RISK AVERSION AND HUMAN CAPITAL INVESTMENT: A STRUCTURAL ECONOMETRIC MODEL*

Thomas Brodaty**
Université de Cergy-Pontoise, THEMA
33, boulevard du Port, 95011 Cergy-Pontoise cedex, France

Robert J. Gary-Bobo***,
Université Paris 1, Centre d'Economie de la Sorbonne, IDEP and CEPR
106, boulevard de l'Hôpital, 75647 Paris cedex 13, France

Ana Prieto****
CNRS, THEMA, Université de Cergy-Pontoise
33, boulevard du Port, 95011 Cergy-Pontoise cedex, France

15 March 2006

Abstract
We propose to model individual educational investments as a rational decision, maximizing expected utility, conditional on some characteristics observed by the student, under the combined risks affecting future wages and schooling duration. Assuming that students' attitudes toward risk can be represented by a CRRA utility, we show that the risk-aversion parameter can be identified in a natural way, using the variation in school-leaving ages, conditional on certified educational levels. Estimation can be performed by means of classic Maximum Likelihood methods. The model can easily be compared with a non-structural, simplified version, which is a standard wage equation with endogenous dummy variables representing education levels, education levels being themselves determined by an Ordered Probit model. We find small but significant values of the coefficient of relative risk aversion, between 0.1 and 0.9. These results are obtained with a rich sample of 12,500 young men who left the educational system in 1992, in France.

* We thank Christian Belzil, Denis Fougère, Marc Gurgand, Guy Laroque, Thierry Magnac, Jean-Marc Robin, Orazio Attanasio, Jose-Victor Rios-Rull, and the participants of the 2005 NBER Summer Workshop, for their help and useful discussions. The present paper is a revised version of a manuscript circulated in January 2005 and entitled: "Risk Aversion, Expected Earnings and Opportunity Costs: A Structural Econometric Model of Human Capital Investment".

** thomas.brodaty@eco.u-cergy.fr

*** garybobo@univ-paris1.fr

**** ana.prieto@eco.u-cergy.fr

1. Introduction

Youths investing in education are facing several kinds of risk. In the long run, there are risks affecting the price of skilled labour of various types on the labour market. In every period of the working life, accidents such as unemployment can reduce earnings. In the short run, there are also non-negligible risks affecting the direct and opportunity costs of education. In some countries at least, there is substantial variation in school-leaving age, among individuals with the same qualifications. The number of years needed to pass some exams, or more generally, to reach a certain certified level of achievement is substantially dispersed: students differ in their learning "speed". It follows that the costs of education, i.e., the human-capital investment-costs, are random. These costs, as well as future earnings, are random from the point of view of each individual, because of risks that cannot be controlled for, and because individuals have an imperfect knowledge of their future "job market", and "academic" abilities.

We propose to model individual educational investments as a rational decision, maximizing expected utility, conditional on some characteristics observed by the student, under the combined risks affecting future wages and schooling duration. Assuming then that the observed individuals' attitudes towards risk can be represented by a CRRA utility defined as a function of earnings, we show that the risk-aversion parameter can be identified in a natural way, using the variation in school-leaving ages, conditional on certified educational levels. Our model has unambiguous comparative-statics properties: an increase in risk-aversion always discourages students to push education further. Using variants of the basic model, we typically always find a small value of the Arrow-Pratt coefficient of relative risk aversion, between 0.1 and 0.9. We thus find that students are less risk-averse than decision-makers endowed with a logarithmic utility function. These results are obtained with a very rich sample, containing 12,500 young men who left the educational system in 1992, in France. As a by-product, we find estimates of the returns to education and of the "ability bias", that is, formally, of the correlation between the error terms of the log-wage and educational choice equations. The instruments used for educational achievement are the number of siblings, and birth-order indicators. We find a positive and highly significant value of the ability bias: in other words, OLS estimates of the returns to education in a log-earnings equation would be biased upwards.

It is important to emphasize that our approach doesn't rely on years-of-schooling as a measure for human capital, as in a number of classic studies. We instead carefully distinguish education *levels* and schooling *durations*. There is a hierarchy of diplomas and certificates that we aggregate in a number of levels (i.e., high-school drop-outs, vocational high-school certificates, high-school diploma, two years of College, four years of College, and graduate studies). A "theoretical duration," or normal school-leaving age, corresponds to each level. The observed school-leaving age is called "duration". We exploit the ratio of theoretical to observed school-leaving age, given the level reached. This ratio is an individual measure of "speed", that captures aspects of the student's "ability". We call it "Pascal-p", because of its connection with the Pascal distribution. We assume that this speed measure is known to the student, as it is to the econometrician, and we use it as a control in the regression of log-wages on dummies indicating education levels. Wages are thus explained by the educational level, observed ability — as measured by student “speed” — and other controls. It happens that both the levels and the speed with which these levels have been reached do matter. Our speed measure is also used (by the student) to predict the duration (and thus costs) of prospective studies. Again, we show that student speed, as measured by Pascal-p, plays an important role, along with family background, in educational investment decisions. We show that identification of the risk-aversion parameter is based on Pascal-p variability in a very natural way.

Of secondary interest, but probably worth mentioning, is the fact that we explore a kind of “median way”, between the most sophisticated structural approaches of returns to education, based on dynamic optimization and nested maximum-likelihood algorithms (e.g., Keane and Wolpin (1997), Eckstein and Wolpin (1999), Belzil and Hansen (2002)), and the “reduced-form” approaches, based on the search for new instruments for education, and natural experiments[1]. Our model has a relatively simple and transparent structure, because it is based on static, expected (and discounted) utility maximization. In other words, we assume that the individual chooses an educational level at the age of 16, and faces the risks associated with this decision — and in particular, the risks due to duration randomness. Dynamic models typically permit a description of sequential choices. Yet, notice that our model has a dynamic element, insofar as it relies on the idea that students do not know their market ability at the time of their education decision, and try to forecast future wages, conditional on known family-background characteristics and known aspects of ability. At the cost of this simplification, we get a model with several important advantages. First, the model can be estimated by straightforward maximum likelihood techniques. Second, the model can be compared, easily and closely, with an endogenous-dummy-variable model à la Heckman[2], in which education levels are determined by a latent variable, through an Ordered Discrete Choice structure. In a sense, our "reduced-form" model can itself be viewed as a generalized version of Cameron and Heckman’s (1998) Ordered Probit model. We compare our structural model with the “less structural”, endogenous-dummy model, by means of a test of non-nested hypotheses. The two approaches yield different results. There are clear cases in which our structural model dominates. A third advantage of the approach is that sources of identification are clear.

Related Literature on Risk and Return to Education

A huge research effort has been devoted to returns to education in many countries, following the pioneering contributions of Becker (1964) and Mincer (1974), but comparatively very few contributions have focused on the riskiness of educational investments. Becker (1964, p. 91-92), mentions the fact that wages should embody a risk and liquidity premium but did not focus on the risk-side of human capital investment. On these questions, the real forerunner is Weiss (1972), with a study of the American scientists’ wages. Weiss (1972) models wages as log-normal variables and utility functions as CRRA, and proposes an attempt to study the risk-premium element in wages. A recent echo of his approach is to be found in the work of Hartog and Vijverberg (2002). These authors estimate risk-aversion and prudence in a model of educational choices, and find values of relative risk-aversion

[1] See the surveys of Card (1999) and Harmon *et al.* (2003).

[2] See Heckman (1978), Lee (1983) and Maddala (1983).

around 0.5. They thus find the same results as us, but with a different methodology and different data. Belzil and Hansen (2004) estimate a sophisticated dynamic-programming model of education, in which a risk-aversion parameter can be inferred from individual schooling decisions. In their model, individuals are heterogenous with respect to ability, but share the same degree of relative-risk aversion. Belzil and Hansen (2004) find an Arrow-Pratt relative risk-aversion coefficient around 0.9 — a small value again. Skyt-Nielsen and Vissing-Andersen (2005) infer risk-aversion by means of numerical methods, using the stochastic-process properties of labour-income time series. Their computations lead to a much higher value of relative risk aversion, of the order of 5.

Other recent approaches of the risk in human capital investment include Belzil and Leonardi (2005), in which measures of individual risk aversion, obtained as answers to survey questions, are used as explanatory variables in a model of schooling attainment. A portfolio approach, inspired by financial asset pricing and the CAPM, has led to several recent contributions by Palacios-Huerta (2003), Hogan and Walker (2003) and Christiansen *et al.* (2006).

In the following, Section 2 presents the model. Section 3 provides a description of the data used for estimation. Section 4 presents the estimation methods, and discusses identification. Section 5 is devoted to estimation results. Concluding remarks are gathered in Section 6.

2. The Model

We start with the basic assumption that individual's educational choices can be described as the result of expected, and discounted, utility maximization. The economic agents' Von Neumann-Morgenstern utility functions exhibit constant relative risk aversion (CRRA), that is, formally,

$$u(w) = \frac{1 - w^{-\alpha}}{\alpha}, \tag{1}$$

where w is the agent's earnings. The Arrow-Pratt index of relative risk aversion for this utility function is $\gamma = 1 + \alpha$, so that we assume $\alpha \geq -1$.

Education levels are discrete variables denoted s, i.e., $s = 0, 1, ..., n$. Yearly wages (or earnings) w_s are determined as a function of education s as follows,

$$\ln(w_s) = f_s + X_0\beta_0 + \nu, \tag{2}$$

where, f_s is a skill premium (depending on s), X_0 is a vector of exogenous explanatory variables, β_0 is an associated vector of parameters, and ν is a random variable, representing unobserved aspects of "market ability" plus some noise ("luck"). To simplify the analysis, we assume that agents expect wages to be constant during the life-cycle, i.e., do not depend on age, but this assumption does not have important consequences.

A student-worker is also characterized by a pair of independent random characteristics (p, ϵ), where p is her (his) "talent" and ϵ represents unobserved effects of family background. We view "talent" p as a measure of ability. More precisely, p is a measure of the speed with which the individual reaches a given educational level s. As will be shown below, this has empirical relevance. An important assumption is that p is observed by the economic agent and the econometrician, while ϵ is observed by the agent (student) only. We thus assume that p is an entry of vector X_0. A student chooses her (his) education level s so as to maximize the expected value of a discounted sum of utilities, knowing "talent" p, observed background factors X, and unobserved factor ϵ. The sum is over periods $t = 1, ..., T$, where T is the individual's "life" duration, in years.

We represent education costs as a forgone fraction of the individual potential earnings, given the level already reached. The opportunity and direct costs of spending one more year in school, when the education level is already s, are expected to be a fraction $1 - h_{s+1}(X_1, p, \epsilon)$ of the wage w_s, that is, $(1 - h_{s+1}(X_1, p, \epsilon))w_s$, where $0 \geq h_s(.) \geq 1$, and $h_s(.)$ is a function of talent p, ϵ, and of the vector of exogenous variables X_1, for all $s = 1, ..., n$. It follows that education costs increase with s if the wage w_s itself increases with s, *ceteris paribus*. Let finally d_s denote the number of years required to reach level s. An agent thus expects to spend $\Delta d_s = d_s - d_{s-1}$ years in school to increase his (her) education level from $s-1$ to s. Durations d_s can be viewed as random variables. We assume for simplicity that durations d_s and wages w_s are independent, *conditional* on (X, p, ϵ), and we later add the assumption that the conditional expected durations $\hat{d}_s = E(d_s \mid X, p, \epsilon)$ are deterministic functions of observable characteristics (X, p). A more complex version of the model would of course view the $\hat{d}_s$ as random, and specify the joint distribution of talent p, unobservable family factors ϵ, and d_s. We can now define the expected discounted sum of utilities of choosing level s, knowing p and ϵ, as follows,

$$V(s \mid p, \epsilon) = E\left[\sum_{t=1+d_s}^{T} \delta^t u(w_s) + \sum_{z=1}^{s} \sum_{t=1+d_{z-1}}^{t=d_z} \delta^t u(h_z w_{z-1}) \,\middle|\, p, \epsilon \right], \tag{3}$$

where δ is a psychological discount factor.

