out of the 20 who participated) for 60 clusters from the original PASEC sample. We then increased the sample size in each cluster to 30 households by adding households drawn randomly from those who had children in the same age range as the PASEC children (14 to 17 years old in 2003) and residing in the catchment areas of the PASEC schools. As a result, the final sample is not a representative sample of the Senegalese population, since it includes children who have on average more schooling than the general population[5]. Hence, this survey cannot be used to compute descriptive statistics that are valid for the country as a whole. This, nevertheless, does not prevent it from being a perfectly suitable and rich database to study household behaviour with regard to education.

For our purpose in this paper, it is noteworthy that the EBMS data contains information on three generations of individuals (grandparents, parents, children), as well as a number of variables on parents' living conditions when they were ten years old. Regarding the first aspect, we know in particular the level of education of grandparents and of parents' siblings. Examples of parents' living conditions are: area of residence (rural/urban), infrastructure (primary school, lower and upper secondary school, health care providers) within 5 kilometers, housing characteristics, whether the parents were alive/healthy or not, number of siblings and birth order of the parent. For the subsample of children belonging to the PASEC dataset, we will use the fact that we have their results on tests in mathematics and French at the beginning of the second class of primary school. This is the earliest level for which such tests are available since before this, children cannot take written exams.

No information on income or consumption was collected in the EBMS survey. Rather, we chose to collect information on housing and durable goods, in order to construct a permanent wealth indicator using principal factor analysis. Appendix A shows the results of the analysis. Household ranking according to this indicator is stable when alternative subsets of the ownership and housing variables are used.

1.2 Some descriptive statistics on education and schooling mobility

In our sample, 18% of the individuals aged between 7 and 17 have never been to school despite the fact that school is compulsory from age 7 onwards. Only 69% of children in this age range are currently enrolled. As is frequently the case in African countries, these average numbers hide important differences between boys (of whom 84% have been to school) and girls (only 80%) and an

[5] The sample is representative of the Senegalese population in terms of religion, ethnic groups and demographic characteristics.

even greater discrepancy between urban and rural areas where those numbers amount respectively to 91% and 76% (see table 1). Table 2 shows that even for those who have some education, the level attained is fairly low: most of them did not complete their primary schooling. Obviously, this is very variable with age within this group since those younger than 13 cannot possibly have completed their primary education anyway. Regarding parental education, the fact that nearly 60% of the fathers and three quarters of the mothers of the children aged 7 to 17 have no schooling at all, testifies of the steep increase in schooling across time in Senegal. As the point of our paper is to examine the impact of parental background on schooling attainment, it is interesting, although not surprising, to notice that current enrollment rates increase markedly with the parents' education level. In fact, table 3 shows that there is a huge step up when comparing children whose parents have no schooling to those with at least some primary schooling and that it then increases regularly with the father's education (the profile with respect to the mother's is flatter). The same upward trend in the probability of having been to school appears when children are sorted by wealth quartile (see table 4).

A descriptive analysis based on the relevant mobility matrices (not reproduced here) also gives some sense of the relatively low schooling mobility that can be observed in Senegal[6]. This analysis that focuses on individuals aged 21 to 49 who are assumed to have completed their education, points to a few key facts. In terms of schooling mobility, origin matters mainly in the sense that the lowest schooling level is very "sticky". For example, if one considers two individuals, one whose father has no education at all and the other whose father

[6] The details of this descriptive analysis can be found in the working paper version of this article available on
http://www.inra.fr/Internet/Departements/ESR/UR/lea/documents/wp/wp0503.pdf.

Table 1: Education status, for children aged between 7 and 17

	Percentage
has been to school	82%
in rural areas	76%
in urban areas	91%
for boys	84%
for girls	80%
is currently enrolled	69%
in rural areas	65%
in urban areas	74%
for boys	71%
for girls	67%

Table 2: Levels of education, for children aged between 7 and 17, and for their parents

Code	Level	Child	Father	Mother
1	no schooling	18.05%	59.90%	73.09%
2	incomplete primary	52.01%	10.46%	10.38%
3	complete primary	14.70%	8.62%	7.82%
4	incomplete lower secondary	13.65%	5.99%	3.67%
5	complete lower secondary	1.11%	5.19%	3.20%
6	incomplete higher secondary	0.47%	3.16%	0.53%
7	complete higher secondary	0.01%	3.00%	0.98%
8	university	0.00%	3.66%	0.34%

Table 3: Proportion of children having been to school, given the education levels of the parents

Parental education	of the father	of the mother
no schooling	73%	77%
incomplete primary	90%	93%
complete primary	92%	93%
incomplete lower secondary	94%	98%
complete lower secondary	97%	96%
incomplete higher secondary	98%	100%
complete higher secondary	98%	97%
university	98%	95%

has been to primary school but didn't complete it, the situation where the first one doesn't get any education and the second some primary education is 6.97 times more likely to happen than the reverse. This is a very high reproduction coefficient (in the extreme case of perfect schooling mobility, reproduction coefficients would be equal to 1).

2 A simple model of intergenerational schooling mobility

To analyze schooling mobility, we conduct an econometric analysis of the impact of variables of origin on schooling achievement. The descriptive results presented in the previous section are not sufficient to assess the impact of parental background on schooling achievement for two reasons. First, the descriptive analysis

Table 4: Distribution of the wealth index and proportion of children having been to school, by wealth quartile

Wealth quartile	Value	Prop. of children
Min	-1.48	
		70%
25%	-0.84	
		76%
Median	-0.17	
		87%
75%	0.78	
		95%
Max	3.16	

is mostly bivariate, while, if only because parental education and parental wealth are correlated, a multivariate analysis is necessary. The descriptive statistics do not permit us to disentangle the effects of the various dimensions of family background. Second, those variables are likely to be endogenous to the schooling outcomes for reasons we detail later. To assess the impact that an economic policy such as monetary transfers or a literacy campaign for adults could have on education demand, we need to estimate the true causal impact of parental background on children's schooling. In this section, we discuss the various biases that could emerge from an estimation without instrumentation, the potential instruments and the estimation methods.

