



IZA DP No. 43

Household Characteristics, Ability and Education: Evidence from a Dynamic Expected Utility Model

Christian Belzil
Joergen Hansen

May 1999

Household Characteristics, Ability and Education: Evidence from a Dynamic Expected Utility Model

**Christian Belzil
Jörgen Hansen**

**Discussion Paper No. 43
May 1999**

IZA

P.O. Box 7240
D-53072 Bonn
Germany

Tel.: +49-228-3894-0
Fax: +49-228-3894-210
Email: iza@iza.org

This Discussion Paper is issued within the framework of IZA's research area *General Labor Economics*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor markets, (2) internationalization of labor markets and European integration, (3) the welfare state and labor markets, (4) labor markets in transition, (5) the future of work, and (6) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character.

ABSTRACT

Household Characteristics, Ability and Education: Evidence from a Dynamic Expected Utility Model

We estimate a Dynamic Programming model of the decision between continuing schooling or entering the labor market using a panel from the National Longitudinal Survey (NLSY). The model, set in an expected utility framework (with a power utility function), fits data on both schooling attainments and wage very well. We find a degree of relative risk aversion much smaller than usually found in the finance literature (around 0.6) and a subjective annual discount rate between 4% and 5%. Various simulations indicate that schooling attainments are elastic with respect to the return to college education and, to a lesser extent, with respect to measured ability (Armed Forces Qualification Test scores) but inelastic with respect to household characteristics (especially household income). The ``true'' intergenerational correlation between schooling attainments and parents' education (after conditioning on observed and unobserved ability) is found to be quite low. Finally, our estimates of the return to schooling indicate strong returns to high school graduation and a clear evidence of a positive correlation between the utility of attending school and unobserved labor market ability (ability bias). Estimates of the return to schooling which take into account the ability bias are found to be between 25% and 30% smaller than those obtained ignoring it.

JEL Classification: J2,J3

Keywords: Dynamic Programming, Returns to Education, Risk Aversion, Rate of Time Preference

Jörgen Hansen
Concordia University
Department of Economics
Montreal, H3G 1M8
Canada
e-mail: hansenj@vax2.concordia.ca

1 Introduction

For a long time, economists have considered the acquisition of general human capital through education as one of the most important activity by which young individuals increase their potential lifetime earnings (see Ben-Porath, 1967, for a classic example). While enrolled in school, individuals typically receive parental support and give up potential earnings in favor of higher future earnings. Parental transfers can take the form of housing services and other living expenses (such as food and transportation) and are most likely unaffected by stochastic elements affecting household income. As opposed to parental transfers, which are therefore most likely non-stochastic from the perspective of young individuals, future earnings are usually unknown. As a consequence, modeling education choices in a dynamic stochastic environment requires a theoretical framework in which both the rate of time preference (the discount rate) and the behavior toward risk (the degree of risk aversion) play a central part. Surprisingly, economists have only recently devoted efforts to design empirical models in which the decision to acquire an additional year of schooling is modeled in a stochastic dynamic framework. Keane and Wolpin (1997) have used a structural dynamic programming model of schooling and occupational decisions using a cohort of the NLSY and payed a particular attention to the capacity of a dynamic programming model to fit data on occupational choices and mobility patterns. Cameron and Heckman (1998) have developed a semi-structural model of schooling choices estimated using hazard function specifications of the transition probability from one grade level to the following and have investigated the effects of parental background on schooling attainments. footnote

Although the estimation of structural dynamic programming models has become increasingly popular, the literature is still dominated by models in which individual preferences are represented by a random utility specification of the instantaneous (per-period) utility function. Random utility specifications are particularly useful since they are based on the assumption that the per-period utility of a given choice is the sum of a structural part (incorporating key preference parameters) and a stochastic shock (typically assumed to follow an extreme value distribution when there are more than two choices). In the context of a dynamic programming model of education choices in which preferences are defined over wages (or productivity) and in which wages are stochastic, a linear utility function would typically restrict individuals preferences to a logarithmic utility function and therefore imply a maintained hypothesis that individuals are risk averse.