We also assume that the joint distribution of ν and ϵ is normal, with correlation coefficient ρ, and independent of p. The variance of ϵ is normalized to one. We can admit the possibility that ν also captures unobserved aspects of ability or talent, provided that these aspects be independent of those captured by the other talent variable p. Formally,

$$(\nu, \epsilon) \sim \mathcal{N}\left(0, \begin{pmatrix} \sigma^2 & \sigma\rho \\ \sigma\rho & 1 \end{pmatrix}\right). \tag{4}$$

Using the identity,

$$\sum_{t=1+d_{z-1}}^{t=d_z} \delta^t = \delta^{(1+d_{z-1})} \left(\frac{1 - \delta^{\Delta d_z}}{1 - \delta} \right),$$

and the conditional independence property of wages and durations, it is possible to simplify the expression of V above. We get,

$$\begin{aligned} V(s \mid p, \epsilon) = {} & E\left[\delta^{1+d_s} \left(\frac{1 - \delta^{T-ds}}{1 - \delta} \right) \,\middle|\, p, \epsilon \right] E\left[u(w_s) \mid p, \epsilon \right] \\ & + \sum_{z=1}^{s} E\left[\delta^{1+d_{z-1}} \left(\frac{1 - \delta^{\Delta d_z}}{1 - \delta} \right) \,\middle|\, p, \epsilon \right] E\left[u(h_z w_{z-1}) \mid p, \epsilon \right]. \end{aligned}$$

2.1. Computation of ΔV

Define $\Delta V(s+1 \mid p, \epsilon) = V(s+1 \mid p, \epsilon) - V(s \mid p, \epsilon)$. To simplify notation, if (x, y) are random variables, define then the mapping

$$\phi(x, y) = E\left[\delta^{1+x}\left(\frac{1-\delta^y}{1-\delta}\right) \,\middle|\, p, \epsilon\right]. \tag{5}$$

Simple computations then yield,

$$\begin{aligned}\Delta V(s+1 \mid p, \epsilon) = {} & \phi(d_{s+1}, T - d_{s+1}) E\left[u(w_{s+1}) \mid p, \epsilon\right] \\ & -\phi(d_s, T - d_s) E\left[u(w_s) \mid p, \epsilon\right] + \phi(d_s, \Delta d_{s+1}) E\left[u(h_{s+1} w_s) \mid p, \epsilon\right]\end{aligned} \tag{6}$$

Notice that ϕ has the following useful property,

$$\phi(d_{s+1}, T - d_{s+1}) - \phi(d_s, T - d_s) + \phi(d_s, \Delta d_{s+1}) = 0. \tag{7}$$

We need to compute terms of the form $E\left[u(\kappa w_s) \mid \epsilon\right]$, with $\kappa = 1$ or $\kappa = h_{s+1}$. Denote now $W_s = f_s + X_0 \beta_0$, to simplify notation. Using (1),(2) and (4), i.e., the CRRA and normality assumptions, combined with linearity of log-wages, and the assumed independence of p and (ϵ, ν), we obtain

$$\begin{aligned}E\left[u(\kappa w_s) \mid p, \epsilon\right] = {} & E\left[(1/\alpha)\left(1 - \kappa^{-\alpha} \exp(-\alpha \ln(w_s))\right) \mid p, \epsilon\right] \\ = {} & \frac{1}{\alpha}\left(1 - \kappa^{-\alpha} E\left[\exp(-\alpha(W_s + \nu)) \mid p, \epsilon\right]\right) \\ = {} & \frac{1}{\alpha}\left(1 - \kappa^{-\alpha} e^{-\alpha W_s} E\left[e^{-\alpha \nu} \mid \epsilon\right]\right).\end{aligned}$$

Given (4), we have $E[\nu \mid \epsilon] = \rho \sigma \epsilon$ and $Var[\nu \mid \epsilon] = \sigma^2(1 - \rho^2)$. A well-known property of the expectation of a log-normal random variable then yields,

$$E[e^{-\alpha \nu} \mid \epsilon] = \exp\Big\{-\alpha\big[\rho \sigma \epsilon - \frac{\alpha}{2} \sigma^2 (1 - \rho^2)\big]\Big\} \equiv \exp[-\alpha C(\epsilon)], \tag{8}$$

and from this latter expression we derive,

$$E\left[u(\kappa w_s) \mid \epsilon\right] = \frac{1}{\alpha}\left(1 - \kappa^{-\alpha} \exp[-\alpha(W_s + C(\epsilon)]\right), \tag{9}$$

for $\kappa = 1$ or $\kappa = h_{s+1}$, where $C(\epsilon)$ is defined by (8).

Using equations (5) to (9) yields an expression of $\Delta V(s \mid \epsilon)$. Easy algebra shows, after some simplifications, that $\Delta V(s+1 \mid p, \epsilon) \leq 0$ is equivalent to,

$$\frac{1}{\alpha}\left[\frac{\phi(d_s, T - d_s) - \phi(d_{s+1}, T - d_{s+1}) e^{-\alpha \Delta f_{s+1}}}{\phi(d_{s+1}, \Delta d_{s+1})}\right] \leq \frac{(h_{s+1})^{-\alpha}}{\alpha}, \tag{10}$$

where, by definition,

$$\Delta f_{s+1} = f_{s+1} - f_s.$$

Expression (10) is crucial, since the choice of level s is optimal for an agent with unobserved family factors ϵ only if $\Delta V(s+1 \mid p\,,\epsilon) \leq 0$ and $\Delta V(s \mid p\,,\epsilon) \geq 0$.

We specify and estimate the function h_s directly as follows.

$$h_s \equiv \exp(-X_1\beta_1 - c_s)\exp(\epsilon). \tag{11}$$

The c_s can be interpreted as cost parameters, for the following reason: it is not difficult to check that, using the definitions given above,

$$\Delta \ln(h_{s+1}w_s) = -\Delta c_{s+1} + \Delta f_s,$$

for $s \geq 1$. We also assume that there is no constant in X_1, and β_1 is a vector of parameters.

Our model is therefore determined by (2), giving wages as a function of schooling and labor market ability, and (10), which characterizes the optimal schooling choice as a function of expected skill premia Δf_s, observable characteristics X_1, and unobservable family factors ϵ. But to complete the model, we need to specify the distribution of durations d_s — this is done below, in particular cases.

We can now derive several variants of the model. The parameters to be estimated are α (the risk aversion), β_0, β_1, δ (the discount factor), the f_s (the skill premia), the c_s (cost parameters), σ (the standard deviation of log-wages), ρ (the correlation of ν and ϵ), and the $\hat{d}_s$ (the conditionally expected number of years required to reach level s), which can be viewed as functions of exogenous variables (X, p). Parameters β_0, β_1, f_s, σ and ρ will not pose any particular identification problem. In contrast, δ and α might be difficult to estimate or identify.

One possibility is to let T go to infinity and study the infinite horizon version of the model with a fixed, given value of δ. A second possibility would of course be to estimate δ in the infinite horizon model with a fixed value of the risk-aversion parameter α. Finally, a third possibility is to fix a finite value of T, say, $T = 65$. It is then possible to set $\delta = 1$, and we will show that the maximum likelihood procedure can estimate (and thus identify) α in a relatively natural way. In this variant, α is estimated, on top of the β, f_s, c_s, σ and ρ.

2.2. The Finite Horizon, $\delta = 1$ Case

Let us now consider the finite-horizon, fully patient version of the model. We set $\delta = 1$ and $T < \infty$. By l'Hôpital's rule, we get for any $x, y > 0$,

$$\lim_{\delta \to 1} \phi(x,y) = E\left[\lim_{\delta \to 1} \delta^{1+x}\left(\frac{1-\delta^y}{1-\delta}\right) \;\middle|\; p\,,\epsilon\right] = E\left[y \mid p\,,\epsilon\right]. \tag{12}$$

Then, inequality (10), combined with (11) and (12), yields the following:

$$\frac{1}{\alpha}\left(\frac{(T-\hat{d}_s) - (T-\hat{d}_{s+1})e^{-\alpha\Delta f_{s+1}}}{\Delta\hat{d}_{s+1}}\right) \leq \frac{e^{\alpha(X_1\beta_1 + c_{s+1} - \epsilon)}}{\alpha}, \tag{13}$$

where $\hat{d}_s = E[d_s \mid p\,,\epsilon]$. This is, in a sense, remarkably simple, because if we assume that $\hat{d}_s$ is not a function of ϵ, the unobserved family factor ϵ intervenes in this expression only indirectly, through its effect on costs h_s. For any positive value of α, and for negative values of α which are sufficiently close to zero, taking logarithms yields,

$$\epsilon \leq X_1\beta_1 + k_{s+1} + c_{s+1}, \tag{14a}$$

an expression which is equivalent to (13), where,

$$k_s = -(1/\alpha)\ \ln\left(\frac{(T-\hat{d}_{s-1}) - (T-\hat{d}_s)e^{-\alpha\Delta f_s}}{\Delta\hat{d}_s}\right) \tag{14b}$$

It follows that the necessary condition for an optimal choice s, that is, $\Delta V(s+1 \mid p\,,\epsilon) \leq 0 \leq \Delta V(s \mid p\,,\epsilon)$ is now equivalent to

$$X_1\beta_1 + k_s + c_s \leq\ \epsilon\ \leq X_1\beta_1 + k_{s+1} + c_{s+1}. \tag{15}$$

Level s is chosen by an individual if her (his) unobserved family factor ϵ falls in the interval $[X_1\beta_1 + k_s + c_s,\ X_1\beta_1 + k_{s+1} + c_{s+1}]$. Therefore, our theory boils down to an Ordered Discrete Choice model, with a particular functional form imposed on the cuts $k_s + c_s$.

2.3. The Pascal Model of Expected Durations

Our model will now be complete if we specify the $\hat{d}_s$, i.e., the expected number of years needed to reach education level s. A difficulty is that, for each individual i, we only observe individual i's observed school-leaving age, i.e., d_{is}, for $s = s_i$. But we need the expected number of years that would have been needed by i had he or she decided to stay in the educational system. In other words we need to specify the expected durations $\hat{d}_s$ of every individual for every education level s. We adopt a simple formulation derived from the Pascal distribution. Let τ_s be the theoretical number of years needed to reach level s. For instance, if s is the high-school diploma (i.e., *baccalauréat*) then $\tau_s = 18$, and so on. Assume that each year, an individual i is promoted to the next grade with a constant probability p_i, and $(1-p_i)$ is the probability of repeating a grade — or the probability of being "held back". These parameters do vary substantially in our sample. If p_i is constant across years for each individual, the probability of reaching level s in k years, with $k \geq \tau_s$, is given by the Pascal distribution (number of independent trials needed to obtain τ_s successes, when the probability of success is constant equal to p), that is,

$$\Pr(d_s = k) = \binom{k-1}{\tau_s - 1} p^{\tau_s}(1-p)^{k-\tau_s}.$$

Using next the formula for the mean of a Pascal distribution, we get the expected duration of studies of level s for an individual who left school at level z, as follows,

$$\hat{d}_s = E[d_s \mid X] = \frac{\tau_s}{p}.$$

This suggests that for each individual, an empirical measure of "talent" p_i —which is in fact a measure of "speed"— is given by the ratio

$$p_i = \frac{\tau_{s_i}}{d_{is_i}} \tag{16}$$

where τ_{s_i} is the "theoretical" school-leaving age of i, given that individual i's observed educational level is s_i, and d_{is_i} is i's observed school-leaving age (at level s_i). So we simply take (16) as defining the observed talent variable p_i, and will refer to it as i's Pascal-p. Given that this is an observation, we can control for its possible direct effect on wages and costs, including p_i in the list of controls X — and thus, these effects are neither in ν nor in ϵ. We assume in addition that each individual i forms duration expectations with the help of this model. Hence, for all s, we set

$$\hat{d}_{si} = \frac{\tau_s}{p_i}. \tag{17}$$

We will show below that there is enough variability in the $\hat{d}_{si}$ to identify the risk-aversion parameter $1+\alpha$ by means of variations in the k_s functions, in the Ordered Probit structure defined by (14) above.

2.4. Comparative Statics

It is now possible to show that the cuts k_s are monotonically increasing functions of risk aversion (i.e., of α). To see this, define

$$A = \frac{T - (T-d)e^{-\alpha f}}{d}. \tag{18a}$$

It is sufficient to show that $K(\alpha) \equiv (-1/\alpha)\ln(A)$ is an increasing function of α for $0 < d < T$ and $f > 0$. Simple calculus yields,

$$\frac{\partial K}{\partial \alpha} = \left(\frac{1}{\alpha^2}\right)\left[\ln(A) - A^{-1}\left(\frac{T-d}{d}\right)\alpha f e^{-\alpha f}\right]. \tag{18b}$$

Using the concavity of the logarithm function, we get $\ln(1) - \ln(A) \leq A^{-1}(1-A)$, which is equivalent to $\ln(A) \geq 1 - A^{-1}$. The term between square brackets at the right-hand side of (18b) can then be bounded below as follows,

$$\ln(A) - A^{-1}\left(\frac{T-d}{d}\right)\alpha f e^{-\alpha f} \geq 1 - A^{-1}\left[1 + \left(\frac{T-d}{d}\right)\alpha f e^{-\alpha f}\right] \equiv B, \tag{19}$$

and, using (18a), the bound B itself boils down to

$$B = \frac{(T-d)(1 - e^{-\alpha f} - \alpha f e^{-\alpha f})}{T - (T-d)e^{-\alpha f}}. \tag{20}$$

Given that we assume $A > 0$, it follows from (20) that $B \geq 0$ if and only if $1 - e^{-\alpha f} \geq \alpha f e^{-\alpha f}$, or equivalently, if and only if $1 - e^{-x} \geq x e^{-x}$, but the latter inequality is true, because $-e^{-x}$ is concave. We conclude that $\partial K / \partial \alpha \geq 0$: when risk aversion increases, it takes a better unobservable family background (i.e., an increased ϵ) to study more. The number of students enrolled in a given level $s > 0$ should therefore tend to fall when risk aversion goes up.