2.1 The model

The final schooling outcome of a child results from earlier schooling decisions and from the successes or failures met by the child during his/her school life. We chose not to model the decision process itself, nor the education production function, but rather to concentrate on a reduced form model where the role played by family background in determining schooling achievement is emphasized. Two aspects of family background are considered here: parental education and parental wealth. Since other observable variables that contribute to determining schooling outcomes are not central to the point we want to discuss here, they are introduced but not detailed in the model. Hence, we write the schooling outcome of child c in a household h, living in generation t as:

$$s_{cht} = \alpha y_{ht-1} + \beta s_{ht-1} + X_{cht}\Gamma + v_{ht} + w_{cht} \qquad (1)$$

where s_{cht} is a measure of schooling achievement of child c, y_{ht-1} is the income of his parents, s_{ht-1} is the schooling outcome of his parents, X_{cht} is a set of other determinants of the child's schooling and the last two terms are unobservable variables affecting the child's outcome. They reflect 2 kinds of effects:

- a household effect (v_{ht}), which is generation-dependent. This can encompass productivity differences in children's upbringings, differences in preferences towards schooling, and/or household specific shocks.
- a child specific effect (w_{cht}), which may reflect specific abilities or preferences of the child.

By definition, the decomposition of the unobservable between these two terms is such that:

$$v_{ht} \perp w_{cht} \quad \forall c, h, t \tag{2}$$

Parental education and wealth may be correlated with unobservable characteristics of the household, such as cognitive ability or preferences towards education, that affect schooling choices for the children. This is notably the case if some capacities or preferences are transmitted across generations. In order to discuss the difficulties that will need to be overcome to identify the above model, it is useful to fully specify the process that determines the educational achievement of a given child. This is done by adding two equations that describe schooling outcomes of the parents and household income or wealth. Parents' educational level is assumed to be determined in a similar way to children's schooling achievement. Household income is assumed to depend on parental schooling outcomes and some unobservables. Hence, the full model is:

$$\begin{aligned}
s_{ht-1} &= \alpha_{t-1} y_{ht-2} + \beta_{t-1} s_{ht-2} + X_{ht-1}\Gamma_{t-1} + v_{ht-1} + w_{ht-1} && (3) \\
y_{ht-1} &= \delta s_{ht-1} + X_{ht-1} B + \zeta_{ht-1} && (4) \\
s_{cht} &= \alpha_t y_{ht-1} + \beta_t s_{ht-1} + X_{cht}\Gamma_t + v_{ht} + w_{cht} && (5)
\end{aligned}$$

ζ_{ht-1} are unobservable variables affecting the income of the parents.

2.2 Potential endogeneity of parental characteristics

The choice of an estimating method for the above model depends on the assumptions made about the correlations between the various residuals. We discuss here potential correlations between residuals that give rise to serious endogeneity problems and that cannot be assumed away.

Transmission effect If preferences towards schooling are transmitted through generations, then these preferences determine on the one hand parents schooling and hence parents income and, on the other, children's schooling. These variables are thus jointly determined by the same unobservable variables. Formally speaking, this means that v_{ht-1} is correlated with v_{ht}.

Parent's ability If parents are academically more able than their siblings, they may also be more able to help their children in their studies and to support them in their schooling. This implies that w_{ht-1} may be correlated with v_{ht}. This is another transmission effect, but this one is specific to the abilities of a parent, contrary to v_{ht}, which is the same across a generation of the dynasty. This correlation also implies that parents' schooling and income are correlated with v_{ht}.

Parent's productivity The same applies if the unobservables affecting parental income (ζ_{ht-1}) are correlated with the household effect in the child schooling outcome equation (v_{ht}). The correlation between these two terms may arise from the impact of unobservables at the level of the household (e.g., possession of productive assets) or at the level of the individual (abilities).

Measurement error Another source of endogeneity is possible measurement errors on parental education and parental wealth. Such measurement errors would be captured in the error term w_{ht-1} and cause it to be correlated with v_{ht} (see Greene (1997), ch.9).

An additional point should be mentioned, relating to the potential bias introduced by the joint determination of time allocation decisions for members of the household. In fact, for various reasons, labour supply decisions of the different household members are likely to be jointly determined and the same could be true for decisions regarding school. In such a case, if in the model y_{ht-1} is measured by income (either current consumption or income flows), since labour supply decisions directly affect income, a naive estimation of the impact of income on child school attendance may be biased. In our case though, this issue will be bypassed since we use a measure of permanent income.

2.3 Choice of instruments

The Ebms data contains some information on parents' living conditions when they were ten years old, such as: area of residence (rural/urban), infrastructure (primary school, lower and upper secondary school, health care providers) within 5 kilometers, housing characteristics, whether their parents were alive/healthy or not, number of siblings and ranking of the parent. We also know the level of education of grandparents and of parents' siblings. Some of these variables may not be valid instruments and we discuss them below.

First, some variables may be correlated with the family's unobserved heterogeneity component (v_{ht}): education of grand-parents or of parents' siblings are a priori correlated with v_{ht-1}, itself potentially correlated with v_{ht}. Thus they are not valid instruments. In the same way, the housing characteristics, as proxies of household wealth for the previous generation, that is partially determined by the level of grandparents' education, may not be exogenous either.

The number of siblings may also reflect some preferences towards education, in a framework where parents trade-off "quantity for quality" of children.

We could also think of using differences in the level of education between grand-parents and parents or between parents and their siblings. Since the household effect is not a fixed effect constant across generations ($v_{ht} \neq v_{ht-1}$), it will not be eliminated by the differencing between parents and grand-parents education[7]. The difference between parents and their siblings will indeed eliminate a household effect (v_{ht-1}), but the individual component (w_{ht-1}) that reflects the differences in capacities or preferences will remain. The endogeneity problem will then persist since, as discussed in the paragraph about the transmission of parental ability, w_{ht-1} and v_{ht} are possibly correlated as well. For these reasons, we did not use any of the previously listed variables as instruments. Those we used and the conditions for their validity are now discussed.

The area of residence and the presence of infrastructures can be considered exogenous, if households do not move in order to allow their children to go to school. If this is not true, these variables are not valid instruments. In fact, we will be able to test overidentification restrictions for some instruments and we will see that exogeneity of area of residence is rejected. Infrastructure variables fare better. For each parent, we use dummies indicating the presence of primary, lower secondary and upper secondary schools in their neighbourhood when they were aged 10, and whether there was a health care centre in the community.

We also use two variables related to parental health. For each parent, we know whether, when they were aged 10, their own parents were alive and whether they were suffering from serious illnesses that prevented them from working normally. Parental health is probably not correlated with preferences and abilities, since it is mainly explained by age and shocks. This being said, one could argue that the richest households have better access to health care and that they have a greater probability of being alive and in good health than the others. Nevertheless, in the Senegalese context, the hypothesis that the probability of being dead or seriously ill before one's child reaches age 10 is not correlated with wealth either is arguably sensible, given the very wide range of possible parent-children age difference[8].

