Are Successive Investments in Education Equally Worthwile? Endogenous Schooling Decisions and Non-linearities in the Earnings-schooling Relationship.

Ali Skalli*
ERMES, UMR CNRS 7017
Université Panthéon-Assas (Paris 2)
12, place du Panthéon - F-75231 Paris Cedex 05
skalli@u-paris2.fr

June 10th, 2004.

Abstract

Most of the studies that account for the endogeneity bias when estimating the returns to schooling assume a linear relationship between education and earnings. Studies that assume the latter relationship to be non-linear simply ignore the endogeneity bias. Moreover, they either assume an ad-hoc non linear relationship or argue that non-linearities are due to sheepskin effects. The approach I adopt in this paper allows estimation of the returns to years of schooling without assuming any explicit form of non-linearity while accounting for the endogeneity of education. The results suggest (i) endogeneity is indeed a crucial issue, (ii) there are sharp non-linearities which do not seem to be due to certification effects only and (iii) it is important to account for the impact of education on the returns to other observable characteristic such as age and seniority.

JEL classification : J21.

Keywords : Returns to schooling, Selectivity, Non-linearity.

*I would like to thank Giorgio Brunello, Peter Dolton, John Treble and participants to the "Labour Market Outcome of Educational Investments" workshop at Paris for valuable discussions and comments. I also thank people at LASMAS-IDL for making the data available. I am also grateful to the European Commission for financial support under grant HPSE-CT-2002-00108 for the EDWIN project.

1 Introduction

The rate of return to schooling measures the extra earnings a worker would get if she or he invests a further year in education. It is therefore an important factor in understanding educational attainment and participation. Moreover, in the framework of human capital theory, it is interpreted as a measure of the impact of education on the productivity individuals bring to the labour market. Hence, by comparison to the social return to education, it also sheds light on the public resources that must be devoted to the funding of education. This justifies the large body of literature that has developped, exploring the importance of a variety of potential sources of bias that might affect estimates of the returns to schooling (see Card, 2000).

A common feature of this literature is the use of Mincer-type equations where earnings or wages are expressed as a function of specific and general human capital measures including education and a number of other observable characteristics. Most of the time, however, the earnings-schooling relationship is assumed to be linear despite the existing evidence that it is not (see for instance Denny and Harmon, 2001 and Park, 1999). Of course, one advantage of the linearity assumption is that it allows one to easily implement standard testing strategies to account for potential sources of bias such as selectivity or the endogeneity of education.[1] Yet, there is a priori no reason why the latter biases would be more or less severe than the bias the linear hypothesis might yield. Moreover, the information provided by the return estimated using a linear relationship is in fact rather poor. For example, there is clear evidence that the average marginal return to schooling, estimated assuming linearity, is increasing over time in some countries and decreasing in others (Ashenfelter *et al*, 1999, Harmon *et al*, 2001). Suppose the marginal returns to education are in fact decreasing with schooling. Then a decrease of one percentage point in the average marginal return might be due to a large decrease in the return to higher education or to a slight decrease only in the return to an extra-year of secondary education. Besides this statistical effect, there are theoretical implications as well. For instance, a concave earnings-schooling relationship could be interpreted in the framework of human capital theory as the result of a decrease in the amount of human capital individuals accumulate each extra-year of school attendance. Interpretation of such a pattern in the framework of signaling theory is, in contrast, more problematic.[2] That the marginal returns to high schooling levels be decreasing could indeed be interpreted as a means potential employers could use to induce low ability individuals not to invest any further.[3] That the marginal returns to lower secondary education be higher than the marginal returns to tertiary education is, however, by no means, a

[1] One exception is Brunello and Miniaci (1999) who use IV methods to estimate the returns to primary, secondary and tertiary education.

[2] See Spence (1973, 1974), Arrow (1973) and Riley (1979). See also Layard and Psacharopoulos (1974) for an analysis of the empirical problems discrimination between human capital theory and signaling theory poses.

[3] This indeed results in marginal returns to the investment in signaling that are lower than the marginal costs low ability individuals would have to incur if they undertake the investment

prediction of signaling theory.

A second problem with the conventional approach to estimating the returns to schooling is that it is often assumed that the returns to all other observed characteristics do not depend on schooling. To be more specific, the return to schooling is in general estimated for individuals that have been made comparable in terms of all observables. This, however, imposes a further restriction besides that of a linear earnings-schooling relationship: that schooling does not affect the returns to other individual endowments and, in particular, that individuals with different levels of schooling face the same age-earnings profiles. Under such a hypothesis, optimizing individuals would invest in education if and only if investment in an extra-year of education yields a higher starting wage. Education would then affect starting earnings and individuals would accumulate general and specific human capital throughout their working lives at the same rythm, whether they are early school leavers or college graduates. This hypothesis is too strong as one would expect the highly educated to be better endowed to 'learn by doing' and hence to accumulate general and specific human capital more efficiently than low educated individuals and therefore to face not only higher starting earnings but steeper age-earnings profiles as well. Even in basic human capital models, individuals undertake a further investment in education if the net present value of the investment is positive even if they expect a lower starting wage. Hence, what determines the decision to invest is not the starting wage only but the expected age-earnings profile, together with individuals' discount rates. Thus, education influences at least the returns to age and possibly the returns to other observables as well. Therefore, an appropriate measure of the returns to education should capture not only the effect on starting wages but also the effect on the returns to all individual endowments. Besides the theoretical justification of such effects, there is also clear evidence pleading for their existence. For instance, Lollivier and Payen (1990) and Lhéritier (1994) show that, in France, there are sometimes large differences in the returns to age, job tenure and a variety of other individual and job characteristics between blue collar and white collar workers. To the extent that occupational status is correlated to educational attainments, such differences are very likely to be due to the effect of education on the returns to these characteristics.

A third problem with the Mincerian approach is the endogeneity of schooling. However individuals' investment in education is measured, it is indeed hard to assume it is exogenous and hence to estimate the associated return using Ordinary Least Squares (OLS). This is why a wide body of literature attempted to examine how severely OLS estimates of the returns to schooling might be biased. The evidence from this literature is, however, mixed. It suggests that whether OLS estimates are upward or downward biased depends on how ability differences are accounted for. For example, studies where endogeneity is accounted for via the inclusion of an explicit measure of ability report an upward bias in OLS estimates (Blackburn and Neumark, 1993) whereas those based on panel data and where ability is captured by individual fixed effects conclude to a downward bias in OLS estimates (Guillotin and Sevestre, 1994). Another approach consists in eliminating differences in innate ability by exploiting differences between twins or siblings in the levels of schooling and earnings. Using U.S. siblings and twins data, respectively, Ashenfelter and Zimmerman (1993) and Ashenfelter and Krueger (1994) report estimates that are much higher than typical OLS ones. In contrast, using U.K. twins data, Blanchflower and Elias (1993) find evidence of an upward bias in OLS estimates. However, studies using Instrumental Variables (IV) by exploiting natural variations in data caused by exogenous influences on the schooling decision systematically conclude to a downward bias in OLS estimates. (see Angrist and Krueger, 1991, Card, 1993, Kane and Rouse, 1995, Dearden, 1995, Harmon and Walker, 1995 and Uusitalo, 1999 among others). Eventhough, there is no unanimity in these studies about the importance of the endogeneity bias. For example, while Angrist and Krueger (1991, 1992) conclude to a limited impact of endogeneity, the results in Butcher and Case (1994) or Kalwij (1996) suggest such an impact is rather large.[4] One common feature of these studies which might explain these differences is that schooling is treated as a continuous variable. Garen (1984) considers the case in which education is coded as an ordered integer and shows that the discrete nature of the educational choice set implies that standard simultaneous-equations estimators are not consistent due to the nature of the disturbances. This suggests that as long as schooling is recorded as an integer in the data, a simultaneous-equations framework where schooling is modeled as a discrete choice variable is preferable to the usual approach.

In this paper, I adopt a testing strategy which simultaneously overcomes these three limitations of the conventional approach to estimating the returns to schooling. By estimating a generalized version of the switching regression model with endogenous switching (see Maddala, 1983), I account for endogeneity through selectivity models while the returns to any characteristic I control for are left free to vary from one schooling level to another. This allows estimation of subsequent marginal returns to schooling as the differences in expected log-wages between two subsequent levels of schooling. Thus, the marginal returns to education are left free to vary across schooling levels, hence obeying to no specific non-linearity scheme. This is different from the conventional approach to testing the linear hypothesis. For instance, Denny and Harmon (2001) test quadratic and cubic earnings-schooling relationships and systematically reject the linear hypothesis. Another approach consists in controling for years of schooling as well as for qualification levels (see Chevalier and Walker, 2001). In the same vein, Park (1999) estimates the returns to qualifications according to the number of years of schooling it takes one to attain them. While Denny and Harmon's (2001) approach is based on ad-hoc earnings-schooling relationships, Park's (1999) highlights sheepskin effects. That is, the existence of 'bonus returns' to attaining a qualification in due time. Sheepskin effects are in general interpreted as the rewards to signaling higher ability than individuals who failed to attain qualifications or to whom it took a longer time to attain them. Thus, the latter approach imposes certification as the only source of non-linearities

[4] Rummery *et al* (1999) adopt an approach based on rank-order instrumental variables and find estimates for Australia that are close to their OLS counterparts.

and, as a matter of fact, signaling effects as the only possible explanation of such non-linearities.

The idea of using a selectivity model to account for the endogeneity of education is not new. Willis and Rosen (1979) estimate a switching regression model with endogenous switching where selectivity is modelled through a probit model describing the probability that individuals attend college or not. They, however, do not estimate marginal returns to schooling, but rather the wage differential between college attendees and those who leave school earlier. Harmon and Walker (1995) make a step further by accounting for selectivity through an ordered probit model the left hand side variable of which is the number of years of schooling treated as an ordered integer. The earnings function they estimate is, however, linear in schooling and assumes identical age-earnings profiles whatever individuals' schooling levels are. To my knowledge, only Vella and Gregory (1996) account for selectivity as well as interaction effects by estimating separate earnings functions for 7 educational groups. Unfortunately, the data they use include male school leavers aged 15 to 26 only and hence do not allow them to assess the association between schooling and age-earnings profiles.

The paper is organized as follows: section 2 describes the data and the empirical setup whereas section 3 discusses the results with respect to the endogeneity issue, to the non-linearity of the earnings-schooling relationship and to the effect of education on age-earnings profiles. Section 4 concludes the paper.

2 Data and Empirical Setup

The conventional IV approach to estimating the returns to schooling consists of the estimation of a two-equations system describing log-earnings, y_i, and the number of years of schooling (NYS), S_i:

$$\begin{aligned} y_i &= X_i'\beta + rS_i + u_i && (1) \\ S_i &= Z_i'\gamma + v_i && (2) \end{aligned}$$

where X and Z are vectors of observed characteristics and where $E(X_i, u_i) = E(Z_i, v_i) = 0$. Only if S_i is exogenous would estimation of (1) yield an unbiased estimate of the return to schooling, r. Otherwise, (1) and (2) are to be estimated simultaneously. In general, S_i is treated as a continuous variable and IV methods are therefore used. A problem with this approach, however, is that it treats schooling as a continuous measure though, in general, no information on the date of leaving school in any year is given so that the schooling measure is in fact an integer. Moreover, there is in general a high proportion of individuals who leave school at the compulsory schooling level. Because of this, the nature of the disturbances is such that the resulting estimates are not consistent (see Garen, 1984). Using U.K. data, Harmon and Walker (1995) estimate a selectivity model where (2) is replaced with an ordered probit equation. While this yields an estimate of the returns to schooling of 16.9%, the IV approach suggests such an estimate to be 15.3%.[5] More specifically, Harmon and Walker (1995) use an

[5] These figures are to be compared to an OLS estimate of 6.1%.

extension of Heckman's two-step approach which they apply to the estimation of the following model:

$$y_i = X_i'\beta + rS_i + u_i \tag{3}$$
$$S_i^* = Z_i'\gamma + v_i \tag{4}$$
$$S_i = j \quad \text{if} \quad \mu^{j-1} < S_i^* \leq \mu^j \tag{5}$$

where S_i^* is the latent variable corresponding to S_i and where $j = 1, 2, \cdots J$. The μ's are unknown parameters to be estimated which indicate the threshold values for moving through the schooling participation decision.