There are some other clear-cut consequences of our theory. The number of years separating two educational levels, denoted d in (18a), has a positive impact on the thresholds K_s, that is,

$$\frac{\partial K}{\partial d} = \frac{T}{\alpha A d^2}(1 - e^{-\alpha f}) \geq 0.$$

Thresholds are also decreasing with respect to the "duration of working life" T, i.e., we easily get,

$$\frac{\partial K}{\partial T} = \frac{-1}{\alpha A d}(1 - e^{-\alpha f}) < 0.$$

This means that an increase in the duration of studies will always discourage some of the individuals with the weakest unobserved family backgrounds, who used to chose level s, and as a consequence, the average value of the enrolled students' unobserved family factors will increase. An increase in the working-life duration raises the value of education, and therefore will attract more individuals to higher education levels.

Our approach suffers from the same limitations as all studies of, say consumption or asset prices, based on an additively separable, intertemporal expected utility maximization problem: risk aversion and intertemporal substitution parameters cannot be disentangled. A more sophisticated model, using a recursive utility structure, could help separating the two. On this question, see Epstein (1992). But our model has the advantage of exhibiting unambiguous comparative statics properties. The theoretical literature on risk aversion and human capital seems limited; see however the classic contributions of Lehvari and Weiss (1974), Williams (1979) and Eaton and Rosen (1980). As far as we know, this classic literature doesn't yield unambiguous predictions of the effects of increased risk-aversion on human capital.

In addition, we must ensure that $c_s + k_s < k_{s+1} + c_{s+1}$ for all $s > 0$ to ensure that the model's probability distributions are well-defined. This latter monotonicity property will be satisfied if $\Delta f_{s+1} \leq \Delta f_s$ (i.e., "concave" returns), $\Delta d_{s+1} > \Delta d_s$ (i.e., "convex" opportunity costs), and $c_{s+1} \geq c_s$ but, these latter conditions are not necessary. Monotonicity of the cuts k_s still holds if the returns to education are increasing, (i.e., if $\Delta f_{s+1} > \Delta f_s$), provided that they don't increase too much.

2.5. The Logarithmic Utility Case (i.e., $\alpha = 0$)

An interesting particular case is obtained by letting $\alpha \to 0$. Using l'Hôpital's rule again, we get with (13)-(14),

$$\lim_{\alpha \to 0^+} \frac{1}{\alpha} \ln \left[\frac{(T - \hat{d}_{s-1}) - (T - \hat{d}_s) e^{-\alpha \Delta f_s}}{\Delta \hat{d}_s} \right] = \frac{\Delta f_s (T - \hat{d}_s)}{\Delta \hat{d}_s}.$$

This yields the logarithmic utility model, which is characterized by the inequalities,

$$c_s + X_1\beta_1 - \frac{(T - \hat{d}_s)\Delta f_s}{\Delta \hat{d}_s} \leq \epsilon \leq X_1\beta_1 - \frac{(T - \hat{d}_{s+1})\Delta f_{s+1}}{\Delta \hat{d}_{s+1}} + c_{s+1}.$$

The analytic expression of the cutoff points k_s is easily interpreted. Since agents are very patient, i.e., $\delta = 1$, the marginal skill-premium gain (per year) of jumping from level s to level $s+1$ is $\Delta f_{s+1}/\Delta \hat{d}_{s+1}$, multiplied by the number of years to go after the end of studies, i.e., $T - \hat{d}_{s+1}$. This expression of marginal benefit must be compared with an expression of marginal costs, which is simply $c_s + X_1\beta_1 - \epsilon$ here.

In the logarithmic utility case, the monotonicity condition $k_{s+1} > k_s$ is equivalent to

$$\frac{\Delta \hat{d}_{s+1}}{\Delta \hat{d}_s} > \frac{\Delta f_{s+1}}{\Delta f_s} \frac{(T - \hat{d}_{s+1})}{(T - \hat{d}_s)},$$

showing that the cutoff monotonicity condition $c_{s+1} + k_{s+1} > c_s + k_s$ will be satisfied under "return concavity", and "cost convexity". But one can allow for increasing returns, i.e., for $\Delta f_{s+1}/\Delta f_s > 1$, if T is not too large, and if $\Delta \hat{d}_{s+1}$ is large enough, for then the ratio $\frac{(T-\hat{d}_{s+1})}{(T-\hat{d}_s)}$ is sufficiently smaller than 1. These remarks continue to hold true, by continuity, for values of α which are close to 0, even if $\alpha < 0$.

2.6. The Infinite Horizon, Discounted Utility Case

To understand the potential of our model, consider the case in which $T \to +\infty$ and $\delta < 1$. To simplify the analysis, assume that the d_s are deterministic functions of (X, p). Then, inequality (10), combined with (11), yields,

$$\frac{1}{\alpha}\left(\frac{1 - \delta^{\Delta d_{s+1}} e^{-\alpha \Delta f_{s+1}}}{1 - \delta^{\Delta d_{s+1}}}\right) \leq \frac{e^{\alpha(X_1\beta_1 + c_{s+1} - \epsilon)}}{\alpha} \tag{21}$$

For any positive value of α, and for negative values of α which are sufficiently close to zero, taking logarithms yields an expression which is equivalent to (21),

$$\epsilon \leq X_1\beta_1 + l_{s+1} + c_{s+1} \tag{22a}$$

where,

$$l_s = -\frac{1}{\alpha} \ln\left(\frac{1 - \delta^{\Delta d_s} e^{-\alpha \Delta f_s}}{1 - \delta^{\Delta d_s}}\right) \tag{22b}$$

It follows that the necessary condition for an optimal choice s, that is, $\Delta V(s+1 \mid \epsilon) \leq 0 \leq \Delta V(s \mid \epsilon)$ is equivalent to

$$X_1\beta_1 + l_s + c_s \leq \epsilon \leq X_1\beta_1 + l_{s+1} + c_{s+1}. \tag{23}$$

We again obtain an Ordered Discrete Choice structure, with a particular functional form imposed on the cuts $l_s + c_s$.

3. The Data

To realize the estimations presented below, we used "Génération 92", a large scale survey conducted in France. The survey and associated data base have been produced by the CEREQ (*Centre d'Etudes et de Recherches sur les Qualifications*), a public research agency, working under the aegis of the Ministry of Education[3]. Génération 92 is a sample of 26,359 young workers of both sexes, whose education levels range from the lowest (i.e., high school dropouts) to graduate studies, and who graduated in a large array of sectors and disciplines. Observed individuals have left the educational system between January 1rst and December 31rst, 1992[4]. They have left the educational system for the first time, and for at least one year, in 1992[5]. The labor market experience of these individuals has been observed during 5 years, until 1997. The survey provides detailed observations of individual employment and unemployment spells, of wages and occupation types, as well as geographical locations of jobs. The personal labor market history of each survey respondent has been literally reconstructed, month after month, during the period 1993-1997, by means of an interview. Before 1992, the individual's educational achievement is also observed. In the present paper, we concentrate on the male sub-sample below[6].

For the purpose of estimation, we have merged education levels into six categories, 1) the high-school dropouts; 2) the vocational high-school degree holders[7]; 3) those who reached[8] grade 12; 4) those who completed two years of College[9]; 5) those who completed four years of college; 6) graduate studies (including professional schools, Masters', etc...). For descriptive statistics, see Table 1.

On top of this, the survey provides information on family background: the father's and the mother's occupation, the father's and the mother's education levels are the most important of these variables. We also know if the parents are unemployed, inactive, retired or deceased, if they work in the public or private sectors. Are also observed, notably: the age at which the individual left the educational system, the number of sisters, the number of brothers, and the rank among siblings (i.e., birth order).

A difficulty with wages is that we do not observe the hours worked (but we know if the individual worked full-time or part-time). We therefore decided to select the individuals

---

[3] Articles and descriptive statistics, concerning various aspects of the survey, are available at www.cereq.fr.

[4] To fix ideas, the number of inhabitants of France who left school for the first time in 1992 is estimated to be of the order of 640,000.

[5] They did not return to school for more than one year after 1992, and they had not left school before 1992 except for compulsory military service, illness, or pregnancy.

[6] For results on women, see Gary-Bobo et al. (2006), Brodaty et al. (2005).

[7] i.e., the so-called *Certificats d'Aptitude* and *Brevet d'Etudes Professionnelles*.

[8] We take grade 12 students in the US to correspond (roughly) to the French *classe terminale*. There is a high proportion of high-school diploma holders among this group. The national high-school degree is called *Baccalauréat*)

[9] The corresponding exam is called DEUG (*Diplôme d'Etudes Universitaires Générales*), which is the equivalent of an Associate's degree, or DUT (*Diplôme Universitaire de Technologie*), which is the equivalent of a technical or vocational Associate's. There are exams at the end of each of the college years in French universities, and the DEUG or DUT corresponds (roughly) to the level reached at the end of grade 14.

who experienced at least a full-time employment spell during the five years observation period. More precisely, we first removed 717 individuals who had never worked (no employment spell recorded during 5 years). The remaining 25642 individuals are the addition of 14213 men and 11429 women who worked at least once during the observation period. We then selected the males who experienced at least one full-time employment spell during the five years. As a consequence, we lost 11.7% of the male sub-sample, but we still have 12,538 men. The possible bias introduced by this selection procedure is thus limited in the case of men. On the other hand, this procedure permits us to compare earnings more precisely, given that full-time employment means 39 hours a week for most wage-earning employees (given the heavily regulated French labor market), and more importantly, it tends to select a relatively homogenous population of young people who really want to work full-time (which has some advantages when it comes to estimate structural parameters).

Each individual's curriculum on the job market is an array of data including a number of jobs, with their corresponding wages and durations in months, and unemployment spells, again with a length in months. To estimate the returns to education, we rely on a single, scalar index of earnings for each worker. To estimate the model, we constructed two different wage variables with the help of the data. The first statistic is simply the arithmetic average of the full-time wages earned during full-time employment spells, weighted by their respective spell durations. In the following, this index is called the "mean wage". The mean wage variable ignores the length of unemployment spells, and the difficulties faced by the individual to find a stable (and well-paid) job. To take the probability of unemployment into account, as well as to capture the effect of job instability on average earnings, we employed a second index, simply called the individual's "earnings". To compute this average, wages and unemployment benefits are weighted by the corresponding employment or unemployment spell duration[10]. Table 2 gives descriptive statistics relative to the two wage indices. See Figure 1 for a plot of the density of mean wages and earnings.

The last important piece of information that we have is school-leaving age and education level reached, for each individual. as emphasized in the introduction we distinguish educational achievement, measured by levels, from school-leaving age. Clearly, we depart from the classic studies in which education is measured by means of a number of years of schooling. Table 3 shows the distribution of school-leaving age, conditional on education level (highest level reached). There is a substantial variability in the data, which is in part due to peculiarities of the French educational system: repeated grades are frequent in school and high school. Freshmen repeating the first and second years of college are also quite common. The summation of these random sources creates the observed variability.