The last set of instruments we have chosen describe the birth order of the parent. It can be argued that the only exogenous position in the birth order is that of the eldest child, while being the youngest may not be (parents have decided to stop having children after this one). Hence, we used two ranking variables that allow us to distinguish the elder son and the elder daughter: "no older brother" and "no older sister".

[7]In practice, even if we assume that household effects are fixed, the generally low levels of education for grandparents prevents us from using this method.

[8]When considering the children of the household head, the average age difference between a household head and his children is 42 years, but it goes from 23 years for the first percentile to 69 years at the 99th percentile with a standard error of 10. The correlation between the wealth index and the children-father age difference is positive but very small (0.12).

2.4 Empirical specification

The model as presented in section 2.1 can not be estimated directly since, as discussed above, grand-parents' wealth and education are also potentially endogenous. Hence we estimate equation (5) using the IV method, the "structural" equations for parents education and wealth being replaced by the corresponding reduced form equations.

The complete model is then:

$$
\begin{aligned}
s_{ht-1} &= X_{ht-1}\tilde{\Gamma}_{t-1} + \tilde{v}_{ht-1} + w_{ht-1} && (6)\\
y_{ht-1} &= X_{ht-1}\tilde{B} + \tilde{\zeta}_{ht-1} && (7)\\
s_{cht} &= \alpha_t y_{ht-1} + \beta_t s_{ht-1} + X_{cht}\Gamma_t + v_{ht} + w_{cht} && (8)
\end{aligned}
$$

The identification hypotheses described in detail in the previous section are still valid and permit us to obtain consistent estimates for the above model[9].

The variables of interest also need to be specified. In particular, there are various ways to measure schooling. The first decision that parents take is whether to send their children to school. The first possible outcome is thus "to have been to school or not", which is measured by our "enrollment" variable. For children who have never been enrolled at the time of the survey, but will enter school in the future, the information regarding enrollment is censored. To avoid the difficulties linked to censoring of this variable, we restrict the sample of interest to children between the ages of 10 and 21 and assume that children who have not yet entered school when they reach age 10 will never do so[10]. The second decision concerns the final level of schooling that children attain. This is measured by an ordinal variable ("level attained") that takes 8 different values corresponding to the education levels given in table 2.

For this later variable, censoring is more of an issue. Indeed, most children under age 21 are still attending school. We thus do not know the final level of schooling they will achieve. Hence, for children still enrolled in school at the time of the survey, we only know that their final schooling outcome will be greater than the observed level at the time of the survey. s_{cht} itself is determined as follows:

$$
s^*_{cht} = \alpha_t y_{ht-1} + \beta_t s_{ht-1} + X_{cht}\Gamma_t + v_{ht} + w_{cht}
$$

$$
s_{cht} \begin{cases} = s^*_{cht} & \text{if child } c \text{ has completed his/her schooling} \\ < s^*_{cht} & \text{if child } c \text{ is still enrolled} \end{cases}
$$

[9]Let's consider the following orthogonal decompositions:

$$
\begin{aligned}
s_{ht-2} &= X_{ht-1}M + \epsilon_{ht-1}\\
y_{ht-2} &= X_{ht-1}N + \eta_{ht-1}\\
\text{with} &\quad X_{ht-1} \perp \epsilon, \eta,
\end{aligned}
$$

where the X_{ht-1} variables represent the set of instruments we retained. Inserting those equations in equation (3), we obtain the model, with $\tilde{v}_{ht-1} = v_{ht-1} + \alpha_{t-1}\epsilon_{ht-1} + \beta_{t-1}\eta_{ht-1}$ and $\tilde{\Gamma}_{t-1} = \Gamma_{t-1} + \alpha_{t-1}M + \beta_{t-1}N$. Hence, $X_{ht-1} \perp \tilde{v}_{ht-1} \Leftrightarrow X_{ht-1} \perp v_{ht-1}$. The same computation can be done for the wealth equation.

[10]In our sample, about 3% of the children who went to school entered at age 10 or later.

We will henceforth assume that the censoring is exogenous and only due to the age of the child, relative to the level s/he wants to reach. This is equivalent to writing a static approximation of a dynamic decision process[11].

There are also various variables to choose from to describe parental background. We can either use the parents' enrollment variable, or the level reached (treated as continuous or as a set of dummies). A specification search showed that, without any instrumentation, the use of these various variables, in the enrollment as well as in the level estimation, does not seem to make much difference: in the level equation, the use of the variable of parents' educational level just slightly enhances the estimation. Therefore, and also for reasons of convenience, we chose in what follows to use only the continuous variable of parental education level to measure parental education. Nevertheless, it should be emphasized straight away that information on level attained might be affected by measurement error, notably because the question was not always answered by the respondents.

The wealth indicator is a continuous variable, normalized to be centred and of variance equal to 1. There is no difference in terms of explanatory power between using it as a continuous indicator or by quartiles. Sample statistics for all the variables used in the estimation are given in appendix B.

To conduct the estimation of censored dependant variables when some explanatory variables are suspected of endogeneity, we follow Smith and Blundell (1986); Rivers and Vuong (1988); Wooldridge (2002), and test exogeneity in a two-step estimation by introducing residuals of the instrumentation equation in the equation of interest. The significance of the coefficient associated to the residual informs whether exogeneity of the variable can be rejected (if coefficient significantly different from zero) or not.

3 Estimated effects of family background

3.1 Instrumentation

First note that the instrumental variables we chose to use are only available for children who live with their parents. Hence, the instrumentation equation is run on the about 2610 children aged 10 to 21 who live with both their parents. The whole sample includes 6884 individuals aged 10 to 21. Such selection based on co-residence with both parents naturally raises a suspicion of endogenous selection. Nevertheless, it should first be noted that these two samples hardly differ in terms of the observables (see table 12, in appendix B). Further, we also estimated our model on the sample of children living with their father (whether their mother was present or not) introducing only the education of the father and did the same for children living with their mother, using the mother's education. The results obtained on those sub-samples are very similar to those presented below. In total, we think the sample selection does not affect

[11]Besides, attempts at estimating a dynamic model showed that there was not much information to be gained by adding this technical complication.

the estimates. Nonetheless, throughout the rest of the paper, all estimations will be conducted for the sake of comparison on both the full and the restricted sample without instrumenting, before being conducted on this restricted sample with instrumentation.

The instrumentation is shown in appendix C. The instrumental variables alone explain respectively 29%, 22% and 35% or the variance of father's education, mother's education and wealth. Given this explanatory power, we can consider that the instruments, taken as a whole, are not weak.