A common feature of the two approaches described above is that earnings are assumed to be linear in schooling. Moreover, both specifications assume that the remuneration of the individual endowments captured in the vector X_i does not depend on individuals' schooling grades. One possible alternative approach could consist in estimating an extended version of the switching regression model with endogenous switching, based on the ordered probit model. The alternative model structure would then be:[6]

$$y_i^j = X_i'\beta^j + u_i^j, \quad j = 1, 2, \cdots, J \tag{6}$$
$$S_i^* = Z_i'\gamma + v_i \tag{7}$$
$$S_i = j \quad \text{if} \quad \mu^{j-1} < S_i^* \leq \mu^j. \tag{8}$$

Suppose that $u_i^j, j = 1, \cdots, J$, and v_i are distributed as $(J+1)$-variate normal. It is shown in the appendix that this model entails the following earnings equations for observed schooling levels $j = 1, \cdots, J$

$$y_i^j = X_i'\beta^j + \rho^j\sigma^j\lambda_i^j + \varepsilon_i^j, \quad j = 1, 2, \cdots, J \tag{9}$$

where,

$$\lambda_i^j = \frac{\phi\left(\mu^{j-1} - Z_i'\gamma\right) - \phi\left(\mu^j - Z_i'\gamma\right)}{\Phi\left(\mu^j - Z_i'\gamma\right) - \Phi\left(\mu^{j-1} - Z_i'\gamma\right)}, \tag{10}$$

μ^0 and μ^J being taken as $-\infty$ and $+\infty$, respectively and $\Phi(\cdot)$ and $\phi(\cdot)$ denoting the standard normal distribution and density functions, respectively. ρ^j, $j = 1, 2, \cdots, J$ are the correlation coefficients of u_i^j with v_i in their respective marginal distributions. The variances of u_i^j, $j = 1, 2, \cdots, J$ and v_i are σ^j and 1 respectively. ε_i^j, $j = 1, 2, \cdots, J$ are zero mean random variables distributed independently of X_i and λ_i^j.

The ordered probit model entails, however, some restrictions on the effects of the elements of the vector Z on the probability of attending school beyond the k^{th} schooling level, $k = 1, \cdots, J-1$, relative to the probability of reaching the j^{th} one, $j = 2, \cdots, J$, $j > k$. In particular, for a given Z, the effect of an element of Z on the probability of reaching the j^{th} schooling level is proportional to its effect on the probability of not leaving school after the k^{th} grade, the

[6] This is formally close to Vella and Gregory's (1996) approach.

proportionality factor being $\phi\left(\mu^j - Z_i'\gamma\right)/\phi\left(\mu^k - Z_i'\gamma\right)$. In order to overcome these proportionality restrictions, I also use a method that has been suggested by Nakamura & Nakamura (1983), albeit in a different context.[7] It consists in estimating $J-1$ two-outcome probit equations, each for whether an individual attends school beyond the j^{th} level, $j = 1, \cdots J-1$, using the full sample of individuals in each case. The model structure in this case is:

$$\begin{aligned} y_i^j &= X_i'\beta^j + u_i^j, \quad j = 1, 2, \cdots, J && (11) \\ S_i^{j*} &= Z_i'\gamma^j + v_i^j, \quad j = 1, 2, \cdots, J-1 && (12) \\ S_i &> j \qquad \text{if} \quad S_i^{j*} > 0 && (13) \end{aligned}$$

the inverse Mill's ratios being then calculated as:

$$\lambda_i^j = \frac{\phi\left(-Z_i'\gamma^{j-1}\right) - \phi\left(-Z_i'\gamma^j\right)}{\Phi\left(-Z_i'\gamma^j\right) - \Phi\left(-Z_i'\gamma^{j-1}\right)} \tag{14}$$

where $\gamma^0 = 0$ and where γ^j, $j = 1, \cdots J-1$, is the parameter vector from the j^{th} probit equation.[8]

In both cases, as in Heckman's (1979) original procedure, while the $\lambda_i^{j\prime}$s are not observed, consistent estimates of them are derived from using consistent estimates of the γ parameter vectors and, in the ordered probit approach, the corresponding μ^{j-1} and μ^j. The existence of sample selection bias (and therefore the endogeneity of schooling decisions) could then be examined via a test of the null hypothesis that the $\rho^{j\prime}$s are zero using the t-distribution.

For each schooling level $j = 1, 2, \cdots, J$, estimation of equations (9) yields a specific estimate $\widehat{\beta}^j$ which in turn, allows simulation of the earnings distribution one would have observed had all individuals had schooling level j as:

$$e_{i|j} = \exp\left\{X_i'\widehat{\beta}^j\right\}, \quad j = 1, 2, \cdots, J. \tag{15}$$

Thus, for each individual i with schooling level j, the (marginal) return to investing a j^{th} year in education is given by:

$$r_{m,i}^j = \frac{e_{i|j}}{e_{i|j-1}} - 1, \tag{16}$$

and the average marginal return r_m^j, associated with schooling level j is simply the sample mean of individual marginal returns. Likewise, cumulative returns to

[7] This method has first been proposed by Nakamura and Nakamura (1983) and used by Ermisch and Wright (1993), both in the context of modelling women decision to participate to the labour market and then to work part or full time.

[8] In contrast to the ordered probit approach, calculation of the λ variables in (14) requires estimation of $J-1$ times as many parameters than in (10), reflecting the proportionality restrictions in the ordered probit model. A problem with this method, however, is that the estimated probability of attending school up to the j^{th} level, $\Phi\left(-Z_i'\gamma^j\right) - \Phi\left(-Z_i'\gamma^{j-1}\right)$ is not constrained between zero and one. Indeed, for some 3,600 individuals out of 297,599, estimation yielded negative probabilities. The corresponding observations have been deleted.

schooling with reference to the lowest schooling level ($j = 1$) could be estimated as:

$$r_{c,i}^{j} = \frac{e_{i|j}}{e_{i|1}} - 1 \tag{17}$$

and the average cumulative return r_c^j, associated with schooling level j as the sample mean of individual cumulative returns.

To test the null hypotheses that the marginal returns are zero, one could think of a series of F-tests, the null of which are $\beta^j = \beta^{j-1}$, $j = 2, \cdots, J$. Such a testing strategy does not, however, account for differences in the characteristics described in the vector X between individuals with different schooling levels. It is therefore preferable to rather envisage a series of t-tests where the null hypotheses are $r_m^j = 0$, $j = 2, \cdots, J$. Similar tests can of course be also performed to test whether the cumulative returns are zero or not.

That the successive marginal returns are not zero does not imply they are not equal to each other. To test whether they differ according to schooling levels or not and therefore whether the earnings-schooling relationship is linear or not, a series of t-tests can again be performed. The null hypotheses are then:

$$\forall j, h = 1, \cdots, J, \quad j \neq h, \quad r_m^j = r_m^h. \tag{18}$$

Besides the simultaneous treatment of endogeneity and non-linearity issues, a further advantage of the endogenous switching approach is that it allows direct examination of differences in the returns to the characteristics described in the vector X. As mentioned above, these could be easily assessed via F-tests. Nevertheless, the characteristics in X are not of the same importance. The returns to some of these characteristics are more likely to be influenced by, or be dependent of, schooling levels. One component I focus my attention on is age as age-earnings profiles are very likely to vary with schooling levels. Thus, F-tests cannot help making such a distinction. Let ϕ^j and Φ^j denote $\phi\left(\mu^j - Z_i^{'}\gamma\right)$ and $\Phi\left(\mu^j - Z_i^{'}\gamma\right)$ respectively in the ordered probit model or $\phi\left(-Z_i^{'}\gamma^j\right)$ and $\Phi\left(-Z_i^{'}\gamma^j\right)$ respectively in the $J-1$ two-outcome probit models. It is shown in the appendix that an alternative testing strategy could consist in estimating the set of $J-1$ equations:

$$\begin{aligned} E\left(y_i \mid S_i = j \cup S_i = j+1\right) \quad = \quad & \left(X_i \cdot \Phi^{j+1}\right)' \beta^{j+1} + \left(X_i \cdot \Phi^{j-1}\right)' \beta^j + \\ & \left(X_i \cdot \Phi^{j}\right)' \left(\beta^j - \beta^{j+1}\right) - \rho^{j+1}\sigma^{j+1} \cdot \phi^{j+1} + \\ & \rho^j\sigma^j \cdot \phi^{j-1} + \left(\rho^{j+1}\sigma^{j+1} - \rho^j\sigma^j\right) \cdot \phi^j \end{aligned} \tag{19}$$

using the sub-sample of individuals having attained the j^{th} or the $j+1^{th}$ grade. This allows one to directly test the null hypotheses that the coefficients on the components of X for grades j and $j+1$ are equal.[9] Note also that if the β's,

[9] Note that with $J = 2$, (19) reduces to Maddala's (1983) equation (8.20) and is therefore equivalent to the switching regression model with endogenous switching, with two regimes.

except the constant terms, are assumed to be equal for grades j and $j+1$, then the equations in (19) reduce to:

$$\begin{aligned} E\left(y_i \mid S_i = j \cup S_i = j+1\right) &= \left(X_i \cdot \Phi^{j+1}\right)' \beta^{j+1} + \left(X_i \cdot \Phi^{j-1}\right)' \beta^j + \\ & \left(X_i \cdot \Phi^j\right)' \left(c^j - c^{j+1}\right) - \rho^{j+1}\sigma^{j+1} \cdot \phi^{j+1} + \\ & \rho^j \sigma^j \cdot \phi^{j-1} + \left(\rho^{j+1}\sigma^{j+1} - \rho^j \sigma^j\right) \cdot \phi^j \qquad (20) \end{aligned}$$

where $c^{j+1} - c^j$ denotes the expected return to investing a further year beyond grade j under the constraint that any other component of X is remunerated the same rate whether individuals undertake the extra investment or not.[10]

Among the β coefficients, those associated with age are of particular importance. Not only would one expect age-earnings profiles to be different from one schooling level to another, but also that they be steeper as one moves from one schooling level to a higher one. For individuals in the j^{th} group, $j = 1, \cdots, J$, starting wages, start_i^j, could be estimated using the j^{th} equation in (9) where age is set equal to the number of years of schooling augmented by 6 years and where job seniority is set equal to 0. For the average individual in that group, the starting wage, start^j, is simply the sample mean of start_i^j. Likewise, assuming Equations (9) are quadratic functions of age, the slope of the corresponding age-earnings profile could be estimated as:

$$\text{slope}_i^j = \exp\left\{\widehat{y}_i^j\right\} \cdot \left(\widehat{a}^j + 2\widehat{b}^j \text{age}_i\right)$$

where $\widehat{y}_i^j$ is the predicted wage of individual i with schooling level j and where $\widehat{a}^j$ and $\widehat{b}^j$ are the estimated coefficients associated with age and its square, respectively. Again, for the average individual in group j, the slope, slope^j, of the age-earnings profile is simply the sample mean of slope_i^j. The age, peak^j, at which the age-earnings profile of individuals in group j peaks is then $-\widehat{a}^j/2\widehat{b}^j$. Rank correlation coeficients of schooling levels $j = 1, 2, \cdots, J$ with start^j, slope^j and peak^j, together with their significance levels could be used to examine how age-earnings profiles vary with schooling levels and, in a sense, the extent to which they are influenced by education.