The main objective of this paper is to estimate a dynamic programming model of education choices in which individual preferences are set in an expected (non-linear) utility framework more general than those typically used. The model is implemented on a panel of individuals taken from the National Longitudinal Survey of Youth (NLSY), and is based on the assumption that individual preferences are representable by a power utility function and that individuals maximize the expected discounted value of lifetime utility over a finite horizon. The finite horizon structure of our model can be used in order to solve the expected utility integral in the last period and solve the dynamic programming problem using backward recursions. There are two fundamental preference parameters in the model; the degree of relative risk aversion and the subjective discount factor. In the model, young individuals must decide on the optimal time spent at school (during which income is non-stochastic) and in the labor market (during which income follows a stochastic process). As a consequence, the theoretical framework provides an opportunity to investigate both the degree of risk aversion and the rate of time preference as separate parameters. Although discount factors and measures of risk aversion play a central role in financial economics, these parameters are practically never investigated jointly in the dynamic discrete choice literature. footnote In the finance literature, the allocation of wealth between risky and risk-free assets is used to make inference about subjective discount rates and the degree of risk aversion. At the outset, and given that our model disregards savings, our measure of risk aversion is clearly different than the one found in the finance literature and, accordingly, our estimates of the degree of relative risk aversion should not necessarily coincide with those obtained in empirical finance.

In order to design an econometric model consistent with observed schooling attainments, we allow both the return to schooling and the utility of attending school to vary with the level of schooling (using splines). Among other things, this allows us to investigate the local return to schooling at those years which correspond to high school or college graduation. The model is used to investigate a relatively well known problem encountered when estimating the true return to schooling; namely the “ability bias”. Despite the large amount of empirical work concerned with the wage return to schooling, obtaining consistent (or unbiased) estimates of the wage return to schooling is still a major challenge for applied econometricians. It is well known that, if individuals with higher unobserved ability (earning higher wages) are also those who tend to obtain more schooling, the utility of attending school is positively correlated with the error term of the wage equation and OLS estimates of the return to schooling will suffer from “ability bias”.

In order to circumvent the problems surrounding the ability bias, several economists have tried to use instrumental variable techniques (IV) or institutional changes in minimum schooling age in order to obtain more reliable estimates of the effect of an additional year of education. Harmon and Walker (1994) have used the change in minimum schooling age and Angrist and Krueger (1991) have used season of birth. As many estimates reported in the literature are meant to measure a local return to schooling (say around high school graduation), it is often difficult to generalize the results to higher and lower levels of schooling. Instrumental variable estimates of the return to schooling above 10% per year (usually around 15%) are often reported. footnote These estimates are however quite puzzling and, indeed, are difficult to reconcile with US data on long run average risk free interest rates, which are likely to be indicative of average individual discount rates. More precisely, given an average completed schooling attainment between 12 and 13 years (as of 1990) and given that the average schooling attainment at age 16 is only slightly below 11 years, estimated returns to education between 10% and 15% are hardly compatible with subjective discount rate values anywhere near the long run average risk free interest rates. Unless individuals experience a particularly high disutility of enrolling in post high school education, this suggests that standard IV estimates tend to overestimate the economic value of a marginal year of schooling (Taber, 1998). footnote

Using data on household characteristics (parents’ education, number of siblings, presence of both biological parents at age 14, geographical location and family income), which are likely to control for household ability and financial resources, and data on Armed Force Qualification Tests (AFQT) score achievements (an observable measure of individual ability), we construct several model specifications able to investigate whether the utility of attending school is correlated with the intercept term of the wage equation (after controlling for household characteristics) and provide a test for the presence of ability bias. In one version of our model, the positive correlation between the utility of attending school and labor market wages arises when AFQT score affects both the level of net earnings while at school and the intercept term of the wage equation. In a second version, we assume that individual ability has an observable component (measured by test scores) and an unobservable component. Using a discrete bi-variate distribution, we allow unobserved ability affecting the utility of attending school to be correlated with unobserved heterogeneity affecting the wage equation and we a test for the presence of ability bias remaining after controlling for test scores.