The instrumental regressions for the three potentially endogenous variables use the whole set of instruments. Notably, the instrumental regression concerning the father's education will not only include variables describing his surroundings at the age of 10, but also that of the mother of the child considered (his wife). In the following comments, we will only discuss the impact of background variables concerning one's own side. In other words, to explain the father's education, we concentrate on the impact of his own background, although sometimes, it will be found that variables regarding his wife's background seem to have an impact. This may be partly due to endogamous marriages. Note also that area of residence of the parent at age 10 is excluded from the final set of instruments used, since the test of over-identifying restrictions rejected it (see section 3.4).

These instrumental regressions show that having grown up in an environment with more infrastructure enhances education, as shown by the positive coefficients associated with urban areas and the presence of primary schools. The results on the presence of a lower secondary school and of a health care facilities are not systematic but generally of the expected sign. Individuals whose father died before they reach age 10 are less educated, but we do not find any effect when the mother is dead. We find some evidence that a parental illness may favor schooling for children of the other gender. Being the oldest boy is also detrimental to education. The estimate of parental wealth gives some similar results, except that the death of the mother seems to be detrimental to wealth.

3.2 Enrollment

A probit estimation of child enrollment as a function of parental background provides the first set of results. The estimated coefficients are given in table 5, with and without instrumentation. A test of exogeneity of parental background gives the following results: father's enrollment and wealth are rejected as exogenous variables, however we cannot reject that maternal enrollment is exogenous (see coefficients for the corresponding residuals in the third column).

Before instrumentation, we find a significantly positive impact of parental education, wealth, being a boy, living in a small household and having older sisters. The impact of control variables seems to remain stable (except for the household size) when we instrument the background variables.

We compute the marginal effects for the three variables of interest, at the mean and for an increase by 1 in education and an increase by 0.7 in wealth, which broadly corresponds to a shift of one quartile. The confidence intervals

Table 5: Child enrollment if $10 \leq \text{age} \leq 21$

	Whole sample	Reduced sample	
Father's education	0.270**	0.238**	0.605**
	(0.029)	(0.050)	(0.136)
Mother's education	0.213**	0.134*	-0.070
	(0.042)	(0.065)	(0.212)
Wealth	0.287**	0.402**	-0.181
	(0.044)	(0.085)	(0.305)
Rural	-0.075	-0.045	-0.441+
	(0.071)	(0.130)	(0.259)
Boy	0.411**	0.288**	0.280**
	(0.043)	(0.069)	(0.073)
Household size	-0.017**	-0.024*	-0.003
	(0.006)	(0.009)	(0.016)
No older brother	-0.167**	-0.054	-0.082
	(0.044)	(0.071)	(0.075)
No older sister	-0.128**	-0.162*	-0.223**
	(0.046)	(0.072)	(0.082)
Father edu. residual			-0.402**
			(0.133)
Mother edu. residual			0.201
			(0.219)
Wealth residual			0.578+
			(0.309)
Constant	0.459**	0.830**	0.282
	(0.112)	(0.187)	(0.385)
Observations	6884	2636	2636
Pseudo-R^2	0.16	0.16	0.17

Note: Estimation performed by probit. Coefficients reported are probit coefficients and not marginal effects. Standard errors in parentheses. * significant at 5%; ** significant at 1%.

are computed by bootstrap and are given for the level of 5%. In table 6, we provide three marginal effects: the first without any instrumentation, the second with all variables instrumented and the last with only the endogenous variables (here, father education and wealth) instrumented.

We find that instrumenting parental background leads us to reevaluate the impact of this background on enrollment. The marginal effect of increasing father's education by 1 increases the probability of enrollment by 3% when not instrumented and by 7% when instrumented. The impact remains quite low,

Table 6: Marginal effects on enrollment

Instrumented variables	None	All	Father education and wealth
Father's education	0.035**	0.077**	0.074**
	(0.005)	(0.015)	(0.015)
Mother's education	0.021*	-0.013	0.020*
	(0.009)	(0.045)	(0.009)
Wealth	0.040**	-0.021	-0.035
	(0.007)	(0.036)	(0.039)

Note: ** (*,+) means respectively that the coefficient is significant at the 1% (5%,10%) level. Standard errors are computed by bootstrap with 200 replications. Marginal effects are given for the mean child, for an increase by 1 in education and for a shift of one quantile in wealth.

but this is partly due to the fact that most children enter school. It is much larger though than the impact of wealth, which is not significantly different from zero, once instrumented. Comparatively, since we do not reject the exogeneity of mother's education, we keep a marginal impact of 2%.

Those results therefore suggest a much stronger impact of father's education than that of the mother's. This finding might seem surprising in light of the previous literature on education demand in Africa and elsewhere. In this literature, mechanisms that are supposed to explain the predominant role of mother's education rely on the fact that mothers dedicate a larger share of their resources to education thereby improving the productivity of the human capital production function. As we will see in section 4.2, in the context under scrutiny here, parental education influences children's schooling mainly through another channel (through a direct preference effect) and this could explain why fathers, acting as the main household decision makers, have a dominant role here.

We found a negative endogeneity bias on father's schooling and a positive one on wealth. Since, as mentioned earlier, information on education is often not provided by the parent him/herself, it is liable to potential measurement error that could explain the negative bias. Whereas, because of the way the wealth variable is constructed, it is potentially much less subject to measurement error. The positive bias on the coefficient of the wealth variable is consistent with the fact that some unobservable variables positively correlated with both parental wealth and children achievement are now accounted for. Recall nevertheless that the instruments for wealth are the same as those used for parental education and they mainly capture the environment of the parents when they were aged 10. This raises the suspicion that the apparent negative correlation between the two could arise from the exploitation as a variation of wealth of what is in fact attributable to education. Nevertheless, since wealth is better predicted than either mother or father education, this explanation can be disregarded. Finally, although the directions of the observed endogeneity biases might still seem surprising, these results are stable across the whole range of specifications we experimented with (differing by either the set of instruments, the set of control variables or the sample).

3.3 Level

We provide here an estimation of the final level of education attained as a function of parental background and the other usual controls. As explained in section 2.4, information about total schooling (final level of education attained) is right-censored for children who are still enrolled. We therefore perform a tobit estimation by maximum likelihood, taking censoring into account.

By looking at the tests of exogeneity provided by the third column of table 7, we conclude that we fail to reject the exogeneity of the 3 variables. Nevertheless, the same pattern as before emerges: the estimated impact of father's education increases, while the two others decrease and become insignificantly different from zero.