To conduct such a testing strategy, I use the French Labour Force Survey (LFS). The sample consists of some 300,000 full-time male workers aged 16-65 in the year of interview, obtained from pooling the eleven consecutive annual LFS cross-sections from 1990 to 2000.[11]

The specification I consider for equation (6) has deflated gross monthly wages on the left hand side.[12] In the right hand side, I include age and its square,

[10] Note that with $J = 2$, (20) reduces to the usual treatment effect model.

[11] This is the so-called *enquête Emploi* which is a rotating panel where individuals are observed thrice. Earnings information in data prior to 1990 is given in earnings bands. Summary statistics are given in Table 1.

[12] Net earnings are not predictable from the data because of the non-neutrality of the French tax system and because taxes are calculated on a household basis. The data provides usually worked hours as well as actuallly worked hours of the last week, so that one could a priori

job tenure and its square, three marital status dummies (married, single, widowed, divorced individuals being the omitted group), the number of children, seven regional dummies, three ethnicity dummies (Immigrants with French citizenship, western countries citizens and immigrants from other countries, the French being the omitted group), four occupational status dummies (tenured contract, temporary worker, trainee, beginner, Fixed-term contract workers being the omitted group) and, finally, ten year dummies (those observed in 1990 being the omitted group).

Equation (7) on the other hand is first estimated as an ordered probit on the left hand side of which is a qualitative variable with nine possible outcomes. The schooling level of those who attended school up to age 16 (10 years of schooling) is coded 1.[13] That of the others is coded as $j = S - 10 + 1$ where $S = 11, \cdots, 17$, is the actual NYS, and $j = 9$ for those who attended school for more than seventeen years. Alternatively, equation (7) is also splitted into eight two-outcome probit equations, each modeling the decision to attend school beyond the j^{th} level, $j = 1, \cdots, 8$, and including the same right hand side variables as the ordered probit equation.

In an individual optimizing framework, the decision to leave or to invest a further year in education depends on individuals' wage expectations. Therefore, the reduced form probit equations I estimate, either ordered or not, include all observable wage detrminants. That is, all the regressors in equation (6) are included in the vector Z. However, like in the IV approach, identification in the model is provided only if the vector Z includes at least one variable that is not included in X. That is, there must exist a variable which is a determinant of schooling participation that can legitimately be omitted from the earnings equation. Here I consider several candidates to play such a role. In an IV framework, Dearden (1995) uses parental education and social class while Uusitalo (1999) considers parental income and education. As an indicator of these family background dimensions, I take four dummies describing fathers' occupational status at the time the individual left school (public sector employee, private sector employee, self-employed with no employees, self-employed with less than 10 employees, those whose father is/was a self-employed with more than 10 employees being the omitted group). Besides parental income and education, Uusitalo (1999) also considers the location of residence. This variable is, however, likely to be highly correlated with the working location and therefore with wages at least via the well known capital-city effect. Instead, I include seven dummies indicating the region of birth. In a sense, this variable might also be considered as a reflecting the effect of the differences in regional endowments in terms of schools and colleges on individuals' educational decisions. This idea underlies the work by Card (1993) who uses an indicator of nearby college in county of residence and by Kane and Rouse (1995) who consider the distance to nearest

---

calculate a proxy of hourly earnings. The resulting variable might, however, suffer from large measurement errors for certain occupational categories such as teachers, white collars, self-employed, seasonal workers, etc.

[13] Compulsory schooling in France has been established at 13 years in 1882, 14 in 1936 and 16 since 1959.

college. Angrist and Krueger (1991) explore how an individual's season of birth may imply that some students reach school leaving age after fewer months of compulsory education than others. I also exploit this idea by including three dummies indicating the season of birth (those born in autumn being the omitted group). Harmon and Walker (1995) analyze the impact of the raise in school living age that occured in the U.K. in 1947 and in 1973 and observe that many of those individuals who would otherwise have left school at the old minimum stayed on beyond the new minimum. This suggests that the exogenous change in the minimum school leaving age is influential in raising participation in post-compulsory education. Following Harmon and Walker (1995), I also include a dummy which distinguishes between individuals who reached the minimum school leaving age before it has been raised from 14 to 16 in 1959 and the others. Finally, although it is likely to be correlated to unemployment levels and therefore to wage levels, I also consider the long-term interest rate which prevailed the year individuals left school to enter the labour market. The motivation for this is simply to better predict the probability that individuals attend school for a given number of years, the idea being based on human capital theory where individuals's investment decisions are based on the comparison of the internal rate of return to the market interest rate.

3 Returns to Schooling and Age-Earnings Profiles

In this section, I report and interpret the results obtained using the data and methodologies described above. I start by identifying the determinants of schooling decisions. I then examine the sensitivity of the estimated returns to schooling to the hypotheses of a linear wage-schooling relationship, of exogenous schooling decisions and of no dependence of the returns to individual characteristics on schooling levels.

3.1 The Determinants of Schooling Decisions

Table 2 reports estimates of reduced form schooling equations using the instrumental variable approach (columns 2 and 3) and the ordered probit model (column 4 and 5), using the same set of right-hand side variables. These include all characteristics which are tought of as wage determinants as well as a set of factors which are meant to potentially influence wages only via their impact on the choice of their schooling levels by individuals.

A remarkable feature of these results is that both the IV and the ordered probit approaches suggest that most of the regressors have qualitatively similar impacts on the choice of an individual level, the only exception being the region of birth, most of the indicators of which have a positive sign in the IV equation and a negative sign in the ordered probit one. This apparent inconsistency is certainly due to the fact that while in the continuous schooling equation, each value of the left-hand side variable refers to the individuals' actual number of years of schooling, in the discrete choice model, the extreme values of the schooling level variable refer to individuals with different numbers of years of schooling. For instance, schooling level 9 refers to all individuals with 18 or more years of schooling. Therefore, it is the difference in the variability of the schooling indicator that yields these seemingly contradictory signs associated with the region of birth dummies. Since all the other regressors have the same sign and have therefore qualitatively similar impacts on the schooling measures considered in the IV and in the ordered probit equations, I shall focus my attention on interpreting the latter to avoid any confusion.

Looking at wage determinants, the signs on age and its square indicate that old individuals are less likely to reach high levels of education than young ones. This is a cohorts effect indicating that young cohorts are better endowed with education than older cohorts. Likewise, job tenure seems to have a decreasing and convex impact on the likelihood that individuals reach the highest schooling levels. This is exactly what one would expect since among two equally aged individuals, the one with the shortest schooling duration is more likely to have a longer job tenure. The negative sign on the number of children indicates either that child care duties limit the likelihood that individuals keep on investing in education or simply that the highly educated have fewer children than individuals with low educational attainments. It seems also that divorced individuals are not likely to attend school for as long as if they were married or singles. This might also suggest that divorce is more likely to occur among low educated individuals. Note also that, compared to the French, immigrants are less likely to reach higher educational levels wherever they come from and whether they are French citizens or not. Probably, immigrants expect lower returns to education than the French, whether they left school in their home country or once in France. It might also be the case that immigrants discount future earnings more heavily than the French. Note finally that job status is also highly correlated with the probability that individuals attain higher educational grades. Compared to individuals with fixed term contracts, those who are tenured or temporary workers are less likely to be highly educated. In contrast, beginners are more likely to have undertaken long educational investments. This is consistent with the job tenure effect discussed above.

Examining the group of variables that do not enter the wage equations, it is worth noting that their effets on schooling decisions are also highly significant. This is the case of father's occupation at the time individuals left school. Indeed, it seems that being the son of a self-employed having more than ten employees significantly increases the probability that an individual reaches high schooling grades. The region of birth has also a significant impact on the likelihood that long educational investments be undertaken. According to the ordered probit model, being born in the south-east of France increases this likelihood. Perhaps more interesting is the positive effect of the interest rate which prevailed during the school living year on individuals' schooling decisions. According to human capital theory, one would expect that the higher is the interest rate, the less likely are individuals to undertake educational investments. Note, however, that the interest rate included here is the one which prevailed the year individuals have decided to enter the labour market, not when they have had to decide upon educational investments. Therefore, the reported positive effect could be interpreted in two ways. It possibly indicates that the higher the interest rate in a given year, the more likely are individuals to stop attending school that year. A probably more convincing interpretation is that the higher is the interest rate in a given year, the more likely it is that it has been previously low enough when individuals have had to decide upon their investment. Besides, the season of birth seems also to be an important determinant of the schooling levels individuals reach. Compared to those born in autumn or in winter, spring and summer natives are more likely to reach high schooling grades. This result is not clearly in line with Angrist and Krueger's (1991) prediction that individuals born in the beginning of the year start school at an older age and are therefore likely to drop out after completing less schooling than those born towards the end of the year. While this argument should imply a negative sign on the spring dummy as well, the results here suggest that spring and summer natives attend school longer than autumn natives. Note, however, that as mentioned by Levin and Plug (1999), educational psychologists acknowledge that an opposite effect might also be at play; that cutoff dates contribute to a relative age effect whereby individuals starting school at an older age tend to enroll in more schooling. The idea is that within classes, older pupils tend to receive better marks which is thought to encourage further schooling.[14] Probably, the signs on the season dummies only reflect the balance between these two opposite effects. Finally, the raise in school leaving age seem to have had a positive impact on educational attainments. Those who were allowed to leave school at 14 are less likely to attain high educational levels than those who have had to wait until age 16.

Though not reported, the results from 8 two-outcome probit equations including the same regressors as the ordered probit equation highlight interesting patterns as well. While most of the right-hand side variables remain highly significant with the same signs, some others seem to have a level-specific effect. For instance, the effect of seasons of birth reveals to be more subtle than what the ordered probit estimates had suggested. Compared to autumn natives, those born in winter are more likely to attend school beyond schooling levels 1 and 2, but are less likely to do better than high school (level 3). In addition, birth in spring increases the probability of reaching higher grades but such an effect decreases steadily until it becomes unsignificant at graduation level (level 6) and beyond. Likewise, the effect of being born in summer is positive only before the high school level (level 3), not beyond. These patterns reflect that the balance between the compulsory school rules effect and the relative age effect is in fact level-specific.

[14] See Sharp (1995).

3.2 Estimates of the Returns to Schooling: Endogeneity and Non-linearity Taken Together

To simultaneously account for possible non-linearities in the wage-schooling relationship as well as for the endogeneity of schooling levels, I estimate a wage equation for each of the 9 educational levels. This requires that selectivity be controlled for. Inverse Mill's ratios should therefore be estimated and this could be done using either the ordered probit equation or the two-outcome probit equations. Again, because the results from both approaches are qualitatively similar, I only report those based on the ordered probit approach in Table 3.

A striking feature of Table 3 is the systematic high significance level of the coefficients associated with inverse Mill's ratios. This is evidence that correction for selectivity is crucial and therefore that the schooling levels individuals reach are indeed endogenous. Besides, the results are again in line with those usually observed. Age-earnings profiles are systematically increasing and concave. Likewise, except for schooling levels 5, 8 and 9 where they seem to be linear, tenure-wage profiles are also increasing and concave. In addition, the impact of the number of children on wages is clearly regime-specific. Positive for some schooling levels, it is unsignificant or even negative in others. A similar pattern applies to the wage differential between widows and divorced individuals which is positive for levels 1 and 9, negative for level 2 and unsignificant otherwise. In contrast, the marriage effect is systematically positive whereas the single earn systematically less than the divorced. The effect of ethnic dummies is also interesting as it highlights that only immigrants from western countries might earn more than the French, especially when they are endowed with higher education. In contrast, immigrants from other countries systematically earn less than the French and the wage gap seems to be increasing with educational levels. This is also the case of immigrants with French citizenship, although the estimated wage differential is lower whatever individuals' educational level is. Looking at job status indicators, it turns out that tenured workers earn significantly more than workers with a fixed term contract, and the difference is larger among highly educated individuals. In contrast, trainees earn less, especially among the highly educated. Less stable is the wage gap between temporary workers and wage earners with a fixed term contract. While the former earn more in general, their gain vanishes for very high schooling levels (7 and 8) and is even negative among those having reached the highest level (9). Finally, only when endowed with higher education (levels 7, 8 and 9) do beginners benefit from a starting wage differential. This premium is probably due to an excess demand for the highly educated. Such an interpretation is, however, less likely to be valid for those who left school immediately after compulsory schooling as there is a positive differential in favour of beginners in this category too. For these workers, the observed differential is more likely to be due to the minimum wage effect.