After investigating the importance of the ability bias, we pay a particular attention to the separate effects of household characteristics and individual ability on schooling attainments. The availability of data on family characteristics and on AFQT score achievements allows us to measure the “true” effects of family background variables (after controlling for individual observed and unobserved ability) and investigate whether factors such as father’s and mother’s education have a larger impact on schooling attainments than family income. In order to compare how education choices are influenced by economic variables external to the family environment (the return to college education and the variance of the stochastic innovation affecting wages) as opposed to household characteristics, we simulate the effects of an increase in the returns to education and an increase in wage dispersion on schooling attainments and compare these

elasticities with those obtained for various household characteristics.

The main results are as follows. We find a degree of risk aversion much smaller than usually found in the finance literature; our estimates of the degree of relative risk aversion are around 0.6 and are quite robust. As well, our estimates of the subjective annual discount rate are remarkably robust and lie between 4% and 5%. Various simulations indicate that schooling attainments are particularly elastic with respect to the return to college education (elasticities are around 1.3) and relatively elastic with respect to wage dispersion (elasticities with respect to the variance of log wages are around -.10) as well as with respect to measured ability (elasticities for test scores are around 0.12). However, schooling attainments are quite inelastic with respect to household characteristics. Among household characteristics, the highest elasticities are for father's education (around 0.03 at the median value of father's education) and mother's education (averaging 0.02 at the median). Schooling attainments are surprisingly inelastic with respect to household income (below 0.02 at median family income). This seems to indicate that policies aimed at raising schooling attainments, which are based on tuition subsidies, may well be ineffective. The true correlation between schooling attainments and parents' education (after conditioning on observed and unobserved ability) is therefore found to be quite low.

Finally, except for a sharp increase in the return to schooling at high school graduation, our estimates indicate that the returns to schooling are slightly declining with the level of schooling. There is a clear evidence that the utility of attending school is positively correlated with labor market ability and the correlation persists after controlling for AFQT scores. Model specifications which control for the ability bias are found to fit the data much better. Overall, and depending on the level of schooling, estimates obtained under the alternative hypothesis are between 40% and 5% lower than those obtained under the null (no ability bias). The most reliable estimates of the returns to schooling are around 6% and 7% per year for basic education (until year 11) and indicate a 2% premium to high school graduation. Despite the introduction of test scores, there is still a clear positive correlation between unmeasured ability affecting the utility of attending school and unmeasured ability affecting wages. Perhaps, the most important achievement of this paper is therefore to obtain structural estimates of the fundamental preference parameters (especially the discount rate) and estimates of the "true" return to schooling which are able to reconcile data on observed schooling attainments with financial data on long run risk free interest rates.

The content of this paper is as follows. Section 2 is devoted to the presentation of the model while Section 3 contains a brief description of the data. The main empirical results are discussed in Section 4. The conclusion is in Section 5.

2 A Dynamic Programming Model

We model education choices as a dynamic discrete choice problem over a finite horizon. The individual can either stay in school and acquire more education (while receiving parental support) or enter the labor market. The decision to remain at school (obtain an additional year of schooling) is modeled as an optimal stopping problem; that is when an individual decides to enter the labor market, it is impossible to return to school.

2.1 Theoretical Structure

In this sub-section, we present the theoretical structure of our empirical model.

- **Assumption 1- The Preferences**

Each individual is endowed with an instantaneous (per period) power utility function $U(I_t)$ given by

$$U(I_t) = \frac{I_t^{1-\alpha} - 1}{1 - \alpha} \quad \#$$

where I_t denotes income at time t . Risk aversion occurs when $\alpha > 0$ while risk loving occurs when $\alpha < 0$. When $\alpha = 0$, individuals are risk neutral. Each individual maximizes his expected discounted lifetime utility by choosing the optimal time to interrupt schooling and enter the labor

market. The discount factor, β , is given by

$$\beta = \frac{1}{1 + \rho}$$

where ρ is the subjective discount rate. The time horizon, T , is finite and is chosen to be when individuals turn 65 years old (a typical retirement age).