4. Estimation Method and Identification of Risk Aversion

We now describe our estimation procedures for the model with $T < +\infty$ and $\delta = 1$. There is a structural and a "reduced-form" version of the model. With the "reduced-form" version of the model, we get estimates of β_0, β_1, σ, ρ, the f_s parameters, and of the cuts $m_s \equiv k_s + c_s$, but neither α nor the c_s will be identified. To estimate the structural

[10] A worker is eligible for unemployment benefits if he or she has worked in the recent past. Students thus get zero before their first job. The unemployment benefits are roughly a half of the lost job's wage.

version of the model, we construct expected durations $\hat{d}_s$ as described above. Technically, all durations $\hat{d}_s$ are deterministic functions of the exogenous variables p_i, d_{is_i} and τ_{s_i}. It happens that this is enough to identify α and the c_s, together with all the other parameters β_0, β_1, σ, ρ, and f_s. Individuals are indexed by $i = 1, ..., N$. For each i, we observe the education level s_i, the wage

$$y_i = \ln(w_i),$$

and the exogenous variables $X_i = (X_{0i}, X_{1i})$, which include p_i as a control. Let ν_i and ϵ_i be i's (jointly normal) unobserved "market ability" and family background variables. For each individual i, $\Pr(s = s_i \text{ and } y = y_i) = \Pr(y = y_i) \Pr(s = s_i \mid y = y_i)$. Using (2), we get,

$$\Pr(s = s_i \mid y = y_i) = \Pr(s_i \mid \nu_i = y_i - X_{0i}\beta_0 - f_s). \tag{24}$$

Using (15) above, we can then express this conditional probability with respect to ϵ_i, as follows,

$$\Pr(s_i \mid y_i) = \Pr(c_{s_i} + k_{s_i} + X_{1i}\beta_1 \leq \epsilon_i \leq c_{s_i+1} + k_{s_i+1} + X_{1i}\beta_1 \mid \nu_i = y_i - X_{0i}\beta_0 - f_{s_i}). \tag{25}$$

Given (4) above, the theoretical regression of ϵ on ν can be written,

$$\epsilon_i = \frac{\rho}{\sigma}\nu_i + \xi_i$$

where ξ_i is normally distributed, $E(\xi_i) = 0$, $E(\nu_i \xi_i) = 0$. We also get the classic results, $E(\epsilon_i \mid \nu_i) = \rho\sigma^{-1}\nu_i$, and $Var(\epsilon_i \mid \nu_i) = Var(\xi_i) = 1 - \rho^2$. With the help of these relations, the conditional probability of choosing s can again be reformulated as follows,

$$\Pr(s_i \mid y_i) =$$
$$\Pr\left[c_{s_i} + k_{s_i} + X_{1i}\beta_1 - \frac{\rho\nu_i}{\sigma} \leq \xi_i \leq c_{s_i+1} + k_{s_i+1} + X_{1i}\beta_1 - \frac{\rho\nu_i}{\sigma} \;\middle|\; \nu_i = y_i - X_{0i}\beta_0 - f_{s_i}\right] \tag{26}$$

Let $\phi(z) = (2\pi)^{-1/2}\exp(-z^2/2)$ denote the normal density and let $\Phi(z) = \int_{-\infty}^{z} \phi(\zeta)d\zeta$ denote the normal cdf. Since ξ_i is normally distributed, we obtain,

$$\Pr(s = s_i \mid y = y_i) = \tilde{\Phi}_{s+1}(y_i, s_i, X_i) - \tilde{\Phi}_s(y_i, s_i, X_i), \tag{27a}$$

where

$$\tilde{\Phi}_s(y_i, s_i, X_i) \equiv \Phi\left[\frac{c_s + k_s + X_{1i}\beta_1 - (\rho/\sigma)(y_i - X_{0i}\beta_0 - f_{s_i})}{\sqrt{1 - \rho^2}}\right]. \tag{27b}$$

Define the indicator variables χ as follows: $\chi_{is} = 1$ if $s_i = s$ and $\chi_{is} = 0$ otherwise. Since y_i is normal with variance σ^2, its density is simply,

$$\frac{1}{\sigma}\phi\left(\frac{y_i - X_{0i}\beta_0 - \sum_s f_s \chi_{is}}{\sigma}\right). \tag{28}$$

Now, the contribution of individual i to likelihood can be expressed as,

$$L_i = \frac{1}{\sigma}\phi\left(\frac{y_i - X_{0i}\beta_0 - \sum_s f_s\chi_{is}}{\sigma}\right)\prod_{z=0}^{n}\left(\tilde{\Phi}_{z+1}(y_i, s_i, X_i) - \tilde{\Phi}_z(y_i, s_i, X_i)\right)^{\chi_{iz}}, \qquad (29)$$

where, conventionally, we set $k_0 = -\infty$ and $k_{n+1} = +\infty$, education levels being numbered from 0 to n. The likelihood function is therefore $L = \prod_{i=1}^{N} L_i$.

Several variants of the model can now be estimated by maximizing (29) with different specifications of the k_s. We have estimated the model when the cuts k_s assume the form given by (14b) above (but estimation could of course also be performed, in principle, with the cuts defined as l_s (i.e., expression (22b)), and with a fixed value of $\delta < 1$). This also allows for a grid-search maximum-likelihood estimation of δ. The cuts k_s are nonlinear functions of the risk aversion parameter α, to be estimated. In the following, this structural model will be called Model A.

From this, we can easily see that our structural model is a generalization of a more standard, latent-variable model with endogenous qualitative variables à la Heckman (1978). To see this, assume that the choice of education level s is determined by a latent variable

$$z_i = -X_{1i}\beta_1 + \epsilon_i, \qquad (30)$$

where ϵ is still the unobservable family factor (or just unobserved "school ability"). Assume in addition that log-wages y are still determined by the regression function $y_i = X_{0i}\beta_0 + \sum_s f_s\chi_{is} + \nu_i$, and finally that the dummy variables χ satisfy the property,

$$\Pr(\chi_{is} = 1) = \Pr(s_i = s) = \Pr(m_s \leq z_i \leq m_{s+1}). \qquad (31)$$

The χs are then determined by an ordered discrete choice structure with endogenous thresholds m_s. Expression (29) above is precisely the contribution to likelihood of this model, if we forget the fact that the cuts can be decomposed as $m_s = k_s + c_s$ and that the k_s satisfy (14b). We can therefore define a model in which the m_s are estimated freely, along with ρ, σ, the f_s, and the β parameters. We then ignore the structural preference parameter α. In the following, this model à la Heckman will be called Model B. Model B is fully justified insofar as it is derived from our theoretical human-capital investment model.

With these definitions, we are able to compare the results of the structural and "reduced-form" estimation methods, i.e., Models A and B, as closely as possible.

Which sources of exogenous variability identify the skill-premia parameters Δf_s (i.e., the returns to levels of education)? As in many studies of returns to education, the endogeneity of education-level dummies is taken care of by some exclusions, i.e; by zero restrictions in the Mincerian log-wage equation. The excluded variables — and thus the instruments for "years of schooling" — here, are, in essence, the birth-order and number-of-siblings variables. We have shown elsewhere that these variables do matter to explain educational achievement with the same data (see Gary-Bobo *et al.* (2006) and Brodaty *et al.* (2005)). This has been shown by others too (see e.g., Black *et al.* (2005). We also think that these variables are not — or very weakly — correlated with the wage-equation error-term ν, and are thus valid instruments: it seems that the fact of having many brothers and sisters (when many other family-background variables are used as controls) has no strong and compelling reason to influence wages directly. All the model parameters except of course risk aversion $1 + \alpha$ can be identified without any problem by a "reduced-form" model à la Heckman, i.e., Model B. Model A therefore shows that risk aversion can be identified by using the variability of school-leaving age d_s, conditional on educational levels s. Intuitively, the risk affecting the direct and opportunity costs of education is due to grade repetitions in an important way, and this kind of risk is itself a source of variability that can be exploited to identify a risk-aversion parameter.

To sum up, our structural approach is relatively simple and transparent; it has a closely comparable "reduced-form" counterpart, and the sources of identification of structural parameters are clear.

5. Estimation Results

We start our discussion of empirical results with some comments on the Pascal-p variable. Figure 2 displays the empirical density of Pascal-p values in the sample, conditional on the student's educational achievement (i.e. level s). There is a remarkable pattern, showing that the expected value of p, knowing s, is increasing. In contrast, the variance of p conditional on s seems to be roughly constant, so that Figure 2 gives the impression that the distributions of p knowing s are obtained by successive translations of the same density function. The Pascal-p is well-dispersed, and as will be shown below, conveys much information. Notice that the empirical Pascal-p takes values greater than one: this is because a non-negligible fraction of students pass their exams sooner than the standard, or "theoretical" age τ_s. The observed dispersion in turn comes from the substantial variance in school-leaving ages, conditional on education level s, as shown in Table 3. Next, Figure 3 shows the results of nonparametric regressions of "theoretical" age τ_s on Pascal-p, in sub-samples determined by different levels of parental educational achievement. Mother's and father's education are observed separately and fall in 6 categories[11]. With the help of these observations, we constructed cells with the mother's and the father's characteristics, such as "(father: went to college (or more)) times (mother: high school graduate)" . We merged the cells of some of these cross-categories in which the numbers of observations were relatively small. This led to the parental education variables listed in Fig. 3. The strong positive correlation of p and s, or τ_s, is again visible. It is also striking to see that there is strict dominance-ranking of the regressions plotted on Figure 3, the higher parental education, the higher the chosen τ_s, for all values of p. The interaction of Pascal-p and parental education seems to have a strong, unambiguous effect on completed education.

The estimation results for Model A and Model B are displayed by Tables 4 and 5. Each table presents the results obtained for Models A and B with one of the two different earnings statistics defined above: mean (full-time) wages, and mean earnings. Recall that

[11] 1) No qualification; 2) Elementary school certificate; 3) Vocational high-school degree; 4) High-school degree (*baccalauréat*); 5) Went to college; 6) Missing. There exists in France an elementary school certificate (*Certificat d'Etudes Primaires*) which is becoming rare but used to be very important in the past.

our structural model is Model A. Model B is the "reduced-form" version "à la Heckman", obtained when the cuts of the ordered schooling choice model are estimated freely, and when no structural connection between cuts, durations and wages is imposed. Model B cannot provide an estimation of risk-aversion, but yields an estimation of the error terms' correlation ρ.

Table 4 gives the results obtained with mean wages. Table 4(i) lists the estimates for the parameters of the Ordered Probit structure; that is, the estimated values of β_1, and the "cuts". These cuts are the c_s in the case of Model A and the m_s in the case of Model B. Risk-aversion $\gamma = 1 + \alpha$ is estimated by Model A only through the k_s functions, while the Pascal-p is introduced as a control in X_1 in Model B only. Model A and Model B cuts are thus difficult to compare since in Model B, $m_s + \beta_p p$ plays the role of Model A's $c_s + k_s(\alpha, \Delta f_s, \hat{d}_s(p))$. The rest of the β_1 parameters are very similar.

Mother's and father's occupations are observed separately, and fall in one of 8 categories[12]. With the help of these observations, we constructed interaction dummies with the mother's and the father's characteristics, such as "(father: middle manager) times (mother: white collar)". We merged (i.e. added the dummies of) some of these cross-categories in which the numbers of observations were too small. This led to the parental occupation variables listed as controls in Tables 4-8[13].

The variables used as instruments are birth-order dummies, the number of brothers, and the number of sisters. Of these, only birth-order dummies are fairly significant in the first-stage regression. This result is close to those of Black *et al.* (2005), Gary-Bobo *et al.* (2006). But it is not the last word about the impact of family size, as will be seen below: the non-significant number-of-siblings parameters are hiding from us a contrasted situation, in which these parameters are positive in some groups, and negative in some others. In any case, the use of birth order as an instrument for education levels is a success.

Notice that risk aversion is small, with a significant value of $1 + \alpha = .41$: this means that individuals are on average less risk-averse than decision-makers endowed with logarithmic utility functions —but they are not risk neutral. The very strong and highly significant coefficient on Pascal-p in Model B shows the empirical relevance of this variable as a control for education levels.

We now turn to Table 4(ii) which summarizes the log-wage equation estimates. This second sub-table also has a number of striking features. First of all, the β_0 parameters, appearing in the middle (parental occupation, parental employment) are roughly similar, not always very significant. A notable difference is the positive direct impact of parents in the executive (and highly educated professionals) category on wages, which is much stronger and precise in Model A than in Model B. But then, in spite of exactly similar specifications, the Model A and Model B wage-equations give very different pictures of the returns to education. According to Model B, returns to education are high — in fact,

[12] 1) Farmers; 2) Craftsmen, Tradesmen and Owners-Managers; 3) Executives, Doctors, Lawyers, Engineers and Teachers; 4) Middle managers, technicians; 5) White Collars; 6) Blue Collars; 7) Missing observations; 8) Deceased or unemployed.

[13] "Father and Mother: white collars" is the reference group. F stands for father and M stands for mother in the names given to controls.

they're a bit higher than the equivalent OLS estimates — and as a consequence, there is a negative "ability bias", as shown by the correlation of ν and ϵ, which is $\rho = -.075$. The Pascal-p variable has a negative, significant, and direct effect on wages: it seems that returns to education are smaller for those who finish school quickly (relative to the chosen education level). In sharp contrast, the Model-A estimate of the same error-term correlation is high, very significant, and positive: we find $\rho = +.36$, indicating a strong "ability bias" effect. According to Model A, OLS estimates of the returns to education are biased upwards, and the Pascal-p variable has a positive effect on wages: our ability measure has a significant coefficient with the expected sign.