The returns to education estimated through the switching regression model with endogenous switching are reported in the column labeled "switching regression" in Table 4. This column comprises 2 sets of estimates each based on a specific choice model to control for selectiviy : the ordered probit model or the two-outcome probit model. Although the two sets of estimates are rather different, they both show that the marginal returns to education are not constant, hence suggesting that the wage-schooling relationship is not linear. While one further year after compulsory schooling yields a wage increase between 15 and 16%, the first year of tertiary education yields an extra wage of less than 8% above the salary of those who undertake no further investment after high school graduation. In addition, the wage-schooling relationship is not concave either. For instance, the fourth year of higher education induces a wage increase of more than 10%. Overall, the estimated marginal returns obey to no specific functional form; rather, they oscillate across educational levels. Note however, that since no marginal return is negative, the cumulative returns are steadily increasing.[15] With an overall cumulative return of more than 102%, those who reach the highest educational levels earn more than twice the average wage of early school leavers.

One important characteristic of these estimates is that they relax the conventional hypothesis that the remuneration of individual endowments does not depend on educational levels. This hypothesis is in fact too strong as one would expect at least that the higher is the level of education, the steeper age-earnings profiles would be. Besides the already mentioned differences between educational levels in the estimated coefficients associated to the various regressors included in the wage equations, estimation of equations of the type (19) allows one to directly test the significance of such differences. Table 5 reports the outcome from such tests for age and job tenure and their squares.[16] Each column in Table 5 is labeled $j+1$ vs. j, $j = 1, \cdots, 8$, and compares the estimated coefficients associated with educational levels $j+1$ and j. It is obvious from this table that age-earnings profiles differ between educational groups. The same holds for job tenure, albeit to a lesser extent. One might of course argue for instance that a difference between groups 1 and 2 and between groups 2 and 3 in terms of age-earnings profiles does not mean that groups 1 and 3 are different as well. However, the point here is that the switching regression approach would have imposed itself even if one group only differed from the others.

Taken together, the wage equations in Table 3 and the marginal returns reported in the "switching regression" column of Table 4 provide one with another means of highlighting the relationship between educational levels and age-earnings profiles. Table 6 reports Spearman rank correlation coefficients between the former and the latter, depending on whether selectivity is controlled for using the ordered probit or the two-outcome probit model. The first row above the diagonal and the first column below the diagonal show that the higher

[15] When selectivity is controlled for using the ordered probit model, the estimated marginal return to the highest educational level is negative (-0.4%). Such an estimate is, however, much different from all those obtained using a variety of other methods. This is a further reason why my preference goes to the two-outcome probit based approach.

[16] Although not reported, the complete regression results are available upon request from the author. Obviously, the choice of including age and tenure in Table 4 relies on their importance from the point of view of human capital theory.

educational levels are, the higher is the starting wage and the steeper is the age-earnings profile. The correlation between schooling levels and the peaking age of age-earnings profiles is low and unsignificant. This is in contradiction with the usual expectation that the higher education is, the later do age-earnings profiles peak.

These results suggest that neglection of differences in the returns to individual endowments might yield to misleading estimates of the returns to education. The columns labeled "Treatment Effect" in Table 4 report estimates of the returns to schooling based on such hypothesis; that is on equation (20). It is clear from these columns that the resulting marginal returns are rather different from those based on switching regressions. In particular, the treatment effect model overestimates the marginal returns to the extreme educational levels (2 and 9). In between, there are educational levels for which the marginal returns are overestimated and others for which they are underestimated. Interestingly, such a pattern results in an overall cumulative return to the highest educational level which is significantly lower that the corresponding switching regression estimate. While the latter is at least 102%, the former is less than 93%.

Another interesting result that Table 6 highlights is that marginal returns to education are not correlated to educational levels. Although this would have been expected from the pattern highlighted in Table 4, of how marginal returns evolve with schooling levels, it could also be considered as evidence that not only is the wage-schooling relationship not linear, but that the marginal returns are neither monotonically increasing nor monotonically decreasing.

Although the switching regression and the treatment effect models differ in the underlying assumption on whether wage determination is level-specific or not, their common feature is that they both account for selectivity. It is therefore important to evaluate the bias that might affect the estimated returns if the endogeneity of schooling decisions is also neglected. The columns labeled "OLS" in Table 4 report Ordinary Least Squares estimates of marginal as well as cumulative returns to education. It can easily be seen that these estimates are in general downward biased. More specifically, while those from the switching regression model are systematically larger, those from the treatment effect model are, only when based on the two-outcome probit model. In contrast, when selectivity is accounted for using the ordered probit model, only for some schooling levels are the estimated returns larger than OLS ones. However, comparison of cumulative returns to the highest level of education shows that OLS estimates tend to be the lowest whatever the selectivity corrected estimates they are compared to.

Overall, a synthetic comparison of the estimates of Table 4 could be based on their average values. These are obtained simply by dividing the cumulative return to the highest level by the number of estimated returns from each model. These mean values are reported in the bottom row of Table 4. They clearly show that both OLS and Treatment effect based estimates are downward biased. While the average return from the switching regression model is around 13%, the one based on the treatment effect model is around 11% whereas the average OLS estimate is 7.3%.

A further advantage of these mean estimates is that they could be easily be compared to those based on the hypothesis of a linear wage-schooling relationship. Indeed, the variation in the returns with educational levels discussed above might be only apparent. A simple means of examining the validity of the non-linear hypothesis could consist in conducting t tests the null of which is that the difference between the marginal returns to reaching two different educational levels is zero. The results from such a testing strategy are reported in Table 7. This table simply compares the various estimates that are reported in the "switching regression" column of Table 4. No doubt, both tables show that the estimated returns are systematically pairwise different and that the differences are systematically highly significant. While this suggests that estimates assuming a linear wage-schooling relationship might be heavily biased, it remains nevertheless interesting to evaluate the importance of such a bias.

In general, the conventional approaches to estimating the returns to schooling assuming a linear wage-schooling relationship are either OLS or IV. Table 8 reports such estimates. It confirms the usual evidence that compared to the IV estimate (7.34%), the OLS one (5.35%) is downward biased. Harmon and Walker (1995) have explored a specification where the schooling variable is treated as an integer since endogeneity of schooling decisions is accounted for through a discrete choice model (an ordered probit in their case), but where it is included as a continuous variable in a wage equation that is assumed to be linear in schooling. Another characteristic of Harmon and Walker's (1995) specification is that, in contrast to the treatment effect model, it restricts the coefficients on the selectivity term to be equal whatever individuals' educational levels are. Estimates based on such an approach are also reported in Table 8. These rely on the two specifications reported under the heading "H & W", which differ only in the choice model correction for selectivity is based one: ordered probit versus two-outcome probit. A distinguishing feature of these estimates is that they are in the range of their OLS counterpart. This is in contrast with Harmon and Walker (1995) who find a selectivity corrected estimate in the range of the IV one (around 14% for the UK). However, the main message from Table 8 could be delivered only if it is compared to the estimates reported in Table 4 or at least the mean returns in its bottom row. Indeed, such a comparison makes it clear that, whether endogeneity of schooling is accounted for or not and, however the way this is done, assuming a linear wage-schooling relationship yields estimates of the returns to schooling that are further downward biased.

3.3 Implications of Varying Marginal Returns

The main message from the previous subsection is threefold. First, schooling decisions are not exogenous. Second, the impact of individual endowments on wages depends on the amount invested in education. Third, the wage-schooling relationship is not linear and obeys to no specific functional form. Therefore, as long as education is coded as an integer, it should be treated as a discrete choice variable and only a switching regression model with endogenous switching is appropriate to account for these features simultaneously. The outcome from such an estimation strategy has been discussed in the previous subsection and, in particular, it has been quantitatively compared to the outcome from alternative approaches. Its economic implications remain, however, to be discussed.[17]

First, the pattern of variation of marginal returns across educational levels indicates that a one percentage point decrease in the average return to schooling could be due to either a relatively lower decrease in the return to the first year after compulsory schooling or to a much higher decrease in the return to the first year after high school degree. More specifically, a drop in the average marginal return from 12.77% to 11.77% is compatible with either a drop in the marginal return to level 2 from 15,37% to 7.6% (approximately 50%) or a drop in the marginal return to level 4 from 7.68% to -0.1% (more than 100%). This means that the decrease over time in the returns to education in France that is reported in the literature (Baudelot and Glaude, 1989) is not necessarily uniform and might be due to a variety of combinations of variations in the marginal returns to specific levels. Moreover, the highlighted pattern could also help evaluating the validity of certain arguments aiming at explaining the observed decrease in the average marginal return to schooling. For instance, one of these arguments has been suggested by Baudelot and Glaude (1989) and relates to the effect of minimum wages. Since these are continuously increasing in France, one would expect the gap between wages of highly educated individuals and those who are at the bottom of the wage distribution to be narrower as minimum wages increase, hence reducing the relative value of high educational levels. The relatively high value of the return to schooling level 2 suggests the minimum wage explanation is likely to be valid since an increase in minimum wages is likely to exert its most important negative effect on this return and therefore, to significantly influence the overall average marginal return.

Second, the estimates reported in Table 4 raise theoretical questions as well. It is common practice in the literature to relate non-linearities in the wage-schooling relationship to sheepskin effects, that is, "bonus" returns to finishing a degree or obtaining a diploma. The results in Table 4 suggest that non-linearities occur even at educational levels which do not correspond to typically required numbers of years to finishing specific diplomas. For instance, while level 3 would typically correspond to the high school level, it only yields a return of 6.38% which is not even half the return to level 2 which yields no certification at all. Likewise, across tertiary education levels (levels 4 and higher), the returns steadily inrease with years of schooling, except for the highest level where there is again a decrease. Probably, the non-liearities which emerge from specifiactions including qualifications simply capture part of the non-linearities highlighted by the estimates in Table 4. Moreover, while sheepskin effects are in general interpreted as due to signaling effects, that is to employers interpreting certification as a signal of the ability to persevere or to jump hurdles, the estimates in Table 4 do not seem to be compatible with basic predictions of signaling theory. The latter predicts indeed that potential employers use wage levels as a screening

[17] Unless explicitly mentioned, I shall focus on the interpretation of the estimates from the two-outcome probit based approach.

device. Wages would then be fixed in such a way that the less able would find the return to their investment in signaling not high enough to compensate them for the high cost they would incur if they undertake the investment. Given that the basic assumption in signaling models is that the cost of investing in signaling is a decreasing function of ability, one would expect the marginal returns to be decreasing with educational levels. Otherwise, higher education could no more serve as a filter since some of the less able might also find the investment worth undertaking. It is exactly the opposite situation that is highlighted in Table 4 where the marginal returns to the successive levels of higher education seem to be increasing up to college graduation (level 8). Of course, advocates of screening theory would argue that this increase is compatible with the less able individuals being endowed with an increasing and convex cost function. This would however not answer the question of why would employers pay increasingly higher wages to the abler who are likely to find it worthwile investing even if the returns were lower. In addition, this would not explain nor why the returns to level 2 are the highest ones nor why the returns to high school levels are more than twice lower.