The same controls as before are significant and of the same sign. An increase by 1 in father education raises by 0.28 the education level of the child (i.e. by 11%), while the same increase in the mother education has two third of this impact (0.18). An increase by one quartile in wealth (ie, an increase by 0.7) augments by 0.36 (14%) the level attained[12].

Even if the impacts remain quite small, they correspond to real differences of chances for two children whose backgrounds differ.

3.4 Test of over-identifying restrictions

To conclude this analysis, we conduct a test of over-identifying restrictions for the instruments we see as potentially the most fragile among those we retained. Further, we also check whether the instruments that we had rejected on theoretical grounds are indeed rejected on an empirical basis.

As discussed earlier, the confidence we have in the theoretical validity of the instruments we use differs according to the instrument. Being the eldest is the variable that is least likely to be endogenous. Then, among the remaining instruments, we trust the exogeneity of the health of the grand-parents more (because of the huge variance in the children-parent age difference) than that of the presence of infrastructure and of the area of residence (rural/urban). We will therefore use the identifying restriction that being the eldest child and grand-parent health are exogenous to test the conditional validity of the remaining instruments. We introduce these variables into the equation of interest (as well as in the instrumentation equation) and check whether their coefficients are significantly different from zero or not. We do this only in the schooling model, since in the attainment model we could not reject the exogeneity of the parental background variables.

[12] The marginal effects are directly given by the estimation of the tobit.

Table 7: Final education level if $10 \leq \text{age} \leq 21$

	Whole sample	Reduced sample	
Father's education	0.280**	0.283**	0.477**
	(0.024)	(0.046)	(0.177)
Mother's education	0.228**	0.179**	0.085
	(0.035)	(0.066)	(0.288)
Wealth	0.452**	0.527**	0.231
	(0.055)	(0.091)	(0.446)
Rural	0.154	0.199	0.017
	(0.097)	(0.157)	(0.360)
Boy	0.580**	0.320**	0.314**
	(0.054)	(0.087)	(0.097)
Household size	-0.025**	-0.037**	-0.026
	(0.007)	(0.010)	(0.020)
No older brother	-0.050	0.186*	0.173+
	(0.057)	(0.092)	(0.095)
No older sister	-0.121*	-0.217*	-0.247*
	(0.061)	(0.095)	(0.114)
Father edu. residual			-0.208
			(0.179)
Mother edu. residual			0.091
			(0.292)
Wealth residual			0.287
			(0.451)
Constant	2.691**	3.006**	2.679**
	(0.132)	(0.211)	(0.539)
Observations	6830	2613	2613
Pseudo-R^2	0.07	0.08	0.08

Note: Estimation performed by maximum likelihood. ** (*, +) means that the coefficient is significantly different from 0 at the 1% (5%, 10%) level.

The results are given in appendix D. The exogeneity of the rural/urban variable is rejected, but not that of the presence of infrastructure. Furthermore, the empirical over-identification test on the variables that we rejected on theoretical grounds (grandparents' and siblings' education, parent-grandparent or parent-sibling differences in education, grandparent housing conditions and number of siblings) confirm our prior assumptions: all are rejected at the 5% level, except for the number of siblings, which is rejected only at the 10% level.

4 Further insights into the role of parental education

We have now established that the causal link between parental background and children's education is quite strong, notably when it comes to father's education. Taking endogeneity of parental background into account led us to revise downward the estimates of schooling mobility (or equivalently, to revise upward the causal impact of parents' education). This section tries to refine this analysis of schooling mobility by looking at two issues of major importance. The first is that of the timing of mobility and the second is the question mentioned before of the role of preferences versus productivity in explaining the impact of parental education.

4.1 Timing

Regarding timing, the question is to know whether background variables (parental wealth and education) play a role mainly before the child enters school or also afterwards. In fact, before the child enters school, his or her main influences are family ones and it is to be expected that individual trajectories will diverge accordingly. Once in school, the child is subject to other influences. In particular, the school as an institution could place some weight in equalizing opportunities among children with similar cognitive levels as they enter school. Although it is likely that school will not be able to counterbalance disadvantages accumulated before school entry, it could try to compensate for differences in backgrounds by distributing resources so as to equalize chances of reaching a given final level of education for children who starting school with the same cognitive level.

In order to study this point, we use one of the unique features of our survey, namely the fact that, for a sub-sample of children, we have data on their test scores from very early in their schooling. The idea is the following. Parental background is likely to influence schooling achievement more or less continuously through a child's life. We can assume that what takes place between birth and entry to school translates into a given test score. Hence, we can use the scores on these tests as control variables in the equation for schooling attainment. What will be measured with such an estimate is the residual impact of parental background on achievement, given that a certain level of cognitive knowledge has already been attained. If the school is able to promote children according to their observed skills when they enter school in a way that does not depend on economic and social origin, this impact should be reduced compared to what is obtained in an estimate that does not control for early scores. The way in which social and economic background affects schooling achievement, controlling for test scores, is therefore of interest.

Table 8 shows the results of this exercise. The first column shows the results of a simple censored regression run on the whole sample of children aged 10 to 21. The last column gives the results of the estimation using scores. It is carried out for children belonging to the PASEC sample. In order to facilitate the

Table 8: Education level attained if $10 \leq \text{age} \leq 21$

	10-21	14-17	14-17 w/ schooling	Pasec children	
Father's education	0.280**	0.239**	0.124**	0.139**	0.131*
	(0.018)	(0.027)	(0.026)	(0.053)	(0.052)
Mother's education	0.229**	0.184**	0.137**	0.112	0.080
	(0.028)	(0.041)	(0.040)	(0.073)	(0.072)
Wealth	0.452**	0.513**	0.334**	0.159	0.117
	(0.038)	(0.059)	(0.062)	(0.112)	(0.110)
Rural	0.157*	0.136	0.362**	-0.031	-0.050
	(0.069)	(0.107)	(0.114)	(0.195)	(0.191)
Boy	0.580**	0.537**	0.297**	-0.180	-0.261
	(0.051)	(0.080)	(0.088)	(0.161)	(0.159)
Household size	-0.025**	-0.010	-0.001	0.014	0.012
	(0.004)	(0.007)	(0.008)	(0.015)	(0.015)
No older brother	-0.050	-0.057	0.075	-0.124	-0.115
	(0.053)	(0.083)	(0.090)	(0.157)	(0.154)
No older sister	-0.121*	-0.066	0.011	0.276+	0.230
	(0.055)	(0.087)	(0.093)	(0.163)	(0.159)
French score					0.008
					(0.007)
Math score					0.017**
					(0.006)
Constant	2.691**	2.882**	3.522**	3.453**	2.623**
	(0.098)	(0.153)	(0.166)	(0.289)	(0.371)
Observations	6830	2986	2535	616	616
Pseudo-R^2	0.07	0.06	0.03	0.03	0.04

Note: Estimation performed by maximum likelihood. Standard errors in parentheses. ** (*, $^+$) means that the coefficient is significantly different from 0 at the 1% (5%, 10%) level.

comparison between these two sets of results, we carry out the same regression on different samples. The last column but one gives the estimate on the PASEC sample without controlling for the scores. The table also includes the same regression (without scores) for a sample of children comparable to the PASEC children (aged 14 to 17, having attended school) and for the sample of all the children aged 14 to 17.