Table 4(ii) lists the Δf_s coefficients directly. It is readily seen that the typical estimate of these skill-premia parameters is 6.5% for Model A and 12 to 14% for Model B, except for the highest level. The value of graduate studies[14] seems to be grossly overstated by Model B. This is a likely result of Model A's nonlinear specification, which is very sensitive around the estimated risk-aversion coefficient. But in an essential way, Model A captures something that is overlooked by Model B, namely, the fact that there are interactions of the skill-premia Δf_s with expected durations $\hat{d}_s$ in the k_s cut-functions. These non-trivial and non-negligible interactions represent the effect of the student's expectations, and are likely to be responsible for part of the differences in estimated returns to education. Now, recalling that the Δf_s coefficients are the value of a jump from level $s-1$ to level s, and this jump takes typically $\Delta \bar{d}_s$ years, where $\bar{d}_s$ denotes average duration, we should in fact compute,

$$(1 + \Delta f_s)^{1/\Delta \bar{d}_s} - 1 \approx \frac{\Delta f_s}{\Delta \bar{d}_s}.$$

Given that $\Delta \bar{d}_s$ is approximately 2 years, we see that the per annum returns $\Delta f_s / \Delta \bar{d}_s$ are of the order of 3.25% in Model A and 6 to 7% in Model B (except for graduate studies, which last one additional year only, in many cases).

The last lines of Table 3 show the Vuong test of non-nested hypotheses, which is based on the likelihood ratio (see Vuong (1989)). In the present case, given the number of variables and observations, the test reads as an ordinary t-statistic: a value greater than 1.96 indicates that Model A is better than Model B, a smaller value says only that Model A doesn't statistically dominates Model B. This is precisely the case with the estimations of Table 3 — but we will find a strict domination of Model A when earnings are used instead of wages.

Table 5 displays the estimation results in the case of the average earnings statistic; it is organized exactly as Table 4. We will briefly comment on the most important similarities and differences with Table 4. On Table 5(i), risk aversion is even smaller, with a value of $1 + \alpha = .17$ in Model A, and Pascal-p is still very strong as a control, in Model B. Birth order also works well as an instrument for education levels. On Table 5(ii), we find the same opposite signs for the Pascal-p control coefficient in Models A and B and the same opposite sign of the ability bias: correlation ρ is still negative and significant in Model B and positive around 34%, and very precisely estimated in Model A. As a consequence, we

[14] Many graduate students spend in fact only one more year at the University, to get a DESS or DEA (i.e., the equivalent of a Master's degree).

find that the Model-A returns to education are small as compared to that of Model B. This time, Vuong's test rejects Model B very clearly in favor of Model A, with a t-value around 8.38. The reason for such a sharp contrast lies in the difference of the average log-likelihoods, $logL_A - logL_B$ say, which is of the order of .003 in Table 4, but approximately 7 times greater, around .02, in Table 5.

We conclude that there are many indications, both quantitative and qualitative or intuitive, which seem to indicate that Model A is a better description of the world than Model B. Table 6 below summarizes the key findings[15].

TABLE 6
Comparison of Models A and B

	Wages		Earnings	
	Model A	Model B	Model A	Model B
Δf_1	0.065**	0.149**	0.048**	0.467**
Δf_2	0.066**	0.126**	0.049**	0.205**
Δf_3	0.068**	0.184**	0.051**	0.309**
Δf_4	0.066**	0.123**	0.051**	0.150**
Δf_5	0.034**	0.270**	0.026**	0.354**
Pascal-p	0.529**	-0.497*	0.659**	-1.595**
ρ	0.362**	-0.075	0.340**	-0.171*
$1+\alpha$	0.411**	—	0.176*	—
Likelihood	-1.410	-1.423	-2.041	-2.060
Vuong test	1.046	—	8.386	—

As a robustness check for some of the features of our theory, we re-estimated Model A in sub-samples determined by parental education. We contrast the results obtained with the sub-sample of individuals whose parents have no qualification on the one hand, and the sub-sample of students whose parents are at least high-school graduates, on the other hand. The results are reported in Table 7 for mean wages, and in Table 8 for mean earnings. Table 7(i) shows that risk aversion varies with parental education: we get a highly significant $1 + \alpha = .91$ when parents are not educated, but only $1 + \alpha = .20$ when parents are at least high-school graduates. Note in passing that the effect of a higher number of siblings on educational achievement is positive and significant in the educated parents' families; this confirms results obtained elsewhere (see Gary-Bobo et al. (2006)). Table 7(ii) shows that the effects of Pascal-p on wages are always significant, but with

[15] Notice that in Table 6, the Pascal-p line gives the coefficient of Pascal-p in the log-wage equation. Likelihood is in fact mean log-likelihood. The Vuong statistic tests Model A against Model B here: a positive value higher than 1.96 says that Model A is preferred; a negative value smaller than -1.96 says that Model B is preferred; none of the models dominates the other if the test value is in between.

different signs: the coefficient's sign is negative in the uneducated parents' sub-sample, whereas it is positive in the other sub-sample. Finishing quickly has a positive effect on wages for the sons of educated parents, but negative effect for the others. This could simply be the mechanical result of the fact that low educational-background students who stay longer at school command a higher wage than the early dropouts, even if they take more time to reach a given level, whereas in educated families, being a laggard is a bad signal. The same sign-pattern holds for the correlation of error terms ρ: the ability bias is positive and significant only in the case of the educated parents' sons, and accordingly, returns to education are low for the latter, but high for the former.

Table 8 displays the results obtained with the mean earnings statistic: the risk-aversion coefficient has very low and non-significant values. But zero is a possible value of risk-aversion, corresponding to risk-neutrality. The fact that earnings are more dispersed than wages is an explanation for the fact that a small value of α is enough to fit the individual's educational choices, whereas a higher value is needed in the case of the wage statistic, because it is less risky. Table 8(ii) does not exhibit the strange opposite signs of the ability bias that we have found in Table 7: this bias is always positive and highly significant, albeit smaller in size in the case of the uneducated parents' sons.

7. References

Becker, G. S. (1964), *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, Third edition, 1993, NBER and the University of Chicago Press, Chicago, Illinois.

Belzil, C. and J. Hansen, (2002), "Unobserved Ability and the Returns to Schooling", *Econometrica*, 70, 2075-2092.

Belzil, C. and J. Hansen, (2004), "Earnings Dispersion, Risk Aversion and Education", *Research in Labor Economics*, 23, 335-358.

Belzil, C. and M. Leonardi, (2005), "Can Risk Aversion Explain Schooling Attainments? Evidence from Italy", *IZA, Institute for the Study of Labour*, Bonn, Germany.

Black, S.E., Devereux, P.J., and K.G. Salvanes (2005), "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education", *Quarterly Journal of Economics*, 120, 669-700.

Brodaty, T., Gary-Bobo, R. J. and A. Prieto (2005), "Risk Aversion, Expected Earnings and Opportunity Costs: A Structural Econometric Model of Human Capital Investment", *University Paris 1, Centre d'Economie de la Sorbonne, CNRS*, manuscript.

Cameron, S. V., and J.J. Heckman (1998), "Life-Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males", *Journal of Political Economy*, 106, 262-333.

Card, D. (1999), "The Causal Effect of Education on Earnings", Chapter 30 in: Ashenfelter, O. and D. Card eds., *Handbook of Labor Economics, Volume 3*, Elsevier Science, Amsterdam.

Christiansen, C., Schröter-Joensen, J. and H. Skyt-Nielsen (2006), "The Risk-Return Trade-Off in Human Capital Investment", IZA Discussion Paper N°1962, *IZA, Institute for the Study of Labour*, Bonn, Germany.

Eaton, J. and H. S. Rosen (1980), "Taxation, Human Capital, and Uncertainty", *American Economic Review*, 70, 705-715

Eckstein, Z. and K. Wolpin (1999), "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities", *Econometrica*, 67, 1295-1340.

Epstein, L. G. (1992), "Behavior under Risk: Recent Developments in Theory and Applications", Chapter 1 in J-J. Laffont ed., *Advances in Economic Theory, 6th World Congress, vol. 2*, Cambridge University Press, Cambridge.

Gary-Bobo, R. J., Prieto, A. and N. Picard (2006), "Birth Order and Sibship Sex Composition as Instruments in the Study of Education and Earnings", CEPR DP n°5514, *Center for Economic Policy Research*, London, UK.

Harmon, C., Oosterbek, H. and I. Walker (2003), "The Returns to Education: Microeconomics", *Journal of Economic Surveys*, 17, 115-155.

Hartog, J. and W. P. M. Vijverberg (2002), "Do Wages Really Compensate for Risk Aversion and Skewness Affection", IZA DP N°426, *IZA Institute for the Study of Labor*, Bonn, Germany.

Heckman, J. J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System", *Econometrica*, 46, 931-959.

Hogan, V. and I. Walker (2003), "Education Choices under Uncertainty: Implications for Public Policy", manuscript, *University College, Dublin*, Dublin, Ireland.

Keane, M. and K. Wolpin (1997), "Career Decisions of Young Men", *Journal of Political Economy*, 105, 473-522.

Lee, L.F. (1983), "Generalized Econometric Models with Selectivity", *Econometrica*, 51, 507-512.

Levhari, D. and Y. Weiss (1974), "The Effect of Risk on the Investment in Human Capital", *American Economic Review*, 64, 950-963.

Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Econometric Society Monographs, Cambridge University Press, Cambridge, UK.

Mincer J. (1974), *Schooling, Experience, and Earnings*, Columbia University Press, New York.

Palacios-Huerta, I. (2003), "An Empirical Analysis of the Risk Properties of Human Capital Returns", *American Economic Review*, 93, 948-964.

Skyt-Nielsen H., and A. Vissing-Jorgensen (2005), "The Impact of Labor Income Risk on Educational Choices: Estimates and Implied Risk-Aversion", manuscript, University of Aarhus, Denmark.

Vuong, Q. H. (1989), "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econometrica*, 57, 307-334.

Weiss, Y. (1972), "The Risk Element in Occupational and Educational Choices", *Journal of Political Economy*, 80, 1203-1213.

Williams, J. (1979), "Uncertainty and the Accumulation of Human Capital over the Life-Cycle", *Journal of Business*, 52, 521-548.

Table 1 : Completed schooling

Level	s	numbers	%
High school dropouts	0	710	6%
Vocational degree	1	4799	38%
High school graduates (grade 12)	2	2770	22%
Two years of college (grade 14)	3	2139	17%
Four years of college (grade 16)	4	698	6%
Graduate studies	5	1422	11%

Table 2 : Summary statistics; wages for each level of education; euros

	mean	standard deviation	min	max
1. Mean earnings	**873**	**461**	**14**	**6943**
High school dropouts	566	335	14	4955
Vocational degree	707	305	20	6943
High school graduates (grade 12)	798	348	16	3773
Two years of college (grade 14)	967	371	51	3714
Four years of college (grade 16)	1103	498	64	3932
Graduate studies	1475	626	35	5495
2. Mean wages	**1153**	**450**	**83**	**9520**
High school dropouts	878	284	83	4955
Vocational degree	950	223	101	4915
High school graduates (grade 12)	1069	317	214	3773
Two years of college (grade 14)	1251	344	366	3714
Four years of college (grade 16)	1420	430	305	4171
Graduate studies	1860	612	183	9520

Table 3 : Empirical distribution of school-leaving age, conditional on education level reached

Age while leaving school	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
High school dropouts	1	**17**	**24**	**33**	**16**	7	1	1	0	0	0	0	0	0	0	0	0	0
Vocational degree	0	0	3	**30**	**37**	**21**	**6**	2	1	0	0	0	0	0	0	0	0	0
High school graduates (grade 12)	0	0	0	2	**12**	**31**	**32**	**15**	6	2	1	0	0	0	0	0	0	0
Two years of college (grade 14)	0	0	0	0	0	**13**	**27**	**28**	**19**	8	2	1	1	0	0	0	0	0
Four years of college (grade 16)	0	0	0	0	0	0	4	**17**	**20**	**27**	**13**	9	4	3	1	1	1	1
Graduate studies	0	0	0	0	0	0	0	4	**24**	**29**	**15**	**11**	7	4	2	2	1	0