Alternatively, one might think of linking the differences in the successive marginal returns to differences in the balance of supply and demand for each educational level in the labour market. While this might be part of the explanation, it is certainly not the main one. At least, it cannot account for the difference between the return to level 2 and the returns to higher levels since one would expect the market effect to be in favour of the highly educated, not the other way around. Certainly, what determines the observed variation is simply the market value of the extra amount of human capital each level endows individuals with.

It is also worth noting that while investing a single further year after compulsory schooling yields a return of more than 15%, the third of sampled individuals have not undertaken the investment. In general, despite increasing cumulative returns, relatively few individuals reach high educational levels. Probably, individuals who do not undertake investments that reveal to be increasingly profitable ex post, ex ante discount future earnings so heavily that the expected wage differential does not suffice to induce them to undertake the investment. High discount rates might of course be due to severe budget constraints and, in general, to unsatisfactory socio-economic conditions. This means that as human capital theory predicts, while high expected wages are likely to positively influence the amount invested in education, they are not the sole determinant of such a decision. As usually in the literature, the reduced forms schooling equations estimated in Table 2 rely on this idea, that the underlying structural model includes expected wages as a determinant of schooling decisions.

Table 9 reports estimates of the structural schooling equations associated with the IV model and the ordered probit model respectively. When the returns to schooling are estimated using the IV approach, it is common practice in the literature to include the expected wage of individuals, given their schooling level, in the structural form schooling equation. The IV specification in Table 9 does the same and the results suggest that educational attainments are positively correlated to expected wages, a result commonly reported in the literature. When endogeneity is accounted for using the ordered probit model, one could, like Harmon and Walker (1995) do, proceed the same way. I have estimated such specification as well and, although not reported in Table 9, the results again suggest that the higher are expected wages, the more likely are individuals to invest further in education.[18] However, the ordered probit model is meant to model schooling decisions as discrete choices and as such is a generalisation of Willis and Rosen's (1979) model. In the latter framework, only two outcomes are considered since the decision that is modeled is whether to attend college or not. Therefore, what determines individuals' decisions is the difference between the wage level individuals would have earned had they chosen to attend college and the one they would have earned had they left school before. Generalizing such model to the case of more than two choice alternatives would therefore imply modeling the choice of individuals comparing pairwise all investment opportunities that are offered to them. To avoid estimating a likely redundant specification, I exploit the ordered nature of the probability model to be estimated and consider the difference between the wage predicted for each individual given her educational attainment and the predicted wage from the best alternative she would have earned had she left school earlier. The results from such a specification are reported in Table 9 and again show that the extra wage individuals expect from a further investment in education positively influences the probability that they undertake that investment.

Though not reported, the results from the structural schooling specifications underlying the reduced form two-outcome probit equations are worth mentioning. Since each of these is a simple probit equation modeling the probabilitiy that individuals keep on attending school beyond level j, $j = 1, 2, \cdots, 8$, they are similar to Willis and Rosen's (1979) model, except that in the latter case, only the high school degree level is considered. Therefore, the wage variable included in each structural form is constructed in the same way as in Willis and Rosen (1979). For each level S, I compute the mean of the wages each individual can expect from schooling levels $j = 1, \cdots, S$ and from schooling levels $j = S + 1, \cdots, 9$. The differences in logs between the latter and the former measure then the average wage differential between educational levels beyond S and lower ones. Interestingly enough, estimates of the impact of such differentials on the probability to attend school beyond each level $j = 1, 2, \cdots, 8$ are not systematically positive. To be more specific, it is positive only for the highest educational levels (6, 7 and 8) and significantly negative for lower levels. Although this might seem astonishing at first sight, it is in fact consistent with the data. While investments to attain any educational level yield systematic wage advantages, the majority of individuals do not undertake them. What these negative signs reflect is that, at low educational levels, while staying at school yields significant wage gains, the proportion of individuals who decide not to stay is too high with respect to the incentive to keep on attending school. As mentioned above, for these individuals, the negative effect of non observed

[18] These results could of course be made available from the author upon request.

factors such as time discounting or family circumstances is too strong to be compensated by the positive effect of expected wages. In contrast, those having reached high educational levels have proven they discount future earnings differently and are probably willing to make the long investment they have already undertaken as profitable as possible. Therefore, the higher is the return they expect, the more likely they are to be willing to benefit from it.

4 Conclusion

In this paper, the returns to schooling in France have been estimated using a procedure that accounts for the endogeneity of schooling as well as for non-linearities in the wage-schooling relationship. Estimation of a switching regression model with endogenous switching allows modeling individual schooling decisions as a discrete choice process. Moreover, it enables one to explore the way education influences the process of wage determination during individuals' working lives and in particular, age and job tenure-wage profiles. Last but not least, it allows estimation of the returns to schooling without imposing any specific form, whether linear or non-linear, on the wage-schooling relationship.

The results show that these three dimensions are indeed important as they suggest the biases they might lead to are not negligible. On the one hand, a carefull testing strategy suggests that (*i*) the wage-schooling relationship is not linear and obeys to no specific non-linearity scheme, (*ii*) the process of wage determination is conditional on indivisuals' educational levels at least because these are highly correlated not only with starting wages but also with the slopes of age-earnings profiles and (*iii*) whether education is treated as a continuous variable or as an ordered integer, endogeneity is crucial. On the other hand, (*i*) the lowest estimates are obtained when assuming wage determination can be modeled within a common linear wage-schooling relationship whether corrected for endogeneity or not. (*ii*) Allowing the returns to schooling to vary with educational grades systematically suggests the average marginal return is higher whether one uses OLS or estimates a treatment effect model. (*iii*) These estimates are, however, still significantly lower than those suggested by the switching regression model with endogenous switching.

In the latter case, the results suggest that the first year after compulsory schooling yields the highest marginal return while the second year yields the lowest one. Beyond, the marginal returns increase steadily as one moves along educational levels and then goes back to its lowest value for the highest schooling level. This decomposition of the usually reported average marginal return has obvious policy implications in terms of resource allocation and incentives to invest. It is also a means of disciminating between human capital and signaling theories. At least, the highlighted pattern seems to be in contradiction with one of the main predictions of the latter theory.

Appendix

Consider the model described by (6, 7, 8). Suppose that $u_i^j, j = 1, \cdots, J$, and v_i are distributed as $(J+1)$-variate normal. The correlation coefficients of u_i^j, $j = 1, 2, \cdots, J$, with v_i in their respective marginal distributions are ρ^j, $j = 1, 2, \cdots, J$, respectively and the variances of u_i^j, $j = 1, 2, \cdots, J$ and v_i are σ^j and 1, respectively. It follows that the expectations of log-earnings conditional on educational attainments are, for $j = 1, 2, \cdots, J$;

$$E\left(y_i^j \mid \mu^{j-1} < S_i^* \leq \mu^j\right) = X_i'\beta^j + E\left(\mu^{j-1} - Z_i'\gamma < u_i^j \leq \mu^j - Z_i'\gamma\right),$$

where μ^0 and μ^J are taken as $-\infty$ and $+\infty$, respectively.

The properties of conditional expectations of normally distributed variables of the truncated normal distribution (see Maddala, 1983, pp. 365-68) imply that the conditional expectations can then be written as :

$$E\left(y_i^j \mid \mu^{j-1} < S_i^* \leq \mu^j\right) = X_i'\beta^j + \rho^j\sigma^j \frac{\phi\left(\mu^{j-1} - Z_i'\gamma\right) - \phi\left(\mu^j - Z_i'\gamma\right)}{\Phi\left(\mu^j - Z_i'\gamma\right) - \Phi\left(\mu^{j-1} - Z_i'\gamma\right)}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal distribution and density functions, respectively. The latter equations suggest estimating Equations (9) in the text.

Let ϕ^j and Φ^j denote $\phi\left(\mu^j - Z_i'\gamma\right)$ and $\Phi\left(\mu^j - Z_i'\gamma\right)$, respectively. We have:

$$\begin{aligned} E(y_i) &= \sum_{j=1}^{J} E\left(y_i \mid \mu^{j-1} < S_i^* \leq \mu^j\right) \cdot \Pr\left(\mu^{j-1} < S_i^* \leq \mu^j\right) \\ &= \sum_{j=1}^{J} E\left(y_i \mid S_i = j\right) \cdot \Pr\left(S_i = j\right), \end{aligned} \tag{21}$$

which, given the equations in (9), implies that:

$$E(y_i) = \sum_{j=1}^{J} \left(X_i'\beta^j + \rho^j\sigma^j \cdot \frac{\phi^{j-1} - \phi^j}{\Phi^j - \Phi^{j-1}}\right) \cdot \left(\Phi^j - \Phi^{j-1}\right).$$

This latter expression could well be written:

$$\begin{aligned} E(y_i) &= X_i'\beta^J + \sum_{j=1}^{J-1} \left(X_i \cdot \Phi^j\right)' \left(\beta^j - \beta^{j+1}\right) \\ &\quad + \sum_{j=1}^{J-1} \left(\rho^{j+1}\sigma^{j+1} - \rho^j\sigma^j\right) \cdot \phi^j, \end{aligned} \tag{22}$$

which is a simple extension of equation (8.19) in Maddala (1983) to the multiple-regimes framework. Estimation of (A2) using the whole sample is equivalent to separately estimating the J equations in (9). This provides a simple tool to testing the null hypotheses:

$$\beta^j - \beta^{j+1} = 0, \quad \forall j = 1, \cdots, J-1$$

using the t-distribution. Note also that a constrained version of (A2) would be one where the parameters in the β's, except the constant terms, are the same for all schooling levels. In this case, equation (A2) reduces to:

$$\begin{aligned} E(y_i) &= X_i'\beta + \sum_{j=1}^{J-1} \left(\Phi^j\right)' \left(c^j - c^{j+1}\right) \\ &\quad + \sum_{j=1}^{J-1} \left(\rho^{j+1}\sigma^{j+1} - \rho^j\sigma^j\right) \cdot \phi^j, \end{aligned} \tag{23}$$

where, for $j = 1, \cdots, J$, c^j denotes the constant term specific to schooling level j. Equation (20) is again an extension to the multiple-regime framework of the treatment effect model (see Maddala's, 1983, equation (8.20)), where marginal returns to schooling are simply $c^{j+1} - c^j$ for $j = 1, \cdots, J-1$. Therefore, as long as the null hypotheses:

$$c^j - c^{j+1} = 0, \quad \forall j = 1, \cdots, J-1$$

are to be rejected, (20) is a non-linear version of the model estimated by Harmon and Walker (1995). Although both models assume the returns to all the characteristics in X are the same for all schooling levels, the distinguishing trait is that, here, the returns to schooling are allowed to vary from one schooling level to another. In fact, (20) is the selectivity-corrected version of a model where each schooling level would be described by a specific dummy variable.

One problem with (22) is that, besides selectivity terms, it includes a number of regressors equal to the number of components of the vector X, times the number of schooling levels. Not only is the output from the regression cumbersome, but such a number of regressors is likely to weaken the statistical properties of the estimates. Fortunately, rather than (A1), one could alternatively use:

$$\begin{aligned} E(y_i \mid S_i = j \cup S_i = j+1) &= E(y_i \mid S_i = j) \cdot \Pr(S_i = j) + \\ &\quad E(y_i \mid S_i = j+1) \cdot \Pr(S_i = j+1) \end{aligned}$$

for $j = 1, \cdots, J-1$, which, given (9), yields Equations (19) in the text and their respective constrained versions (20).

References

[1] Angrist, J. D. and A. B. Krueger, "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, November 1991, 106(4), pp. 979-1014.

[2] Angrist, J. D. and A. B. Krueger, "Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery." National Bureau of Economic Research Working Paper N° 4067, 1992.