● Assumption 2- The Utility of Attending School

While at school, every individual receive income from parental transfers and must bear direct costs (such as books, transportation or other similar costs) as well as other psychological costs associated to the disutility of learning. The net income while at school, ξ_t , is assumed to be non-stochastic. This reflects that parental transfers are most likely unaffected by those stochastic elements affecting household income. The per period utility of attending school is therefore

$$U(\xi_t) = \frac{\xi_t^{1-\alpha} - 1}{1 - \alpha}$$

and is allowed to be nonstationary in order to capture the fact that parental transfers may vary according to the level of schooling considered. When an individual leaves school, ξ_t is set to 0.

● Assumption 3- The Return to Human Capital

Once the individual has entered the labor market, he receives income generated from the level of human capital accumulated up to this time. The return to human capital is captured in the equation for the log wage (w_t) paid in the labor market. As w_t plays the role of the state variable, the wage regression equation is the transition equation and is given by

$$w_t = \varphi_0 + \varphi_1(S_t) \cdot S_t + \varphi_2 \cdot Exp_t + \varphi_3 \cdot Exp_t^2 + \varepsilon_t^w \quad \#$$

where $\varphi_1(S_t)$ represents the returns to schooling (which may vary with the level of schooling), and

$$\varepsilon_t^w \sim i.i.d. N(0, \sigma_w^2)$$

is a stochastic shock which represents wage dispersion. In the case where we allow the returns to schooling to vary with the level of schooling, we estimate the returns using splines. The variables are defined as follows;

$$\begin{aligned} Education_{6-11} &= Education \\ Education_{12} &= Education - 11 \text{ if } Education > 11 \text{ and } 0 \text{ if not.} \\ Education_{13-15} &= Education - 12 \text{ if } Education > 12 \text{ and } 0 \text{ if not.} \\ Education_{16} &= Education - 15 \text{ if } Education > 15 \text{ and } 0 \text{ if not.} \\ Education_{17-more} &= Education - 16 \text{ if } Education > 16 \text{ and } 0 \text{ if not.} \end{aligned}$$

The return for each level of schooling are obtained after adding up all their splines footnote .

● Assumption 4- Optimal Stopping

We assume that once an individual has entered the labor market, it is impossible to return to school. footnote

This means that each individual maximizes his expected lifetime discounted utility by choosing the optimal time to interrupt schooling and enter the labor market. The control variable, d_t , is such that

$d_t = 1$ if an individual invests in an additional year of schooling at the beginning of period t.
When $d_t = 1$, $I_t = \xi_t$.

$d_t = 0$ if an individual leaves school at the beginning of period t (to enter the labor market).
When $d_t = 0$, $I_t = \exp(w_t)$.

2.2 The Solution

At the beginning of period t, the amount of general human capital acquired, denoted S_t , is simply the total number of years spent at school as of t. That is

$$S_t = \sum_{s=0}^t d_s$$

Experience acquired at t, (Exp_t), is then simply

$$Exp_t = \sum_{s=0}^t (1 - d_s)$$

As it is done often in dynamic optimization problems, the solution to the stochastic dynamic problem can be characterized using recursive methods. Working backward from T , the value of going to school for an additional year and the value of entering the labor market can be characterized using backward recursions. There are two main differences with most empirical dynamic programming models; first, we must solve for the expected instantaneous (per period) utility (since we do not use a random utility specification) and, secondly, we need to isolate the stochastic shock (ε_t^w) in order to obtain a closed-form solution for the probability of choosing to continue school or to enter the labor market.