Table 4 : Mean Wages

Table 4(i) : Ordered schooling choice

	Model A		Model B	
School leaving choice (s)	**Coefficient**	**Student-t**	**Coefficient**	**Student-t**
Parental occupation				
(F=Farmer) x (M = Farmer)	0.106	1.959	0.140	2.584
(F=Farmer) x (M = Other)	0.037	0.359	0.073	0.701
(F=Craftsman) x (M = Farmer, Craftsman, Blue C,)	0.063	1.077	0.059	1.009
(F=Craftsman) x (M = Exec,, Middle, White C,)	0.120	2.557	0.123	2.606
(F=Executive) x (M = Farmer, Craftsman, Middle)	0.746	17.435	0.741	17.358
(F=Executive) x (M = Exec,)	1.083	22.236	1.057	21.796
(F=Middle Manager) x (M = Farmer, Craftsman, Middle, Blue C,)	0.283	4.482	0.281	4.464
(F=Middle Manager) x (M = Exec,)	0.417	5.971	0.433	6.207
(F=Middle Manager) x (M = White C,)	0.309	6.073	0.308	6.071
(F=White Collar) x (M = Farmer, Blue C,)	-0.147	-2.944	-0.147	-2.942
(F=White Collar) x (M = Craftsman, Exec,, Middle)	0.375	4.441	0.378	4.487
(F=White Collar) x (M =White C,), Reference group				
(F=Blue Collar) x (M = Farmer, Craftsman, Blue C,)	-0.310	-7.972	-0.318	-8.177
(F=Blue Collar) x (M = Exec,, Middle, White C,)	-0.189	-4.727	-0.186	-4.659
(F=Missing or deceased) x (M = Crafts,, Exec,)	-0.137	-1.334	-0.126	-1.232
(F=Missing or deceased) x (M = Missing or deceased)	-0.348	-3.587	-0.331	-3.407
Parental employment				
Both parents in private sector , Reference group				
(F=Private sector) x (M=Public s,)	0.131	3.404	0.135	3.500
(F=Public sector) x (M=Private s,)	0.007	0.151	0.005	0.089
Both parents in public sector	0.106	2.558	0.106	2.552
(F=Private sector) x (M=Unemployed)	0.135	4.629	0.149	5.096
(F=Public sector) x (M=Unemployed)	0.098	2.365	0.109	2.627
(F=Employed) x (M=Deceased or Missing)	-0.320	-3.354	-0.327	-3.424
(F=Unemployed) x (M=Deceased or Missing)	0.070	0.922	0.076	1.003
Both parents unemployed	0.174	2.430	0.201	2.797
(F=Retired) x (M=Employed,)	0.829	11.712	0.844	11.948
(F=Retired) x (M=Unemployed,)	0.899	20.530	0.953	21.715
(F=Deceased or Missing) x (M=Employed,)	0.301	2.730	0.298	2.701
(F=Deceased or Missing) x (M=Unemployed,)	0.511	4.925	0.527	5.066
(F=Unemployed, retired or deceased) x (M=Deceased or Miss)	0.296	2.662	0.294	2.633
Number of brothers	0.000	0.037	-0.008	-0.716
Number of sisters	-0.001	-0.043	-0.014	-1.151
Birth order				
Only child	0.492	7.082	0.567	7.794
1rst	0.409	6.101	0.456	6.458
2nd	0.346	5.378	0.410	6.071
3rd	0.244	3.760	0.264	3.843
4th	0.169	2.362	0.207	2.731
5th and higher, Reference group				
Pascal-*p*			9.492	54.743
Estimated Risk Aversion $1+\alpha$	0.411	8.751		
Cut1	1.727	10.736	7.415	41.100
Cut2	3.247	21.067	9.197	49.851
Cut3	3.979	25.055	9.940	53.266
Cut4	4.125	30.507	10.650	56.499
Cut5	4.346	32.928	10.970	57.950

Table 4(ii) : Log-mean-wages equation

Log-mean-wages	Model A Coefficient	Model A Student-t	Model B Coefficient	Model B Student-t
Constant	8.234	174.078	9.019	96.360
σ	0.282	117.500	0.260	136.947
Completed Education				
High school dropouts, Reference group				
Vocational degree	0.065	14.705	0.149	5.795
High school graduates (grade 12)	0.066	14.886	0.126	7.308
Two years of college (grade 14)	0.068	14.870	0.184	14.240
Four years of college (grade 16)	0.066	16.146	0.123	8.877
Graduate studies (grade 17 or more)	0.034	16.800	0.270	15.970
Parental Occupation				
(F=Farmer) x (M = Farmer)	-0.083	-5.873	-0.086	-6.477
(F=Farmer) x (M = Other)	-0.026	-0.942	-0.035	-1.395
(F=Craftsman) x (M = Farmer, Craftsman, Blue C.)	0.015	0.993	0.006	0.408
(F=Craftsman) x (M = Exec., Middle, White C.)	0.031	2.520	0.008	0.704
(F=Executive) x (M = Farmer, Craftsman, Middle)	0.129	11.017	0.030	2.000
(F=Executive) x (M = Exec.)	0.192	14.459	0.048	2.565
(F=Middle Manager) x (M = Farmer, Craftsman, Middle, Blue C.)	0.057	3.419	0.027	1.673
(F=Middle Manager) x (M = Exec.)	0.086	4.659	0.030	1.674
(F=Middle Manager) x (M = White C.)	0.061	4.519	0.025	1.864
(F=White Collar) x (M = Farmer, Blue C.)	-0.017	-1.290	0.000	-0.033
(F=White Collar) x (M = Craftsman, Exec., Middle)	0.030	1.347	-0.008	-0.377
(F=White Collar) x (M =White C.), Reference group				
(F=Blue Collar) x (M = Farmer, Craftsman, Blue C.)	-0.035	-3.480	-0.003	-0.327
(F=Blue Collar) x (M = Exec., Middle, White C.)	-0.017	-1.615	0.000	-0.030
(F=Missing or deceased) x (M = Crafts., Exec.)	0.012	0.433	0.014	0.546
(F=Missing or deceased) x (M = Missing or deceased)	-0.068	-2.665	-0.033	-1.364
Parental employment				
Both parents in private sector, Reference group				
(F=Private sector) x (M=Public sector)	-0.008	-0.752	-0.022	-2.358
(F=Public sector) x (M=Private sector)	-0.009	-0.713	-0.012	-1.033
Both parents in public sector	-0.019	-1.722	-0.030	-3.010
(F=Private sector) x (M=Unemployed)	-0.006	-0.800	-0.021	-3.028
(F=Public sector) x (M=Unemployed)	-0.021	-1.963	-0.034	-3.434
(F=Employed) x (M=Deceased or Missing)	-0.035	-1.401	-0.010	-0.442
(F=Unemployed) x (M=Deceased or Missing)	-0.032	-1.596	-0.039	-2.104
Both parents unemployed	-0.057	-3.086	-0.071	-4.152
(F=Retired) x (M=Employed)	0.068	3.629	-0.026	-1.308
(F=Retired) x (M=Unemployed)	0.042	3.684	-0.055	-3.753
(F=Deceased or Missing) x (M=Employed)	-0.017	-0.578	-0.042	-1.559
(F=Deceased or Missing) x (M=Unemployed)	-0.018	-0.680	-0.068	-2.659
(F=Unemployed, retired or deceased) x (M=Deceased or Miss.)	-0.029	-0.979	-0.048	-1.756
Pascal-*p*	0.529	9.441	-0.497	-3.629
Estimated correlation ρ	0.362	21.076	-0.075	-1.251
Sample size	12,538		12,538	
Mean Log-Likelihood	-1.41995		-1.42307	
Vuong Test (Model A vs Model B)	1.0463			

Note: Vuong's model selection test for non-nested models statistics is $(LR-K)/(n^{1/2}\omega)$ where LR is the log-likelihood ratio, n is the sample size, ω is the estimated standard error of the log-likelihood ratio and K=0 because of identical number of parameters.
A significant positive (resp. negative) t-value indicates that model A is preferred to B (resp. B is preferred to A); non significant t-value (positive or negative) indicates that no model is preferred to the other. T-value for significance at the 95%-level is around 1.96.

Table 5 : Mean Earnings

Table 5(i) : Ordered schooling choice

School leaving choice (s)	Model A		Model B	
	Coefficient	Student-t	Coefficient	Student-t
Parental occupation				
(F=Farmer) x (M = Farmer)	0.106	1.947	0.141	2.610
(F=Farmer) x (M = Other)	0.030	0.287	0.071	0.681
(F=Craftsman) x (M = Farmer, Craftsman, Blue C.)	0.061	1.048	0.058	1.002
(F=Craftsman) x (M = Exec., Middle, White C.)	0.117	2.473	0.123	2.613
(F=Executive) x (M = Farmer, Craftsman, Middle)	0.733	17.124	0.742	17.372
(F=Executive) x (M = Exec.)	1.064	21.885	1.056	21.763
(F=Middle Manager) x (M = Farmer, Craftsman, Middle, Blue C.)	0.280	4.448	0.282	4.480
(F=Middle Manager) x (M = Exec.)	0.406	5.824	0.431	6.195
(F=Middle Manager) x (M = White C.)	0.305	6.006	0.309	6.099
(F=White Collar) x (M = Farmer, Blue C.)	-0.148	-2.956	-0.146	-2.924
(F=White Collar) x (M = Craftsman, Exec., Middle)	0.377	4.461	0.378	4.483
(F=White Collar) x (M =White C.), Reference group				
(F=Blue Collar) x (M = Farmer, Craftsman, Blue C.)	-0.313	-8.051	-0.317	-8.178
(F=Blue Collar) x (M = Exec., Middle, White C.)	-0.189	-4.763	-0.185	-4.656
(F=Missing or deceased) x (M = Crafts., Exec.)	-0.145	-1.411	-0.127	-1.242
(F=Missing or deceased) x (M = Missing or deceased)	-0.361	-3.706	-0.327	-3.369
Parental employment				
Both parents in private sector, Reference group				
(F=Private sector) x (M=Public sector)	0.128	3.319	0.135	3.504
(F=Public sector) x (M=Private sector)	0.002	0.037	0.005	0.111
Both parents in public sector	0.102	2.446	0.106	2.553
(F=Private sector) x (M=Unemployed)	0.132	4.507	0.150	5.158
(F=Public sector) x (M=Unemployed)	0.093	2.251	0.110	2.647
(F=Employed) x (M=Deceased or Missing)	-0.325	-3.386	-0.326	-3.410
(F=Unemployed) x (M=Deceased or Missing)	0.069	0.899	0.077	1.017
Both parents unemployed	0.166	2.313	0.203	2.827
(F=Retired) x (M=Employed)	0.821	11.632	0.845	11.983
(F=Retired) x (M=Unemployed)	0.888	20.316	0.952	21.737
(F=Deceased or Missing) x (M=Employed)	0.302	2.721	0.299	2.712
(F=Deceased or Missing) x (M=Unemployed)	0.518	4.971	0.524	5.048
(F=Unemployed, retired or deceased) x (M=Deceased or Miss.)	0.302	2.705	0.291	2.618
Number of brothers	-0.002	-0.162	-0.010	-0.881
Number of sisters	0.001	0.099	-0.018	-1.479
Birth order				
Only child				
1rst	0.495	6.947	0.552	7.635
2nd	0.414	6.025	0.438	6.220
3rd	0.342	5.198	0.405	6.055
4th	0.243	3.702	0.252	3.689
5th and higher, Reference group	0.159	2.189	0.208	2.775
Pascal-*p*			9.488	54.716
Estimated Risk Aversion $1+\alpha$	0.176	3.343		
Cut1	1.161	8.268	7.394	41.054
Cut2	2.768	21.274	9.175	49.782
Cut3	3.483	26.605	9.917	53.203
Cut4	3.770	32.419	10.627	56.439
Cut5	4.029	34.918	10.947	57.891