[3] Arrow, K. J., "Higher Education as a Filter," *Journal of Public Economics*, 1973, 2(3), pp. 193-216.

[4] Ashenfelter, O. and A. B. Krueger, "Estimates of the Economic Return to Schooling from a New Sample of Twins." *American Economic Review*, December 1994, 84(5), pp. 1157-73.

[5] Ashenfelter, O., C. Harmon and H. Oosterbeek, "A review of estimates of the schooling/earnings relationship, with tests for publication bias", *Labour Economics,* November 1999, 6(4), pp. 453-70.

[6] Ashenfelter, O. and D. J. Zimmernan, "Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons and Brothers." National Bureau of Economic Research Working Paper N° 4491, 1993.

[7] Baudelot, C. and M. Glaude, "Les diplômes se dévaluent-ils en se multipliant.", *Economie et Statistique,* 1989, pp. 3-16.

[8] Blackburn, M. L. and D. Neumark, "Omitted Ability Bias and the Increase in the Return to Schooling." *Journal of Labor Economics*, July 1993, 11(3), pp. 521-43.

[9] Blanchflower, D. G. and P. Elias, "Ability, Schooling and Earnings: Are Twins Different?" Mimeo, Darmouth College, U.S., 1993.

[10] Brunello, G. and R. Miniaci, "The economic returns to schooling for Italian men. An evaluation based on instrumental variables", *Labour Economics,* November 1999, 6(4), pp. 509-20.

[11] Butcher, K. F. and A. Case, "The Effect of Sibling Composition on Women's Education and Earnings." *Quarterly Journal of Economics*, Fall 1994, 109(3), pp. 531-63.

[12] Card, D., "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." National Bureau of Economic Research Working Paper N° 4483, 1993.

[13] Card, D., "The Causal Efect of Education on Earnings," in Ashenfelter, O. and D. Card eds, *The Handbook of Labor Economics*, 2000, vol. 3, part 7, 1801-63.

[14] Chevalier, A. and I. Walker, "Chapter 16: United Kingdom" in Harmon, C., I. Walker and N. Wetergärd-Nielsen eds., *Education and Earnings in Europe: A cross Country Analysis of the Returns to Education*, Edward Elgar, Cheltenham, UK, 2001, pp. 302-30.

[15] Denny, K. J. and C. P. Harmon, "Testing for Sheepskin Effects in Earnings Equations: Evidence from Five Countries." *Applied Economics Letters, 2001.*

[16] Ermisch, J. F. and R. E. Wright, "Wage Offers and Full-Time and Part-Time Employment by British Women." *The Journal of Human Resources,* Winter 1993, 28(1), pp. 111-33.

[17] Garen, J., "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable." *Econometrica*, September 1984, 52(5), pp. 1199-1218.

[18] Guillottin, Y. and P. Sevestre, "Estimations de fonctions de gains sur données de panel: Endogénéité du capital humain et effets de la sélection, *Economie et Prévision*, 1994, 116, pp. 119-35.

[19] Harmon, C. and I. Walker, "Estimates of the Economic Return to Schooling for the United Kingdom." *American Economic Review*, 1995, 85(5), pp. 1278-86.

[20] Harmon, C., I. Walker and N. Wetergärd-Nielsen eds., *Education and Earnings in Europe: A cross Country Analysis of the Returns to Education*, Edward Elgar, Cheltenham, UK, 2001.

[21] Heckman, J. J., "Sample Selection Bias as a Specification Error." *Econometrica*, January 1979, 47(1), pp. 153-61.

[22] Kalwij, A., "Estimating the Economic Returns to Schooling on the Basis of Panel Data", Mimeo, Tilburg University,1996.

[23] Kane, T. J. and C. E. Rouse, "Labour Market Returns to Two- and Four-Year College: Is a Credit a Credit and do Degrees Matter? " *American Economic Review*, 1995, 85, pp. 600-14.

[24] Layard, R. and G. Psacharopoulos, "The Screening Hypothesis and the Return to Education", *Journal of Political Economy*, 1974, 82(5), pp. 985-98.

[25] Levin, J. and E. J. S., Plug, "Instrumenting education and the returns to schooling in the Netherlands", *Labour Economics,* November 1999, 6(4), pp. 521-34.

[26] Lhéritier, J. L., "Les déterminants du salaire", *Economie et Statistique*, 1994, 257, pp. 9-22.

[27] Lollivier, S. and J-F. Payen, "L'hétérogénéité des carrières individuelles mesurée sur données de panel", *Economie et Prévision*, 1990, 92/93, pp. 87-95.

[28] Maddala, G. S., *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press, 1983.

[29] Nakamura, A. and M. Nakamura, "Part-Time and Full-Time Work Behaviour of Married Women: A Model with a Doubly Truncated Dependent Variable." *Canadian Journal of Economics*, 16(2), pp. 229-57.

[30] Park, J. H., "Estimation of Sheepskin Effects using the Old and New Measures of Educational Attainment in the Current Population Survey", *Economics Letters*, 1999, 62(12), pp. 237-40.

[31] Riley, J. G., "Testing the Educational Screening Hypothesis." *Journal of Political Economy*, October 1979, Part 2, 87(5), pp. S227-S252.

[32] Sharp, C., "A study of patterns of school entry and the impact of season of birth on school attainment", *Education Research*, 37, pp. 251-65.

[33] Spence, M., "Job Market Signaling", *Quarterly Journal of Economics*, 1973, 87(3), pp. 355-74.

[34] Spence, M., *Market Signaling*, Cambridge, Harvard University Press,1974.

[35] Uusitalo, R. "Returns to Education in Finland", *Labour Economics,* November 1999, 6(4), pp.569-80.

[36] Vella, F. and R. G. Gregory, "Selection Bias and Human Capital Investment: Estimating the Rates of Return to Education for Young Males", *Labour Economics*, September 1996, 3(2), pp. 197-219.

[37] Willis, R. J. and S. Rosen, "Education and Self-Selection." *Journal of Political Economy*, October 1979, Part 2, 87(5), pp. S7-S36.

Table 1. Descriptive statistics.

	$S = 1$	$S = 2$	$S = 3$	$S = 4$	$S = 5$	$S = 6$	$S = 7$	$S = 8$	$S = 9$	All
Mean and standard deviation of continuous variables										
Monthly wages	7008.35	7620.73	7757.11	8048.96	8739.54	9144.00	10140.26	11619.58	12838.82	8245.75
	(1873.76)	(2233.3)	(2456.51)	(2725.69)	(3150.8)	(3416.98)	(4010.35)	(4838.18)	(5289.57)	(3337.05)
Number of years of schooling	8.72	11.00	12.00	13.00	14.00	15.00	16.00	17.00	19.70	11.99
	(1.29)	-	-	-	-	-	-	-	(2.04)	(3.40)
Age	42.13	38.36	36.29	35.21	35.79	34.53	35.57	36.51	39.38	38.62
	(9.66)	(9.20)	(9.20)	(9.43)	(9.64)	(9.14)	(9.16)	(9.42)	(9.46)	(9.85)
Job tenure	13.24	12.11	10.69	9.95	9.88	8.67	8.56	8.83	8.93	11.29
	(10.03)	(9.57)	(9.03)	(8.90)	(8.97)	(8.44)	(8.33)	(8.63)	(8.61)	(9.51)
Number of children	1.46	1.36	1.34	1.26	1.23	1.15	1.11	1.09	1.44	1.33
	(1.28)	(1.14)	(1.13)	(1.14)	(1.13)	(1.14)	(1.125)	(1.15)	(1.16)	(1.2)
Interest rate	7.61	8.834	9.36	9.17	8.90	8.96	8.90	8.75	8.79	8.53
	(2.59)	(2.95)	(2.98)	(2.80)	(2.74)	(2.70)	(2.63)	(2.60)	(2.62)	(2.84)
Frequencies of qualitative variables										
Married	0.7233	0.6766	0.6124	0.5681	0.5654	0.5263	0.5402	0.5678	0.6264	0.6459
Single	0.2140	0.2695	0.3362	0.3876	0.3921	0.4390	0.4174	0.3967	0.3282	0.3018
Widowed	0.0076	0.0039	0.0038	0.0028	0.0033	0.0038	0.0031	0.0024	0.0038	0.0050
Divorced	0.0550	0.0550	0.0476	0.0414	0.0392	0.0309	0.0393	0.0330	0.0416	0.0474
Immig. with citizenship	0.0211	0.0144	0.0136	0.0141	0.0150	0.0162	0.0166	0.0159	0.0335	0.0183
Immig. from western countries	0.0538	0.0146	0.0163	0.0135	0.0115	0.0099	0.0095	0.0109	0.0213	0.0282
Immig. from other countries	0.0451	0.0177	0.0205	0.0200	0.0240	0.0186	0.0180	0.0189	0.0370	0.0299
French native	0.8800	0.9532	0.9496	0.9523	0.9495	0.9553	0.9560	0.9542	0.9082	0.9236
Tenured	0.9378	0.9334	0.9217	0.9010	0.8972	0.8846	0.9004	0.9127	0.9088	0.9218
Interim	0.0195	0.0189	0.0232	0.0279	0.0255	0.0293	0.0175	0.0117	0.0074	0.0203
Apprentice	0.0024	0.0031	0.0031	0.0059	0.0057	0.0065	0.0063	0.0055	0.0033	0.0036
Beginner	0.0059	0.0071	0.0085	0.0108	0.0122	0.0143	0.0164	0.0153	0.0163	0.0093
Fixed Term Contract	0.0280	0.0312	0.0370	0.0468	0.0512	0.0597	0.0502	0.0461	0.0515	0.0378
Father was public sector emp.	0.1189	0.1602	0.1766	0.2025	0.2205	0.2316	0.2266	0.2513	0.2709	0.1735

Father was private sector emp.	0.6390	0.6421	0.6262	0.5951	0.5676	0.5566	0.5455	0.5184	0.4897	0.6060
Father self-employed , no emp.	0.1746	0.1377	0.1290	0.1265	0.1217	0.1220	0.1164	0.1144	0.1117	0.1427
Father self-emp., - than 10 emp.	0.0647	0.0554	0.0622	0.0692	0.0797	0.0792	0.0961	0.0959	0.1077	0.0707
Father self-emp., + than 10 emp.	0.0028	0.0028	0.0060	0.0067	0.0105	0.0106	0.0154	0.0200	0.0200	0.0071
Born in the Great Paris	0.0898	0.1070	01145	0.1250	0.1352	0.1379	0.1547	0.1738	0.1786	0.1167
Born in the Centre of France	0.2189	0.2113	0.2054	0.1910	0.1739	0.1807	0.1739	0.1602	0.1459	0.1992
Born in the North of France	0.0754	0.0785	0.0797	0.0817	0.0745	0.0792	0.0718	0.0707	0.0587	0.0756
Born in the East of France	0.1092	0.1307	0.1087	0.0991	0.1008	0.1111	0.1077	0.1040	0.0950	0.1095
Born in the West of France	0.1433	0.1352	0.1385	0.1480	0.1341	0.1399	0.1303	0.1233	0.1023	0.1366
Born in the South-west	0.0807	0.0958	0.0953	0.1013	0.1030	0.0932	0.0902	0.0877	0.0854	0.0900
Born in the Centre-East	0.0781	0.1004	0.1019	0.0968	0.1129	0.1094	0.1121	0.1122	0.1006	0.0948
Born in south-East	0.053	0.0619	0.0701	0.0691	0.0693	0.0642	0.0696	0.0762	0.0752	0.0632
Lives in the Great Paris	0.1463	0.1349	0.1478	0.1642	0.1789	0.1931	0.2223	0.2767	0.3073	0.1706
Lives in the Centre of France	0.2342	0.2194	0.2162	0.2047	0.1859	0.1854	0.1799	0.1606	0.1598	0.2109
Lives in the North of France	0.0686	0.0653	0.064	0.0643	0.0593	0.0627	0.0515	0.0533	0.0435	0.0630
Lives in the East of France	0.1284	0.1374	0.1138	0.1049	0.1074	0.1143	0.1082	0.097	0.0931	0.1186
Lives in the West of France	0.1395	0.1299	0.1333	0.1418	0.1273	0.1330	0.1158	0.1120	0.0936	0.1308
Lives in the South-west	0.0917	0.1044	0.1054	0.1094	0.1117	0.104	0.0975	0.0965	0.0942	0.0997
Lives in the Centre-East	0.1043	0.1189	0.1206	0.1170	0.1312	0.1256	0.1335	0.1224	0.1158	0.1156
Lives in south-East	0.0871	0.0898	0.0990	0.0937	0.0974	0.0819	0.0907	0.0816	0.0926	0.0909
Born in winter	0.2551	0.2589	0.2646	0.2492	0.2523	0.2540	0.2556	0.2540	0.2412	0.2555
Born in Spring	0.2490	0.2533	0.2657	0.2655	0.2597	0.2716	0.2660	0.2716	0.2569	0.2576
Born in Summer	0.2492	0.2513	0.2482	0.2500	0.2569	0.2482	0.2402	0.2392	0.2588	0.2499
Born in Autumn	0.2467	0.2365	0.2214	0.2354	0.2310	0.2262	0.2383	0.2350	0.2431	0.237
School leaving age, 16	0.8116	0.9504	0.9682	0.9801	0.9828	0.9924	0.9918	0.9912	0.9937	0.9179
Number of observations	104,882	41,100	52,737	22,903	20,316	13,590	12,608	10,298	22,834	301,268