Reducing the original sample to children aged 14 to 17 does not affect the results very much. By contrast, for children who have been enrolled in school, the residual impact of parental education and wealth is markedly reduced.

On the PASEC sample, the impact of wealth becomes insignificant, while results concerning the two variables of parental education are quite similar to what is obtained for the sample of 14-17 year olds who have attended school. We cannot explain this change in the impact of wealth and will therefore not elaborate on this result. Hence, the only background variable that has a significant impact in this sample is the father's education. Note that the corresponding coefficient is likely to be unbiased, even though parental education is not instrumented for, since, as shown in the previous section, exogeneity of background variables cannot be rejected in the achievement equation.

When controlling for scores, the striking result is that the estimated impact of wealth or parental education does not change significantly. The impact of one more unit of father's schooling is not significantly different when adding the scores as controls for unobservables and it remains significantly greater than zero. This is the residual impact of background characteristics, once the child is enrolled in school and his/her capacities are observed, and it is very similar to the impact over his or her whole life time. Hence, the impact of father's education remains the same through the school life of the child as it is before the child enters school. To put it differently, the difference of achievement between two children enrolled in school with different economic and social backgrounds is not reduced by the fact that they have the same cognitive skills at school entry as measured by the test scores at the beginning of their primary education.

4.2 Preferences or productivity in the human capital production process

As explained earlier, unobserved family preferences regarding education could be a source of bias in the estimation of the impact of parental education if not accounted for. More specifically, it is the transmissible nature of preferences that raises the issue of endogeneity. In fact, the preferences that we want to get rid of when trying to correct the estimate from the endogeneity bias are dynastic preferences towards education that exist *before* the realization of parental education. In other words, if parents have inherited from their own parents preferences towards education that explain both their own education and that of the child, estimates of the impact of parental education on the child's schooling achievement will be biased. On the other hand, preferences towards school that could emerge *from* the schooling of the parents are part of the impact of parental education we want to measure.

In fact, a question of interest in explaining the effect of parental education on schooling achievement is the role taken by those preferences towards education shaped by education itself. The question is that of the nature of schooling mobility. Is mobility the result of a dynastic accumulation of skills in terms of production of human capital, or does it result mainly from a gradual shaping of preferences towards education? Boudon (1973) discusses the fact that both mechanisms are at play. First, more educated parents become more efficient at helping their children through their schooling years thanks to a better intellectual and material environment. Second, when a choice needs to be made between different trajectories (dropping out or repeating a class; entering secondary school or professional training; etc.), parents with different educational levels will have different preferences or/and anticipations regarding these alternatives. More educated parents will be more likely to choose a trajectory leading to greater formal education. This is a direct effect of preferences on schooling outcome, as opposed to an indirect effect that would go through the fact that parental preferences induce a greater effort from the child or greater investment in schooling inputs such as books and hence a greater productivity in the human capital production function. It can also be called a "pure" preference effect.

Obviously, the implications of these two mechanisms differ. If productivity in the human capital production function is at stake, public education policy might be able to act by offering compensating resources. If the reason why children from unfavourable backgrounds drop out early is because of their households' preferences, then adequate public policy might be less straightforward to implement.

A way of gaining insight into this question would be to compare the scores obtained by children, or the yearly increase in the scores. If more educated and wealthier households are more efficient at producing human capital with a given amount of school input, their children should improve their scores faster at a given level. Conversely, if they push their children to go as far as possible but are not more efficient, these children should not fare better than the others. We will not use such data, but the idea remains the same. In Senegal, a lot of children are asked to retake a class during their primary school because the cognitive level they attained is deemed insufficient to follow the next grade (65% of people aged under 21 and having completed their primary schooling have repeated at least one class). It is quite unlikely that some parents prefer to have their children progress slowly, rather than follow the normal pattern to reach a given level. We thus want to compare the number of grade repetitions for children having completed the primary cycle[13] for different parental backgrounds[14] . The assumption we make is that the estimated coefficient in this model will mainly reflect the productivity impact of parental schooling.

Table 9 gives the results of the estimation of the number of grade repetitions during primary schooling for children who completed this cycle. We performed an ordered probit estimation, the dependant variable being the number of classes repeated. This variable takes values from 0 (for children who never repeated a class during their primary schooling) to 3 (for those who repeated 3 times or more)[15]. The ordered probit estimates of the model without any instrumentation gives a significant impact for the wealth and the father's education variables. It is noticeable that contrary to what was found in the

[13]This could be done for any level attained.

[14]We also analyzed repetitions at levels where no one dropped out in the subsample of interest (5 first levels of schooling for children who completed primary schooling). The rationale behind this is that, if poor cognitive achievement leaves the choice between repeating a class or dropping out, strong parental preferences toward education will increase the probability of repetition. Hence, considering only levels where no child dropped out limits possible interferences of preferences with poor cognitive achievement in explaining repetition. Results obtained with this exercise were very similar to those presented here

[15]35% of the children who completed their primary schooling did not repeat any class, while 40% repeated one class, 18% repeated two classes, 6% repeated three times or more.