Table 5(ii) : Log-mean-earnings equation

	Model A		Model B	
Log-mean-wages	**Coefficient**	**Student-t**	**Coefficient**	**Student-t**
Constant	7.803	107.929	9.305	49.079
σ	0.522	153.618	0.497	87.246
Completed Education				
High school dropouts, Reference group				
Vocational degree	0.048	16.586	0.467	8.860
High school graduates (grade 12)	0.049	17.500	0.205	5.849
Two years of college (grade 14)	0.051	17.586	0.309	11.934
Four years of college (grade 16)	0.051	18.741	0.150	5.595
Graduate studies (grade 17 or more)	0.026	18.500	0.354	10.634
Parental Occupation				
(F=Farmer) x (M = Farmer)	-0.010	-0.361	-0.020	-0.793
(F=Farmer) x (M = Other)	0.061	1.203	0.042	0.861
(F=Craftsman) x (M = Farmer, Craftsman, Blue C.)	0.027	0.965	0.012	0.433
(F=Craftsman) x (M = Exec., Middle, White C.)	0.097	4.278	0.062	2.845
(F=Executive) x (M = Farmer, Craftsman, Middle)	0.181	8.796	-0.006	-0.198
(F=Executive) x (M = Exec.)	0.262	11.351	-0.003	-0.073
(F=Middle Manager) x (M = Farmer, Craftsman, Middle, Blue C.)	0.089	2.889	0.022	0.734
(F=Middle Manager) x (M = Exec.)	0.093	2.740	-0.013	-0.379
(F=Middle Manager) x (M = White C.)	0.107	4.324	0.032	1.262
(F=White Collar) x (M = Farmer, Blue C.)	0.012	0.494	0.049	2.094
(F=White Collar) x (M = Craftsman, Exec., Middle)	0.049	1.166	-0.039	-0.937
(F=White Collar) x (M =White C.), Reference group				
(F=Blue Collar) x (M = Farmer, Craftsman, Blue C.)	-0.053	-2.886	0.022	1.069
(F=Blue Collar) x (M = Exec., Middle, White C.)	-0.026	-1.396	0.012	0.647
(F=Missing or deceased) x (M = Crafts., Exec.)	0.024	0.467	0.041	0.855
(F=Missing or deceased) x (M = Missing or deceased)	-0.131	-2.762	-0.048	-1.041
Parental employment				
Both parents in private sector, Reference group				
(F=Private sector) x (M=Public sector)	-0.030	-1.604	-0.059	-3.253
(F=Public sector) x (M=Private sector)	-0.046	-1.929	-0.049	-2.149
Both parents in public sector	-0.038	-1.896	-0.061	-3.155
(F=Private sector) x (M=Unemployed)	-0.041	-2.929	-0.070	-5.140
(F=Public sector) x (M=Unemployed)	-0.078	-3.960	-0.098	-5.163
(F=Employed) x (M=Deceased or Missing)	-0.100	-2.179	-0.032	-0.727
(F=Unemployed) x (M=Deceased or Missing)	-0.096	-2.627	-0.105	-3.009
Both parents unemployed	-0.143	-4.196	-0.165	-5.031
(F=Retired) x (M=Employed)	0.071	2.100	-0.113	-2.915
(F=Retired) x (M=Unemployed)	0.012	0.605	-0.177	-6.055
(F=Deceased or Missing) x (M=Employed)	-0.072	-1.314	-0.122	-2.359
(F=Deceased or Missing) x (M=Unemployed)	-0.110	-2.163	-0.200	-4.096
(F=Unemployed, retired or deceased) x (M=Deceased or Miss.)	-0.094	-1.725	-0.138	-2.672
Pascal-*p*	0.659	8.479	-1.595	-5.697
Estimated correlation ρ	0.340	33.683	-0.171	-2.672
Sample size	12,538		12,538	
Mean Log-Likelihood	-2.04099		-2.06049	
Vuong Test (Model A vs Model B)	8.3865			

Note: Vuong's model selection test for non-nested models statistics is (LR-K)/($n^{1/2}\omega$) where LR is the log-likelihood ratio, n is the sample size, ω is the estimated standard error of the log-likelihood ratio and K=0 because of identical number of parameters.
A significant positive (resp. negative) t-value indicates that model A is preferred to B (resp. B is preferred to A); non significant t-value (positive or negative) indicates that no model is preferred to the other. T-value for significance at the 95%-level is around 1.96.

Table 7 : Mean Wages for Various Parental Education Levels

Table 7(i) : Ordered schooling choice

	Parental Education:			
	No qualification		At Least High-School Degree	
School leaving choice (s)	Coefficient	Student-t	Coefficient	Student-t
Parental occupation				
(F=Farmer) x (M = Farmer)	-0.152	-0.935	0.466	2.346
(F=Farmer) x (M = Other)	-0.762	-2.220	0.728	3.636
(F=Craftsman) x (M = Farmer, Craftsman, Blue C.)	0.170	0.735	0.497	3.353
(F=Craftsman) x (M = Exec., Middle, White C.)	-0.321	-1.951	0.427	4.196
(F=Executive) x (M = Farmer, Craftsman, Middle)	0.336	1.525	0.726	9.088
(F=Executive) x (M = Exec.)	0.343	0.821	0.982	11.937
(F=Middle Manager) x (M = Farmer, Craftsman, Middle, Blue C.)	-0.130	-0.460	0.516	4.246
(F=Middle Manager) x (M = Exec.)	0.046	0.142	0.610	5.003
(F=Middle Manager) x (M = White C.)	0.106	0.502	0.333	3.116
(F=White Collar) x (M = Farmer, Blue C.)	-0.289	-1.885	-0.046	-0.297
(F=White Collar) x (M = Craftsman, Exec., Middle)	0.712	1.431	0.318	2.321
(F=White Collar) x (M =White C.), Reference group				
(F=Blue Collar) x (M = Farmer, Craftsman, Blue C.)	-0.362	-2.978	-0.499	-3.254
(F=Blue Collar) x (M = Exec., Middle, White C.)	-0.334	-2.598	-0.169	-1.569
(F=Missing or deceased) x (M = Crafts., Exec.)	-0.091	-0.315	-0.125	-0.523
(F=Missing or deceased) x (M = Missing or deceased)	-0.348	-1.466	0.217	0.789
Parental employment				
Both parents in private sector, Reference group				
(F=Private sector) x (M=Public sector)	0.116	0.738	0.143	2.007
(F=Public sector) x (M=Private sector)	-0.242	-1.071	0.000	-0.004
Both parents in public sector	0.002	0.008	0.201	2.770
(F=Private sector) x (M=Unemployed)	0.170	1.959	0.220	3.239
(F=Public sector) x (M=Unemployed)	-0.102	-0.800	0.360	4.269
(F=Employed) x (M=Deceased or Missing)	-0.403	-1.004	0.036	0.141
(F=Unemployed) x (M=Deceased or Missing)	-0.013	-0.059	0.013	0.079
Both parents unemployed	0.283	2.032	0.417	1.333
(F=Retired) x (M=Employed)	0.542	2.392	0.898	6.544
(F=Retired) x (M=Unemployed)	0.792	7.132	0.962	9.137
(F=Deceased or Missing) x (M=Employed)	0.043	0.137	0.668	2.521
(F=Deceased or Missing) x (M=Unemployed)	0.266	0.996	1.255	4.663
(F=Unemployed, retired or deceased) x (M=Deceased or Miss.)	-0.091	-0.326	0.484	1.597
Number of brothers	-0.019	-0.744	0.064	2.556
Number of sisters	-0.029	-1.032	0.089	3.395
Birth order				
Only child	0.463	3.163	0.460	2.333
1rst	0.413	2.919	0.395	2.041
2nd	0.384	2.839	0.311	1.652
3rd	0.293	2.170	0.262	1.347
4th	0.367	2.596	0.419	1.781
5th and higher, Reference group				
Estimated Risk Aversion $1+\alpha$	0.910	24.149	0.202	2.418
Cut1	3.815	7.751	1.551	5.219
Cut2	4.792	13.217	2.761	9.688
Cut3	6.005	14.253	3.506	12.368
Cut4	5.212	13.708	3.725	13.985
Cut5	5.423	12.351	4.010	15.242

Table 7(ii) : Log-mean-wages equation

	Parental Education:			
	No qualification		At Least High-School Degree	
Log-mean-wages	**Coefficient**	**Student-t**	**Coefficient**	**Student-t**
Constant	9.219	55.606	7.824	82.014
σ	0.250	42.407	0.334	65.431
Completed Education				
High school dropouts, Reference group				
Vocational degree	0.195	6.662	0.050	10.617
High school graduates (grade 12)	0.178	8.184	0.051	10.563
Two years of college (grade 14)	0.204	6.990	0.052	10.714
Four years of college (grade 16)	0.164	7.870	0.052	11.304
Graduate studies (grade 17 or more)	0.081	7.017	0.027	11.565
Parental Occupation				
(F=Farmer) x (M = Farmer)	-0.040	-1.064	0.007	0.120
(F=Farmer) x (M = Other)	0.007	0.097	0.081	1.299
(F=Craftsman) x (M = Farmer, Craftsman, Blue C.)	0.025	0.468	0.128	2.754
(F=Craftsman) x (M = Exec., Middle, White C.)	0.083	2.150	0.129	4.063
(F=Executive) x (M = Farmer, Craftsman, Middle)	-0.023	-0.427	0.176	6.980
(F=Executive) x (M = Exec.)	0.151	1.527	0.219	8.477
(F=Middle Manager) x (M = Farmer, Craftsman, Middle, Blue C.)	-0.009	-0.138	0.119	3.117
(F=Middle Manager) x (M = Exec.)	0.239	3.204	0.129	3.397
(F=Middle Manager) x (M = White C.)	0.046	0.916	0.088	2.616
(F=White Collar) x (M = Farmer, Blue C.)	0.030	0.852	0.052	1.060
(F=White Collar) x (M = Craftsman, Exec., Middle)	-0.037	-0.317	0.064	1.482
(F=White Collar) x (M =White C.), Reference group				
(F=Blue Collar) x (M = Farmer, Craftsman, Blue C.)	0.033	1.152	-0.102	-2.144
(F=Blue Collar) x (M = Exec., Middle, White C.)	0.008	0.267	0.026	0.762
(F=Missing or deceased) x (M = Crafts., Exec.)	0.142	2.063	0.092	1.208
(F=Missing or deceased) x (M = Missing or deceased)	0.036	0.657	0.194	2.237
Parental employment				
Both parents in private sector, Reference group				
(F=Private sector) x (M=Public sector)	-0.029	-0.818	-0.012	-0.552
(F=Public sector) x (M=Private sector)	-0.067	-1.315	-0.012	-0.411
Both parents in public sector	-0.083	-1.775	-0.026	-1.173
(F=Private sector) x (M=Unemployed)	-0.081	-4.159	0.032	1.514
(F=Public sector) x (M=Unemployed)	-0.066	-2.290	0.004	0.155
(F=Employed) x (M=Deceased or Missing)	0.000	-0.004	-0.070	-0.871
(F=Unemployed) x (M=Deceased or Missing)	-0.040	-0.800	-0.059	-1.133
Both parents unemployed	-0.127	-4.068	-0.038	-0.388
(F=Retired) x (M=Employed)	-0.025	-0.465	0.133	3.236
(F=Retired) x (M=Unemployed)	-0.106	-3.809	0.083	2.730
(F=Deceased or Missing) x (M=Employed)	-0.154	-2.034	-0.022	-0.265
(F=Deceased or Missing) x (M=Unemployed)	-0.112	-1.806	0.022	0.260
(F=Unemployed, retired or deceased) x (M=Deceased or Miss.)	-0.093	-1.470	-0.127	-1.330
Pascal-*p*	-0.786	-3.547	1.005	9.960
Estimated correlation ρ	-0.188	-1.878	0.495	24.384
Sample size	1,508		2,848	
Mean Log-Likelihood	-1.29288		-1.59962	

Table 8 : Mean Earnings for Various Parental Education Levels

Table 8(i) : Ordered schooling choice

	Parental Education:			
	No qualification		At Least High-School Degree	
School leaving choice (s)	Coefficient	Student-t	Coefficient	Student-t
Parental occupation				
(F=Farmer) x (M = Farmer)	-0.160	-0.999	0.466	2.343
(F=Farmer) x (M = Other)	-0.743	-2.184	0.735	3.674
(F=Craftsman) x (M = Farmer, Craftsman, Blue C.)	0.208	0.903	0.506	3.408
(F=Craftsman) x (M = Exec., Middle, White C.)	-0.304	-1.860	0.432	4.235
(F=Executive) x (M = Farmer, Craftsman, Middle)	0.320	1.455	0.730	9.077
(F=Executive) x (M = Exec.)	0.378	0.919	0.987	11.914
(F=Middle Manager) x (M = Farmer, Craftsman, Middle, Blue C.)	-0.063	-0.220	0.520	4.268
(F=Middle Manager) x (M = Exec.)	0.001	0.002	0.613	5.014
(F=Middle Manager) x (M = White C.)	0.128	0.609	0.333	3.107
(F=White Collar) x (M = Farmer, Blue C.)	-0.281	-1.849	-0.044	-0.274
(F=White Collar) x (M = Craftsman, Exec., Middle)	0.600	1.227	0.331	2.411
(F=White Collar) x (M =White C.), Reference group				
(F=Blue Collar) x (M = Farmer, Craftsman, Blue C.)	-0.362	-3.028	-0.471	-3.068
(F=Blue Collar) x (M = Exec., Middle, White C.)	-0.328	-2.580	-0.168	-1.552
(F=Missing or deceased) x (M = Crafts., Exec.)	-0.133	-0.442	-0.136	-0.564
(F=Missing or deceased) x (M = Missing or deceased)	-0.401	-1.679	0.221	0.799
Parental employment				
Both parents in private sector, Reference group				
(F=Private sector) x (M=Public sector)	0.092	0.588	0.136	1.922
(F=Public sector) x (M=Private sector)	-0.275	-1.228	-0.008	-0.089
Both parents in public sector	-0.029	-0.144	0.197	2.731
(F=Private sector) x (M=Unemployed)	0.136	1.576	0.218	3.244
(F=Public sector) x (M=Unemployed)	-0.131	-1.033	0.360	4.297
(F=Employed) x (M=Deceased or Missing)	-0.366	-0.914	0.030	0.115
(F=Unemployed) x (M=Deceased or Missing)	0.002	0.012	0.008	0.051
Both parents unemployed	0.245	1.768	0.449	1.426
(F=Retired) x (M=Employed)	0.525	2.307	0.900	6.572
(F=Retired) x (M=Unemployed)	0.713	6.357	0.967	9.197
(F=Deceased or Missing) x (M=Employed)	0.082	0.248	0.679	2.559
(F=Deceased or Missing) x (M=Unemployed)	0.249	0.909	1.307	4.841
(F=Unemployed, retired or deceased) x (M=Deceased or Miss.)	-0.064	-0.228	0.514	1.686
Number of brothers	-0.015	-0.579	0.054	2.065
Number of sisters	0.001	0.050	0.091	3.353
Birth order				
Only child	0.386	2.673	0.531	2.582
1rst	0.420	3.066	0.465	2.307
2nd	0.301	2.256	0.375	1.906
3rd	0.294	2.245	0.354	1.749
4th	0.339	2.474	0.500	2.049
5th and higher, Reference group				
Estimated Risk Aversion $1+\alpha$	0.189	0.873	0.129	1.534
Cut1	1.059	2.181	1.430	4.846
Cut2	2.606	6.582	2.717	9.689
Cut3	3.254	8.023	3.480	12.481
Cut4	3.389	10.571	3.753	14.140
Cut5	3.667	11.333	4.060	15.408