The value functions associated to the decision to remain in school, $V_t^s(S_t)$, given that an individual has already acquired S_t years of schooling, can be expressed as

$$V_t^s(S_t) = \frac{\xi_t^{1-\alpha} - 1}{1 - \alpha} + \beta E \max[V_t^s(S_t + 1), V_t^w(S_t + 1)]$$

or, more compactly as

$$V_t^s(S_t) = \frac{\xi_t^{1-\alpha} - 1}{1 - \alpha} + \beta E(V_{t+1} \mid d_t = 1) \quad \#$$

where $E(V_{t+1} \mid d_t = 1)$ denotes the value of following the optimal policy next period (either remain at school or start working) and the expected value is taken over the distribution of potential labor market income. Setting d_t to 1 reduces the total time in the labor market but raises entering wages.

The value of stopping schooling accumulation is the value of entering the labor market with S_t years of schooling and no labor market experience, $V_t^w(S_t)$, and is given by

$$V_t^w(S_t) = \frac{(\exp(\varphi_0 + \varphi_1(S_t) \cdot S_t + \varepsilon_t^w))^{1-\alpha} - 1}{1 - \alpha} + \beta E(V_{t+1} \mid d_t = 0) \quad \#$$

where $E(V_{t+1} \mid d_t = 0)$ denotes the discounted expected value of lifetime earnings of starting work in the labor market with t years of schooling, no labor market experience and $T-t$ years of potential specific human capital accumulation ahead. Clearly,

$$E(V_{t+1} \mid d_t = 0) = E \sum_{j=t+1}^T \beta^{j-(t+1)} \left\{ \frac{(\exp(\varphi_0 + \varphi_1(S_t) \cdot S_t + \varphi_2 \cdot Exp_j + \varphi_3 \cdot Exp_j^2 + \varepsilon_t^w))^{1-\alpha} - 1}{1 - \alpha} \right\} \quad \#$$

Closed-form solution to the problem can be obtain by noting that

$$E(V_T) = EU(\exp(w_T)) = E\left(\frac{(\exp(w_T))^{1-\alpha} - 1}{1 - \alpha}\right) \quad \#$$

and that

$$\int_{-\infty}^{+\infty} \frac{(\exp(w_T))^{1-\alpha} - 1}{1 - \alpha} f_T(w) dw = \frac{\exp\{\mu_T(1-\alpha) + \frac{1}{2}\sigma_w^2 \cdot (1-\alpha)^2\} - 1}{1 - \alpha}$$

where $f_T(w)$ is the normal density with parameters μ_T and σ_w^2 and where

$$\mu_T = \varphi_0 + \varphi_1(S_T) \cdot S_T + \varphi_2 \cdot Exp_T + \varphi_3 \cdot Exp_T^2 \quad \#$$

The expected utility of entering any period can be solved using recursive methods (see Stokey and Lucas, 1989).

2.3 Econometric Specification

In the sample data, everyone has at least 6 years of education. As a consequence, we only model the decision to acquire schooling beyond six years. Finally, we assume that the returns to accumulated education and experience at 65 (upon retirement) is 0. footnote

● Assumption 5- Household Characteristics and the Utility of Attending School

The per period utility of attending school is obviously a complicated relationship to model. Both direct and psychic learning costs, which can vary with the level of schooling, are likely to affect schooling decisions. Because individuals raised in wealthy families are likely to receive higher intergenerational transfers or experience a lower level of disutility of attending school, parents' human capital and financial wealth are expected to have a strong effect on the probability of transiting from one grade level to the next. As it is impossible to write down the "full" structural model which would include all these aspects, we specify a reduced-form function for ξ_{it} (the net earnings while at school). The function is allowed to depend on various family background variables and household characteristics as well as measured and unmeasured ability so that the effects of household ability on schooling attainments can be identified from individual own ability.