Table 9: Estimation of the number of grades repeated for children having completed their primary schooling

	Whole sample	Restricted sample	
Father's education	-0.024*	-0.054*	0.004
	(0.012)	(0.022)	(0.100)
Mother's education	-0.016	0.004	-0.136
	(0.017)	(0.033)	(0.135)
Wealth	-0.262**	-0.293**	-0.116
	(0.036)	(0.066)	(0.269)
Rural	-0.047	-0.119	0.055
	(0.067)	(0.118)	(0.271)
Boy	-0.063	-0.096	-0.044
	(0.044)	(0.073)	(0.087)
Household size	0.008+	0.015+	0.007
	(0.005)	(0.008)	(0.014)
No older brother	0.065	0.098	0.060
	(0.044)	(0.077)	(0.078)
No older sister	-0.013	0.032	0.071
	(0.046)	(0.077)	(0.083)
Father edu residual			-0.071
			(0.102)
Mother edu residual			0.145
			(0.137)
Wealth residual			-0.206
			(0.270)
Nb. obs.	2799	974	974
Pseudo-R^2	0.02	0.04	0.04

Note: Estimation performed by ordered probit (2 stages for instrumentation). ** (*, +) means that the coefficient is significantly different from 0 at the 1% (5%, 10%) level.

school achievement equation, parental education does not seem to be a strong determinant of speed of progression through the education system. This result is confirmed by the two regressions that deal with endogeneity of background variables. In the instrumented regression (which indicates that exogeneity of the background variables cannot be rejected), the impact of all the background variables is wiped out. These results are consistent with a situation in which wealth, hence material living conditions, affects progress through its impact on the human capital production process (also through the fact that it allows children to concentrate on their schooling and not divert their time to income earning activities for example), while parental education shapes parental preferences with respect to education and hence induces parents to decide in favor of longer schooling. This gives some support to Boudon's intuition. It nevertheless suggests there is room for substituting public inputs to private ones in the human capital production function.

5 Conclusion

This paper examines the relationship between economic and social backgrounds and schooling attainment in Senegal. It focusses more specifically on intergenerational schooling mobility.

The data we use, an original survey conducted in 2003 and to which we contributed, provides instruments that allows us to deal with the issue of endogeneity of background variables. We can therefore exhibit causal relationships between background variables and educational attainment, a result which is rarely found in the existing literature due to the lack of appropriate instruments in most available datasets. Instrumenting correctly proved important since the estimated effect of father's education on enrollment more than doubles when its endogeneity is accounted for.

The results underline that children are not on an equal footing with regard to their chances of ever going to school and of attaining a given grade. Origin matters, in particular regarding the probability of enrolling in school at all. Econometrically, we find a positive effect on enrollment and level attained of having an educated father and also positive effects of mother's education and wealth on level attained. This means that children who differ in their household characteristics do not have the same chances of going to school and then reaching a given level of education.

Interestingly, we found that father's education matters much more than mother's education. Although the literature on this issue generally stresses the crucial role of mothers in determining their children's outcomes, this result is not that surprising in the Senegalese context. In fact, the likelihood of this result is supported by the last set of results that suggests that the direct effect of preferences is an important part of the impact of parental education on children's achievement. This means that the effect of parental education does not mainly go through the human capital production function, as would be the case if it translated into additional time dedicated to children's homework or additional educational expenditure for example, but it directly affects decisions taken with regard to school entry or continuation. Hence, since in Senegal, the decision to keep children in or out of the formal schooling system is very much in a father's hands, it is not really surprising that the influence of the mother's education is somewhat limited.

We discuss the fact that an unfavorable background affects the trajectory of children as much once s/he has entered school as it does before entry, when the only influences a child can receive are from his/her family. This is in particular the case for father's education, which influences outcomes in the same way whether we control for initial cognitive level or not. This suggests that there is scope for the school itself to improve equality of opportunities, even conditional on initial cognitive level.

Finally, our results also suggest that the detrimental impact of low parental education is channeled in part by parental preferences towards education, while poverty prevents families from providing their children with an environment favorable to learning. Hence, there is room for economic policies that would reduce such inequalities. In the short run, redistributive policies will be an effective tool. In the long run, the impact of those policies is likely to be persistent thanks to the role of education in shaping preferences.

References

Behrman, J., 1999. New markets, new opportunities? Economic and social mobility in a changing world, chap. Social mobility: Concepts and measurement in Latin America and the Caribbean. The Brookings Institution and the Carnegie Endowment for International Peace, Washington, D.C.

Behrman, J., Birdsall, N., Szekely, M., 1999. New markets, new opportunities? Economic and social mobility in a changing world, chap. Intergenerational mobility in Latin America: deeper markets and better schools make a difference. The Brookings Institution and the Carnegie Endowment for International Peace, Washington, D.C.

Behrman, J., Rosenzweig, M., 2002. Does increasing women's schooling raise the schooling of the next generation. American Economic Review 92, 323–334.

Behrman, J., Wolfe, B., 1987. Investments in schooling in two generations in prerevolutionary Nicaragua: the roles of family background and school supply. Journal of Development Economics 27, 395–419.

Black, S., Devereux, P., Salvanes, K., 2005. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. American Economic Review 95, 437–449.

Blau, D., 1999. The effect of income on child development. The Review of Economics and Statistics 81, 261–276.

Blow, L., Goodman, A., Kaplan, G., Walker, I., Windmeijer, F., 2005. How important is income in determining outcomes? a methodology review of econometric approaches.

Boudon, R., 1973. L'inégalité des chances. A. Colin.

Chevalier, A., 2004. Parental education and child's education: a natural experiment. Discussion Paper 1153, IZA.

Cogneau, D., Maurin, E., 2001. Parental income and school attendance in a low-income country: A semiparametric analysis. Working Paper 16, DIAL.

Dahl, G., Lochner, L., 2005. The impact of family income on child achievment. Working Paper Series 11279, NBER.

Fryer, R., Levitt, S., 2004. Understanding the black-white test score gap in the first two years of school. The Review of Economics and Statistics 86, 447–464.

Greene, W., 1997. Econometric Analysis -3rd edition. Prentice Hall.

Morris, P., Duncan, G., Rodrigues, C., 2004. Does money really matter? estimating impacts of family income on children's achievment with data from random-assignment experiments.

Oreopoulos, P., Page, M., Stevens, A., 2003. Does human capital transfer from parent to child? the intergenerational effects of compulsory schooling.

Rivers, D., Vuong, Q., 1988. Limited information estimators and exogeneity tests for simultaneous probit models. Journal of Econometrics 39, 347–366.

Smith, R., Blundell, R., 1986. An exogeneity test for a simultaneous equation tobit model with an application to labor supply. Econometrica 54, 679–685.

Wooldridge, J., 2002. Econometric Analysis of cross section and panel data. MIT Press.

A Construction of the wealth indicator

Table 10 gives the results pertaining to the creation of the wealth indicator.