Table 8(ii) : Log-mean-earnings equation

	Parental Education:			
	No qualification		At Least High-School Degree	
Log-mean-earnings	**Coefficient**	**Student-t**	**Coefficient**	**Student-t**
Constant	8.032	36.792	7.597	49.491
σ	0.532	52.186	0.540	72.919
Completed Education				
High school dropouts, Reference group				
Vocational degree	0.049	4.058	0.046	11.244
High school graduates (grade 12)	0.049	4.384	0.047	11.167
Two years of college (grade 14)	0.051	4.331	0.049	11.302
Four years of college (grade 16)	0.049	4.861	0.049	12.125
Graduate studies (grade 17 or more)	0.025	5.020	0.025	12.450
Parental Occupation				
(F=Farmer) x (M = Farmer)	-0.004	-0.055	0.082	0.816
(F=Farmer) x (M = Other)	0.152	0.939	0.238	2.355
(F=Craftsman) x (M = Farmer, Craftsman, Blue C.)	0.073	0.641	0.218	2.905
(F=Craftsman) x (M = Exec., Middle, White C.)	0.152	1.880	0.280	5.456
(F=Executive) x (M = Farmer, Craftsman, Middle)	0.036	0.328	0.257	6.354
(F=Executive) x (M = Exec.)	0.461	2.208	0.296	7.197
(F=Middle Manager) x (M = Farmer, Craftsman, Middle, Blue C.)	0.055	0.395	0.214	3.456
(F=Middle Manager) x (M = Exec.)	0.358	2.271	0.221	3.608
(F=Middle Manager) x (M = White C.)	0.218	2.064	0.157	2.890
(F=White Collar) x (M = Farmer, Blue C.)	0.091	1.222	0.130	1.657
(F=White Collar) x (M = Craftsman, Exec., Middle)	0.208	0.854	0.113	1.605
(F=White Collar) x (M =White C.), Reference group				
(F=Blue Collar) x (M = Farmer, Craftsman, Blue C.)	0.006	0.103	-0.124	-1.613
(F=Blue Collar) x (M = Exec., Middle, White C.)	-0.008	-0.125	0.075	1.369
(F=Missing or deceased) x (M = Crafts., Exec.)	0.291	1.975	0.090	0.748
(F=Missing or deceased) x (M = Missing or deceased)	0.022	0.186	0.176	1.264
Parental employment				
Both parents in private sector, Reference group				
(F=Private sector) x (M=Public sector)	-0.145	-1.908	-0.040	-1.124
(F=Public sector) x (M=Private sector)	-0.369	-3.429	-0.050	-1.055
Both parents in public sector	-0.126	-1.274	-0.030	-0.834
(F=Private sector) x (M=Unemployed)	-0.166	-4.019	0.018	0.522
(F=Public sector) x (M=Unemployed)	-0.213	-3.529	-0.014	-0.345
(F=Employed) x (M=Deceased or Missing)	-0.069	-0.357	-0.018	-0.137
(F=Unemployed) x (M=Deceased or Missing)	-0.161	-1.535	-0.106	-1.273
Both parents unemployed	-0.235	-3.566	-0.131	-0.826
(F=Retired) x (M=Employed)	-0.064	-0.579	0.130	1.958
(F=Retired) x (M=Unemployed)	-0.161	-3.120	0.091	1.866
(F=Deceased or Missing) x (M=Employed)	-0.274	-1.699	-0.025	-0.185
(F=Deceased or Missing) x (M=Unemployed)	-0.252	-1.893	0.079	0.593
(F=Unemployed, retired or deceased) x (M=Deceased or Miss.)	-0.161	-1.198	-0.125	-0.816
Pascal-*p*	0.434	1.855	0.842	5.279
Estimated correlation ρ	0.281	8.634	0.420	23.071
Sample size	1,508		2,848	
Mean Log-Likelihood	-2.01820		-2.11142	

Figure 1: Wage distributions (in euros)

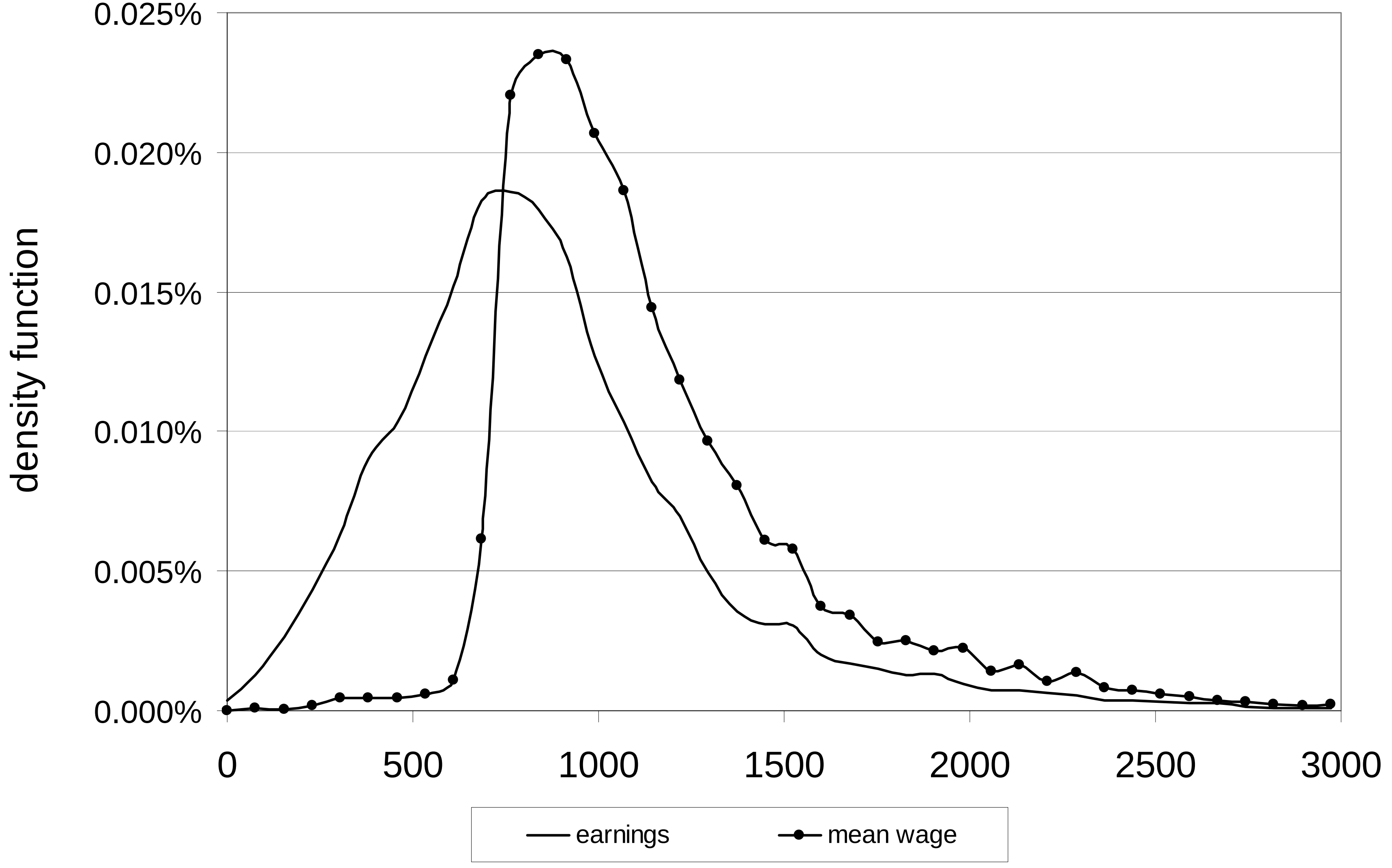

Figure 2: Empirical Distribution of Pascal P for various observed schooling levels

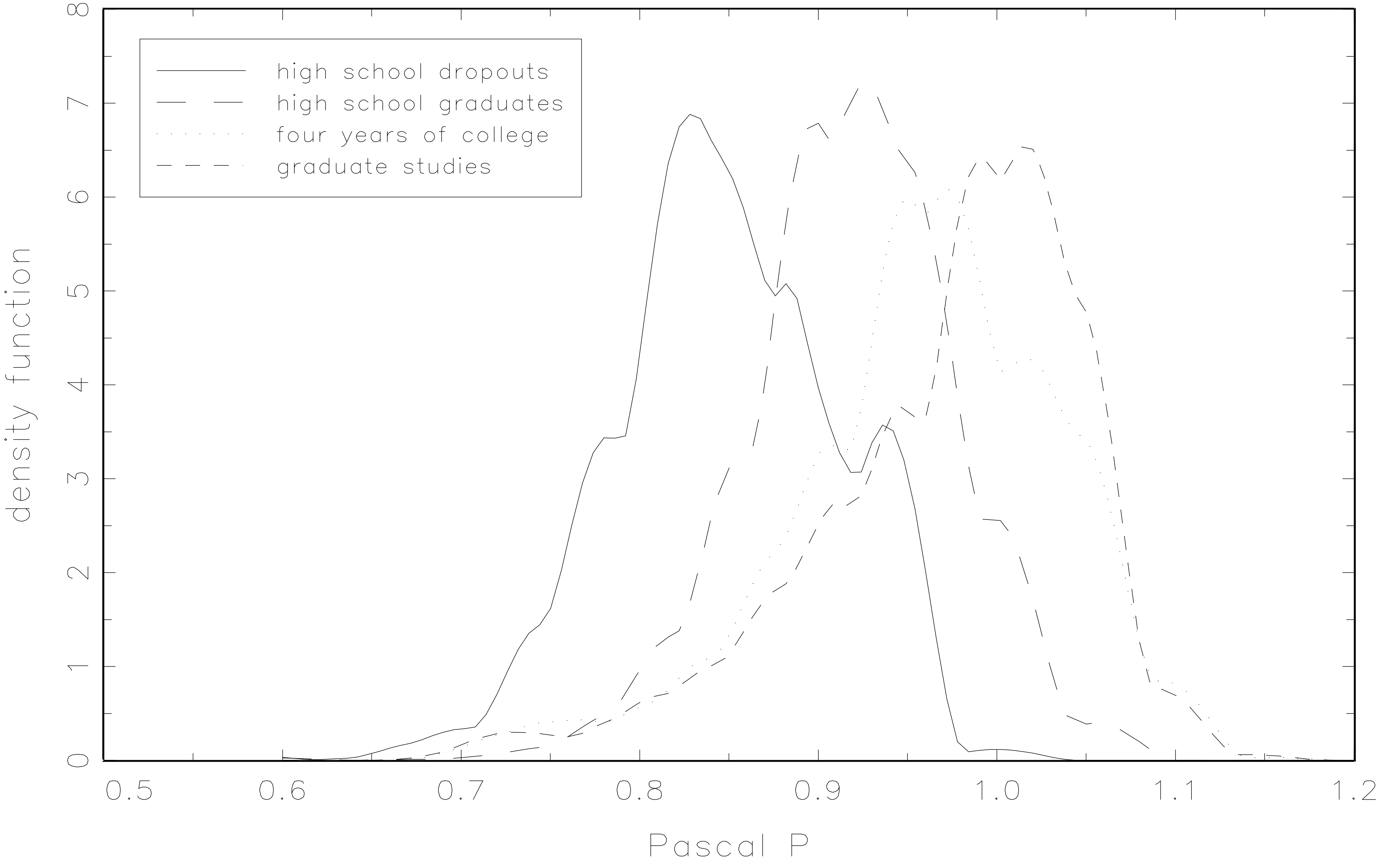

Figure 3: Non parametric regression of chosen schooling levels on Pascal P for various parental education levels