Table 2. Reduced form schooling equations using the selectivity and the IV approaches.

	IV (2-SLS) model		Ordered Probit model	
	Coefficients	Std. Error	Coefficients	Std. Error
Age	-0.2273	(0.0057)	-0.1019	(0.0021)
Age squared	0.0031	(0.0001)	0.0014	(0.0000)
Job tenure	-0.0623	(0.0022)	-0.0175	(0.0008)
Job tenure squared	0.0010	(0.0001)	0.0003	(0.0000)
Number of children	-0.2106	(0.0051)	-0.0817	(0.0018)
Married	0.4321	(0.0270)	0.1603	(0.0098)
Single	0.5460	(0.0292)	0.1744	(0.0105)
Widowed	0.0598	(0.0832)	0.0150	(0.0318)
Immigrants with citizenship	-0.0638	(0.0438)	-0.0450	(0.0158)
Immigrants from western countries	-2.4375	(0.0379)	-0.7555	(0.0150)
Immigrants from other countries	-1.3476	(0.0381)	-0.3838	(0.0141)
Tenured	-0.2830	(0.0291)	-0.0607	(0.0101)
Temporary worker	-0.9437	(0.0477)	-0.2756	(0.0167)
Trainee	-0.1075	(0.0959)	-0.0025	(0.0330)
Beginner	0.3532	(0.0637)	0.1284	(0.0221)
Father was public sector employee	-1.1631	(0.0677)	-0.3618	(0.0234)
Father was private sector employee	-2.1941	(0.0678)	-0.7162	(0.0230)
Father self-employed , no employees	-2.2351	(0.0681)	-0.7183	(0.0236)
Father self-employed, less than 10 employees	-1.2616	(0.0697)	-0.3934	(0.0241)
Born in the Great Paris	0.9754	(0.0262)	-0.0781	(0.0092)
Born in the Centre of France	0.1638	(0.0276)	-0.2405	(0.0092)
Born in the North of France	0.0644	(0.0421)	-0.0783	(0.0135)
Born in the East of France	-0.0600	(0.0334)	-0.0805	(0.0116)
Born in the West of France	0.1190	(0.0326)	-0.1711	(0.0108)
Born in the South-west of France	0.2527	(0.0322)	-0.0907	(0.0111)
Born in the Centre-East of France	0.3491	(0.0307)	-0.0430	(0.0109)
Interest rate of the school leaving year	0.1584	(0.0024)	0.0500	(0.0008)
Born in winter	-0.0182	(0.0159)	0.0061	(0.0057)
Born in Spring	0.0619	(0.0158)	0.0351	(0.0056)
Born in Summer	0.0310	(0.0159)	0.0151	(0.0057)
School leaving age was 16	3.6306	(0.0288)	1.4792	(0.0122)
Intercept	13.4890	(0.1272)	-0.8804	(0.0449)
Cut 2	-	-	0.2247	(0.0022)
Cut 3	-	-	0.4396	(0.0028)
Cut 4	-	-	0.6323	(0.0031)
Cut 5	-	-	0.8764	(0.0034)
Cut 6	-	-	1.1177	(0.0035)
Cut 7	-	-	1.6204	(0.0038)
Cut 8	-	-	2.0207	(0.0040)
Likelihood ratio	-	-	-538332.16	
Hausman F for overidentification	59.22	Pr < 0.0001	-	-
Adjusted R squared	0.18		-	
Number of observations	297,599			

Note : The coefficients on 7 dummies for region of residence and 10 year dummies are not reported in the table although included in the regressions.

Table 3. Selectivity corrected earnings equations using the ordered probit model.

	$J = 1$	$J = 2$	$J = 3$	$J = 4$	$J = 5$	$J = 6$	$J = 7$	$J = 8$	$J = 9$
Intercept	8.1169	8.0571	8.0292	7.8782	7.7646	7.6123	7.4335	7.6560	7.7820
	(0.0146)	(0.0231)	(0.0205)	(0.0322)	(0.0391)	(0.0536)	(0.0648)	(0.0850)	(0.0670)
Age	0.0237	0.0300	0.0335	0.0420	0.0522	0.0580	0.0739	0.0647	0.0593
	(0.0006)	(0.0011)	(0.0010)	(0.0016)	(0.0020)	(0.0027)	(0.0032)	(0.0040)	(0.0029)
Age squared / 100	-0.0259	-0.0291	-0.0315	-0.0397	-0.0512	-0.0545	-0.0746	-0.0618	-0.0056
	(0.0008)	(0.0014)	(0.0013)	(0.0022)	(0.0025)	(0.0035)	(0.0040)	(0.0049)	(0.0039)
Job tenure	0.0087	0.0094	0.0102	0.0108	0.0079	0.0108	0.0070	0.0042	0.0034
	(0.0003)	(0.0000)	(0.0005)	(0.0007)	(0.0009)	(0.0012)	(0.0014)	(0.0017)	(0.0012)
Job tenure sq. / 100	-0.0059	-0.0034	-0.0069	-0.0093	-0.0038	-0.0232	-0.0100	-0.0027	-0.0061
	(0.0008)	(0.0014)	(0.0015)	(0.0024)	(0.0029)	(0.0040)	(0.0046)	(0.0056)	(0.0039)
Nb. of children	-0.0060	-0.0039	-0.0058	0.0000	-0.0032	-0.0027	0.0063	-0.0029	0.0159
	(0.0006)	(0.0011)	(0.0011)	(0.0016)	(0.0019)	(0.0023)	(0.0028)	(0.0034)	(0.0025)
Married	0.0223	0.0287	0.0336	0.0468	0.0361	0.0573	0.0452	0.0903	0.0809
	(0.0032)	(0.0056)	(0.0052)	(0.0086)	(0.0102)	(0.0138)	(0.0142)	(0.0190)	(0.0125)
Single	-0.0774	-0.0437	-0.0518	-0.0344	-0.0576	-0.0458	-0.0496	-0.0473	-0.0704
	(0.0036)	(0.0060)	(0.0056)	(0.0091)	(0.0109)	(0.0145)	(0.0151)	(0.0199)	(0.0132)
Widowed	0.0137	-0.0568	0.0288	-0.0181	-0.0106	-0.0240	0.0460	0.0090	0.1001
	(0.0086)	(0.0193)	(0.0180)	(0.0324)	(0.0343)	(0.0400)	(0.0492)	(0.0694)	(0.0416)
Immig. with citiz.	-0.0291	-0.0750	-0.0585	-0.1182	-0.1279	-0.1081	-0.1459	-0.2137	-0.2118
	(0.0049)	(0.0098)	(0.0093)	(0.0140)	(0.0158)	(0.0183)	(0.0208)	(0.0261)	(0.0133)
Imm. West. count.	-0.0132	0.0043	0.0250	0.0141	-0.0461	0.0712	0.1063	0.0959	0.1149
	(0.0035)	(0.0102)	(0.0090)	(0.0151)	(0.0189)	(0.0248)	(0.0283)	(0.0334)	(0.0178)
Imm. other count.	-0.1430	-0.1611	-0.1768	-0.2110	-0.3069	-0.3066	-0.4192	-0.3045	-0.3414
	(0.0037)	(0.0092)	(0.0079)	(0.0123)	(0.0131)	(0.0176)	(0.0207)	(0.0244)	(0.0131)
Tenured	0.0665	0.0586	0.0531	0.0535	0.0721	0.1043	0.1520	0.1802	0.1720
	(0.0042)	(0.0066)	(0.0057)	(0.0079)	(0.0088)	(0.0100)	(0.0122)	(0.0154)	(0.0106)
Temporary	0.0547	0.0299	0.0481	0.0451	0.0418	0.0282	0.0310	-0.0460	-0.1005
	(0.0065)	(0.0106)	(0.0089)	(0.0125)	(0.0146)	(0.0166)	(0.0232)	(0.0334)	(0.0293)
Trainee	-0.1424	-0.1299	-0.1682	-0.1725	-0.1655	-0.1984	-0.1410	-0.1976	-0.2822
	(0.0149)	(0.0218)	(0.0198)	(0.0227)	(0.0263)	(0.0299)	(0.0350)	(0.0457)	(0.0418)
Beginner	0.0211	0 .0010	-0.0149	-0.0257	0.0003	0.0265	0.0547	0.0820	0.0932
	(0.0099)	(0.0152)	(0.0127)	(0.0174)	(0.0189)	(0.0214)	(0.0235)	(0.0298)	(0.0210)
Lambda	-0.0878	-0.0278	-0.0283	-0.0379	-0.0432	-0.0452	-0.0690	-0.0765	-0.0457
	(0.0061)	(0.0041)	(0.0040)	(0.0066)	(0.0076)	(0.0099)	(0.0115)	(0.0134)	(0.0108)
N. of observations	103,769	40,757	52,256	22,710	20,111	13,484	12,449	10,182	21,878
Adjusted R squared	0.22	0.32	0.36	0.43	0.43	0.47	0.43	0.39	0.31

Note : The coefficients on 7 dummies for the region of residence and 10 year dummies are not reported in the table although included in the regressions.