To preserve positivity of the level of income while at school, we assume that

$$\log \xi_{it} = X'_{it} \delta + \lambda(S_t) \cdot S_t \quad \#$$

The vector X_{it} contains the following variables; parents' education (both mother and father), household income, number of siblings, family composition at age 14 and regional controls while $\lambda(S_t) \cdot S_t$ captures potential changes in parental transfers with the level of schooling or any other

type of nonstationarity in net income while at school. The marginal effect of schooling level on parental transfers, $\lambda(S_t)$, is represented by spline functions. In order to be as flexible as possible, we model the marginal effect using 5 splines (6 to 11 years, 12 years, 13 to 15 years, 16 years and 17 years or more). The number of siblings is used to control for the fact that, other things equal, the amount of parental resources spent per child is declining with the number of siblings. The household composition variable (Nuclear Family) is equal to 1 for those who have been raised with both biological parents (at age 14) and is likely to be correlated with the psychic costs of attending school. The geographical variables are introduced in order to control for the possibility that direct (as well as psychic) costs of schooling may differ between those raised in urban areas and those raised in rural areas and between those raised in the South and those raised in the North. Finally, in order to control for individual ability (affecting the psychic costs of obtaining schooling), we use AFQT scores. footnote Yearly family income is measured in \$1,000 while AFQT scores (reported as a number between 0 and 100) are divided by 10.

2.4 Constructing the Likelihood

Dropping the individual subscript, the probability of investing in schooling in a given year is given by

$$\Pr(d_t = 1) = \Pr[V_t^s(S_t) \geq V_t^w(S_t)] = \Pr\{\frac{\xi_t^{1-\alpha} - 1}{1 - \alpha} + \beta E(V_{t+1} \mid d_t = 1) \geq \frac{(\exp(w_t))^{1-\alpha} - 1}{1 - \alpha} + \beta E(V_{t+1} \mid d_t = 0)\} \quad \#$$

or, equivalently, as

$$\Pr(d_t = 1) = \Pr\{(1 - \alpha)w_t \leq \log[\xi_t^{1-\alpha} + (1 - \alpha)\beta[E(V_{t+1} \mid d_t = 1) - E(V_{t+1} \mid d_t = 0)]]\}$$

and will be expressed as follows, footnote

$$\Pr(d_{it} = 1) = \Pr(\varepsilon_t \leq [h(S_t)]) = \Phi\left(\frac{h(S_t)}{\sigma_w}\right) \quad \#$$

where $h(S_t)$ is given by

$$h(S_t) = \frac{1}{1 - \alpha} \cdot \log\left[(1 - \alpha) \cdot (V_t^s(S_t) - \beta E(V_{t+1} \mid d_t = 0)) + \frac{1}{1 - \alpha}\right] - \varphi_0 - \varphi_1(S_t) \cdot S_t$$

The likelihood function is constructed from data on years of education completed as well as data on labor market wages observed over the panel. The probability of stopping school after having completed $S_t = \tau$ years is given by

$$\Pr(S_{it} = \tau) = \left\{ \prod_{s=1}^{\tau} P(d_{i,s} = 1) \right\} P(d_{i,\tau+1} = 0) \quad \#$$

Note that (13) can easily be adjusted to take into account censoring which occurs for those who are still in school at the end of the panel (1990). The contribution to the likelihood for log wages observed from entrance in the labor market ($\tau_i + 1$) until 1990, $\Pr(w_{i,\tau+1}, \dots, w_{i,1990})$, is given by

$$\Pr(w_{i,\tau+1}, \dots, w_{i,1990}) = \prod_{s=\tau+1}^{1990} \Pr(w_{i,s}) \quad \#$$

where the contribution to the likelihood for each observed wage is given by

$$Pr(w_{it}) = \frac{1}{\sigma_w} \phi\left(\frac{w_{it} - (\varphi_0 + \varphi_1 (S_t) \cdot S_t + \varphi_2 \cdot Exp_t + \varphi_3 \cdot Exp_t^2)}{\sigma_w}\right) \quad \#$$