Table 10: Wealth indicator: weights given to the variables

Durable goods		Housing	
Cooker	0.310	Walls in concrete	ref
Fridge, freezer	0.691	Walls in bricks	-0.140
Coal oven, "gaz butane"	0.441	Walls in earth	-0.356
Electric, gas, micro-wave oven	0.166	Walls in bamboo, canvas, other	-0.068
Sewing machine	0.260	Walls in wood, galvanized iron	0.017
Fan	0.688	Walls in stone	-0.204
Air conditioned	0.158	Floor in concrete, cement	ref
Radio	0.004	Floor in sand, earth, bamboo	-0.578
Tape, record player	0.398	Floor in wood	0.012
Television	0.775	Floor in stone, tile	0.445
Video cassette recorder	0.445	Ceiling in galvanized iron	ref
CD player	0.263	Ceiling in leaves, earth	-0.526
Camera, video	0.165	Ceiling in wood, canvas	0.043
Computer	0.185	Ceiling in concrete, cement	0.531
Bicycle	0.101	Ceiling in tile	0.244
Motorcycle, scooter	0.115	Windows poorly protected (c)	-0.557
Car	0.340	Multi-storied house or apartment	ref
Cable TV	0.472	One-story house	0.273
Electric iron	0.289	Apartment(s)	0.209
Furniture	0.369	Number of rooms (c)	0.139
		Well without pump	ref
		Inside private tap	0.525
		Water seller	0.019
		Outside private tap	0.282
		Well with pump	-0.107
		River, rain water, other	-0.059
		Public standpipe	-0.147
		Low hygienic toilets (c)	-0.725
		Electricity	0.814
		Kitchen	0.129
		Remote telephone (c)	-0.595

Note: (c) means that this variable take several ordered modalities.

B Sample statistics

Table 11: Levels of education, if $10 \leq$ age ≤ 21

Level	Proportion
no schooling	18.34%
incomplete primary	39.55%
complete primary	19.10%
incomplete lower secondary	16.71%
complete lower secondary	3.31%
incomplete higher secondary	2.21%
complete higher secondary	0.66%
university	0.13%

Table 12: Sample statistics

	Sample 1			Sample 2		
Variable	Sample size	Mean	Std. Dev.	Sample size	Mean	Std. Dev.
Attended school	6884	0.81	0.38	2636	0.83	0.36
Level attained	6830	2.57	1.25	2613	2.54	1.23
Father's level	6884	2.24	1.94	2636	2.32	2.01
Father has attended school	6884	0.39	0.48	2636	0.43	0.49
Mother's level	6884	1.59	1.23	2636	1.61	1.24
Mother has attended school	6884	0.26	0.43	2636	0.27	0.44
Wealth	6884	0.05	0.95	2636	−0.03	0.96
Rural	6884	0.52	0.49	2636	0.55	0.49
Boy	6884	0.51	0.49	2636	0.54	0.49
Household's size	6884	12.77	6.01	2636	12.99	5.93
No older brother	6884	0.52	0.49	2636	0.42	0.49
No older sister	6884	0.58	0.49	2636	0.50	0.50
French score	674	37.80	12.41	294	37.87	12.83
Math score	706	39.85	13.69	307	39.96	13.09

Note: Sample 1: children aged 10 to 21 for whom all the explicative variables are not missing; Sample 2: children aged 10 to 21 for whom the instrumentation variables are available, i.e parents live at home.

C Instrumentation

Table 13: Instrumentation of parental background variables

		Father education	Mother education	Wealth
Father side	primary school	0.821**	0.189**	0.167**
		(0.103)	(0.066)	(0.040)
	low. 2ndary school	0.396**	0.095	0.097+
		(0.139)	(0.090)	(0.055)
	upp. 2ndary school	0.134	0.216*	0.145*
		(0.145)	(0.094)	(0.057)
	health care	0.084	-0.065	0.016
		(0.119)	(0.076)	(0.047)
	father dead	-0.182	0.180*	0.027
		(0.114)	(0.073)	(0.045)
	father ill	0.014	-0.119	0.084
		(0.232)	(0.149)	(0.091)
	mother dead	0.026	-0.220+	-0.247**
		(0.178)	(0.114)	(0.070)
	mother ill	0.041	-0.284+	0.221*
		(0.226)	(0.146)	(0.089)
	no older brother	-0.364**	-0.129**	-0.033
		(0.069)	(0.044)	(0.027)
	no older sister	0.077	-0.002	-0.041
		(0.070)	(0.045)	(0.028)
Mother side	primary school	-0.115	0.123+	0.089*
		(0.106)	(0.068)	(0.042)
	low. 2ndary school	0.159	0.332**	0.030
		(0.130)	(0.084)	(0.051)
	upp. 2ndary school	0.814**	0.214*	0.367**
		(0.131)	(0.085)	(0.051)
	health care	0.345**	0.241**	0.116*
		(0.118)	(0.076)	(0.046)
	father dead	-0.087	-0.310**	0.107*
		(0.129)	(0.083)	(0.051)
	father ill	0.187	0.283+	0.058
		(0.233)	(0.150)	(0.091)
	mother dead	0.298	0.069	-0.131
		(0.208)	(0.134)	(0.082)
	mother ill	-1.142**	-0.184	-0.059
		(0.234)	(0.151)	(0.092)
	no older brother	-0.027	0.029	0.013
		(0.069)	(0.044)	(0.027)
	no older sister	-0.199**	0.002	-0.084**
		(0.069)	(0.044)	(0.027)
	Observations	2645	2646	2655
	R-squared	0.30	0.25	0.54

Note: The instrumentation of parental enrollment is done by probit estimation, the coefficients reported are not marginal effects. The instrumentation of parental wealth and of parental education level is done by OLS. Control variables of the interest regression are included but omitted in the table. + significant at 10%; * significant at 5%; ** significant at 1%. Standard errors in parentheses. Joint significance of the instruments in the 3 estimations: $\chi^2(54) = 1071.71$, $P > \chi^2 = 0.0000$.

D Test of over-identifying restrictions

Table 14: Test of over-identifying restrictions

Variables	Number	Identifying hypotheses	χ^2	$P > \chi^2$
Urban/rural	2	Infrastructures+Z	9.57	0.01
Infrastructures	8	Z	11.63	0.16
Sibling's education	2	Infrastructures+Z	7.17	0.03
Grand-parents' education	4	Infrastructures+Z	12.73	0.01
Grand-parents' housing	2	Infrastructures+Z	10.05	0.01
Number of siblings	2	Infrastructures+Z	5.21	0.07
Diff in educ with grand-parent	2	Infrastructures+Z	6.29	0.04
Diff in educ with siblings	2	Infrastructures+Z	19.26	0.00

Note: Variables noted Z are grand-parental health variables and the ranking of the parents in terms of birth order. The RHS columns present the result to the test that all coefficients associated to the variables under scrutiny are jointly non significantly different from zero in the schooling equation. Variables used as instruments for these tests are listed in the third column.