Table 4. Estimates of the returns to schooling assuming the wage-schooling relationship is not necessarily linear.

j	Marginal returns					Cumulative returns				
	OLS	Switching Regressions		Treatment Effect		OLS	Switching Regressions		Treatment Effect	
		Ordered	Simple	Ordered	Simple		Ordered	Simple	Ordered	Simple
2	0.1062 (0.0016)	0.1595 (0.0001)	0.1537 (0.0001)	0.2677 (0.0038)	0.2610 (0.0029)	0.1062 (0.0016)	0.1595 (0.0001)	0.1537 (0.0001)	0.2678 (0.0038)	0.2610 (0.0029)
3	0.0472 (0.0017)	0.0560 (0.0000)	0.0638 (0.0000)	0.0406 (0.0040)	0.0759 (0.0020)	0.1534 (0.0014)	0.2252 (0.0001)	0.2280 (0.0001)	0.3083 (0.0042)	0.3368 (0.0031)
4	0.0537 (0.0021)	0.0782 (0.0001)	0.0768 (0.0001)	0.0422 (0.0060)	0.0634 (0.0029)	0.2070 (0.0020)	0.3324 (0.0001)	0.3234 (0.0002)	0.3505 (0.0063)	0.4002 (0.0041)
5	0.0685 (0.0025)	0.0892 (0.0001)	0.0752 (0.0001)	0.0811 (0.0086)	0.0734 (0.0046)	0.2755 (0.0021)	0.4412 (0.0002)	0.4234 (0.0002)	0.4316 (0.0076)	0.4736 (0.0050)
6	0.0624 (0.0029)	0.0962 (0.0001)	0.0861 (0.0001)	0.0438 (0.0122)	0.0760 (0.0069)	0.3379 (0.0024)	0.5823 (0.0003)	0.5479 (0.0003)	0.4754 (0.0109)	0.5494 (0.0068)
7	0.0756 (0.0032)	0.1340 (0.0001)	0.1020 (0.0001)	0.1244 (0.0164)	0.1215 (0.0090)	0.4135 (0.0025)	0.7961 (0.0004)	0.7066 (0.0004)	0.5998 (0.0137)	0.6709 (0.0076)
8	0.1071 (0.0035)	0.1572 (0.0001)	0.1121 (0.0001)	0.1758 (0.0214)	0.1603 (0.0109)	0.5206 (0.0027)	1.0796 (0.0005)	0.8981 (0.0004)	0.7756 (0.0179)	0.8312 (0.0091)
9	0.0659 (0.0031)	-0.0039 (0.0001)	0.0658 (0.0001)	0.0636 (0.0236)	0.0975 (0.0127)	0.5865 (0.0020)	1.0691 (0.0005)	1.0218 (0.0005)	0.8392 (0.0172)	0.9287 (0.0099)
M	0.0733 (0.0020)	0.1336 (0.0005)	0.1277 (0.0005)	0.1049 (0.0172)	0.1161 (0.0099)	- -	- -	- -	- -	- -

Note : Standard errors in parentheses. OLS estimates estimates of the cumulative returns are from specifications including 8 schooling level dummies. OLS marginal returns are obtained from the same specification, except that the omitted group is replaced according to the marginal return to be estimated. The M row reports the mean of the corresponding column .

Table 5. Differences between level *j* and level *j* – 1, in the returns to age and job tenure. Estimates based on the ordered probit selectivity model.

	2 vs. 1	3 vs. 2	4 vs. 3	5 vs. 4	6 vs. 5	7 vs. 6	8 vs. 7	9 vs. 8
Age	0.0063 (0.0013)	0.0035 (0.0015)	0.0085 (0.0019)	0.0102 (0.0025)	0.0058 (0.0033)	0.0159 (0.0042)	-0.0092 (0.0050)	-0.0054 (0.0050)
Age squared / 100	-0.0032 (0.0016)	-0.0024 (0.0019)	-0.0082 (0.0025)	-0.0115 (0.0033)	-0.0034 (0.0043)	-0.0200 (0.0053)	0.0128 (0.0063)	0.0057 (0.0062)
Tenure	0.0007 (0.0005)	0.0008 (0.0006)	0.0006 (0.0009)	-0.0029 (0.0011)	0.0030 (0.0015)	-0.0038 (0.0018)	-0.0028 (0.0021)	-0.0008 (0.0021)
Tenure squared / 100	0.0025 (0.0016)	-0.0034 (0.0020)	-00025 (0.0028)	0.0055 (0.0038)	-0.0193 (0.0050)	0.0131 (0.0061)	0.0073 (0.0072)	-0.0034 (0.0070)

Standard deviations in parentheses.

Table 6. Spearman rank correlation coefficients matrix. (Ordered probit selectivity model above the diagonal and two-outcome probit selectivity model below the diagonal)

	Schooling	Marginal returns	Starting wages	Slope	Peaking age
Schooling	1 (0.0000)	0.0500 (0.8984)	1 (0.0000)	0.9500 (0.0000)	0.1667 (0.6682)
Marginal returns	0.2667 (0.4879)	1 (0.0000)	0.0500 (0.8984)	0.2667 (0.4879)	-0.1500 (0.7001)
Starting wages	1 (0.0000)	0.2667 (0.4879)	1 (0.0000)	0.9500 (0.0000)	0.1667 (0.6682)
Slope	0.9500 (0.0000)	0.4167 (0.2646)	0.9500 (0.0000)	1 (0.0000)	0.0833 (0.8312)
Peaking age	0.2000 (0.6059)	-0.1167 (0.7650)	0.2000 (0.6059)	0.0333 (0.9322)	1 (0.0000)

Significance level in parentheses.

Table 7. *t*-tests of the equality of marginal returns to schooling as estimated from the ordered probit based switching regression model (above the diagonal) and from the two-outcome probit based switching regression model (below the diagonal).

	2	3	4	5	6	7	8	9
2		-0.1035 < 0.0001	-0.0813 < 0.0001	-0.0703 < 0.0001	-0.0633 < 0.0001	-0.0255 < 0.0001	-0.0023 < 0.0001	-0.1634 < 0.0001
3	-0.0898 < 0.0001		0.0222 < 0.0001	0.0331 < 0.0001	0.0401 < 0.0001	0.0780 < 0.0001	0.1012 < 0.0001	-0.0600 < 0.0001
4	-0.0769 < 0.0001	0.0130 < 0.0001		0.0110 < 0.0001	0.0180 < 0.0001	0.0558 < 0.0001	0.0790 < 0.0001	-0.0821 < 0.0001
5	-0.0785 < 0.0001	0.0113 < 0.0001	-0.0016 < 0.0001		0.0070 < 0.0001	0.0445 < 0.0001	0.0680 < 0.0001	-0.0931 < 0.0001
6	-0.0675 < 0.0001	0.0223 < 0.0001	0.0093 < 0.0001	0.0110 < 0.0001		0.0378 < 0.0001	0.0610 < 0.0001	-0.1001 < 0.0001
7	-0.0517 < 0.0001	0.0381 < 0.0001	0.0252 < 0.0001	0.0268 < 0.0001	0.0158 < 0.0001		0.0232 < 0.0001	-0.1379 < 0.0001
8	-0.0415 < 0.0001	0.0483 < 0.0001	0.0353 < 0.0001	0.0370 < 0.0001	0.0260 < 0.0001	0.0102 < 0.0001		-0.1611 < 0.0001
9	-0.0879 < 0.0001	0.0020 < 0.0001	-0.0110 < 0.0001	-0.0094 < 0.0001	-0.0203 < 0.0001	-0.0362 < 0.0001	-0.0463 < 0.0001	

Note : Above (below) the diagonal, each cell (*i,j*) reports the difference in marginal returns $r^j - r^i$ ($r^j - r^i$) and, in parentheses, the corresponding marginal significance level.

Table 8. Estimates of the returns to schooling assuming a linear wage-schooling relationship.

	OLS	IV	H & W	
			Ordered	Simple
Intercept	7.2444	7.0176	7.3066	7.2565
	(0.0094)	(0.0113)	(0.0099)	(0.0100)
Years of schooling	0.0535	0.0734	0.0472	0.0527
	(0.0002)	(0.0005)	(0.0004)	(0.0003)
Age	0.0386	0.0366	0.0387	0.0386
	(0.0004)	(0.0005)	(0.0004)	(0.0004)
Age squared / 100	-0.0382	-0.0340	-0.0387	-0.0382
	(0.0005)	(0.0001)	(0.0005)	(0.0005)
Job tenure	0.0086	0.0098	0.0083	0.0086
	(0.0002)	(0.0000)	(0.0002)	(0.0004)
Job tenure squared / 100	-0.0070	-0.0090	-0.0065	-0.0069
	(0.0006)	(0.0000)	(0.0006)	(0.0006)
Number of children	-0.0037	0.0006	-0.0048	-0.0039
	(0.0004)	(0.0005)	(0.0004)	(0.0004)
Married	0.0386	0.0315	0.0406	0.0388
	(0.0023)	(0.0024)	(0.0023)	(0.0023)
Single	-0.0603	-0.0685	-0.0575	-0.0599
	(0.0025)	(0.0026)	(0.0025)	(0.0025)
Widowed	0.0072	0.0088	0.0073	0.0057
	(0.0072)	(0.0074)	(0.0072)	(0.0072)
Immigrant with citizenship	-0.0819	-0.0827	-0.0808	-0.0812
	(0.0036)	(0.0037)	(0.0036)	(0.0036)
Immig. From western countries	0.0243	0.0767	0.0120	0.0229
	(0.0030)	(0.0033)	(0.0030)	(0.0031)
Immigrant from other countries	-0.1804	-0.1555	-0.1866	-0.1806
	(0.0029)	(0.0031)	(0.0030)	(0.0030)
Tenured	0.0909	0.0962	0.0895	0.0908
	(0.0025)	(0.0026)	(0.0025)	(0.0025)
Temporary worker	0.0409	0.0623	0.0350	0.0396
	(0.0041)	(0.0043)	(0.0041)	(0.0041)
Trainee	-0.1708	-0.1661	-0.1718	-0.1715
	(0.0083)	(0.0085)	(0.0083)	(0.0083)
Beginner	0.0236	0.0172	0.0255	0.0240
	(0.0055)	(0.0057)	(0.0055)	(0.0055)
Heckman's Lambda	-	-	0.0222	0.0039
	-	-	(0.0012)	(0.0011)
Number of observations	298,052	297,599	297,599	297,087
Adjusted R squared	0.45	0.30	0.45	0.45

Note : The coefficients on 7 dummies for the region of residence and 10 year dummies are not reported in the table although included in the regressions. H & W refers to Harmon and Walker's (1995) specification.

Table 9. Structural form schooling equations using the selectivity and the IV approaches.

	IV (2-SLS) model		Ordered Probit model	
	Coefficients	Std. Error	Coefficients	Std. Error
Expected wage	8.7139	0.0132	-	-
Expected lowest wage differential	-	-	5.8499	0.0328
Father was public sector employee	-0.8103	0.0444	-0.3972	0.0232
Father was private sector employee	-1.0099	0.0439	-0.7420	0.0229
Father self-employed , no employees	-1.3234	0.0447	-0.7678	0.0234
Father self-employed, less than 10 employees	-0.7688	0.0457	-0.4331	0.0240
Born in the Great Paris	-0.0751	0.0139	0.1696	0.0075
Born in the Centre of France	-0.0159	0.0120	-0.1679	0.0066
Born in the North of France	0.2546	0.0159	-0.0905	0.0087
Born in the East of France	-0.1603	0.0141	-0.0899	0.0077
Born in the West of France	0.2438	0.0132	-0.1200	0.0073
Born in the South-west of France	0.3579	0.0150	-0.0256	0.0082
Born in the Centre-East of France	0.2260	0.0147	0.0703	0.0080
Interest rate of the school leaving year	0.2392	0.0013	0.0462	0.0007
Born in winter	-0.0054	0.0104	-0.0005	0.0057
Born in Spring	0.0551	0.0104	0.0347	0.0057
Born in Summer	0.0052	0.0105	0.0141	0.0057
School leaving age was 16	2.5595	0.0138	1.2077	0.0100
Intercept	-69.5049	0.1293	-2.7440	0.0258
Cut 2	-	-	0.1973	0.0019
Cut 3	-	-	0.4035	0.0025
Cut 4	-	-	0.5946	0.0029
Cut 5	-	-	0.8389	0.0032
Cut 6	-	-	1.0814	0.0034
Cut 7	-	-	1.5757	0.0037
Cut 8	-	-	2.0065	0.0040
Likelihood ratio	-		-531,525.06	
Adjusted R squared	0.65		-	
Number of observations	297,559			

Note : The coefficients on 7 dummies for region of residence and 10 year dummies are not reported in the table although included in the regressions.