The model can be improved substantially by the inclusion of observed as well as unobserved ability. As it is usually the case in structural models, unobserved heterogeneity can take multiple forms. In this paper, we restrict ourselves to 2 sources of individual heterogeneity; labor market ability (\tilde{v}_i^w), which affects the intercept term of the wage equation and ability at school (\tilde{v}_i^ξ), which affects the utility of attending school. The distinction between \tilde{v}_i^w and \tilde{v}_i^ξ might be justified if, for instance, skills that are required to perform well in the labor market are not necessarily identical to those skills required at school. As discussed before, the availability of AFQT scores is promising as it is most certainly a proxy for individual ability. To be as general as possible, we assume that individual ability has a measured component (measured by AFQT scores) and an unmeasured component. That is

$$w_{it} = \varphi_0 + \varphi_1 (S_{it}) \cdot S_{it} + \varphi_2 \cdot Exp_{i,t} + \varphi_3 \cdot Exp_{i,t}^2 + \tilde{v}_i^w + \varepsilon_{it}^w$$

$$\log \xi_{it} = X_i' \delta_t + \tilde{v}_i^\xi$$

where each individual is endowed with a pair $(\tilde{v}_i^w, \tilde{v}_i^\xi)$ such that

$$\begin{aligned} \tilde{v}_i^w &= \varphi^w \cdot AFQT_i + v_i^w \\ \tilde{v}_i^\xi &= \delta^\xi \cdot AFQT_i + v_i^\xi \end{aligned}$$

where v_i^w and v_i^ξ represent the unobserved component of individual ability. We assume that v_i^w and v_i^ξ are jointly distributed and that this joint distribution can be approximated with a bi-variate discrete distribution.

● Assumption 7- Ability Bias

In order to investigate the ability bias, we consider the correlation between v_i^{ξ} and v_i^w . The probabilities are expressed as follows:

$$Pr(v^w = v_1^w, v^{\xi} = v_1^{\xi}) = p_1$$

$$Pr(v^w = v_2^w, v^{\xi} = v_1^{\xi}) = p_2$$

$$Pr(v^w = v_1^w, v^{\xi} = v_2^{\xi}) = p_3$$

$$Pr(v^w = v_2^w, v^{\xi} = v_2^{\xi}) = p_4$$

where $v_1^w > v_2^w$ and $v_1^{\xi} > v_2^{\xi}$ and

$$\sum_{j=1}^J p_j = 1 \text{ and } p_j \geq 0 \text{ } j=1,2,\dots,J \text{ (where J=4)} \quad \#$$

The covariance can be expressed as

$$cov(v^w, v^{\xi}) = (p_1.p_4 - p_2.p_3)(v_1^w - v_2^w).(v_1^{\xi} - v_2^{\xi})$$

and, conditional on $(v_1^w \neq v_2^w)$ and $(v_1^{\xi} \neq v_2^{\xi})$, testing for independence (or no correlation) requires testing for $p_1.p_4 - p_2.p_3 = 0$. footnote

Using (13) and (14) and noting that the joint probability of entering the labor market after τ_i years of schooling and observing an entry wage $w_{i,\tau+1}$, is simply given by

$$Pr(d_{i,\tau+1} = 0, w_{i,\tau+1}) = Pr(w_{i,\tau+1})$$

then the likelihood function (conditional on unobserved heterogeneity) is given by

$$L_i(\alpha, \beta, \xi \mid v^w, v^{\xi}) = \left\{ \prod_{s=1}^{\tau_i} P(d_{i,s} = 1) \bullet Pr(w_{i,\tau+1}, \dots, w_{i,1990}) \right\} \quad \#$$

Maximizing the log likelihood function simply requires to write the individual contributions (17) conditional on every possible combinations $\vartheta_j = (v^w, v^{\xi})_j$ and taking a weighted average of all contributions according to the p 's. The contribution to the likelihood, for individual i, is given by

$$\log L_i = \log \sum_{j=1}^J p_j \bullet L_i(\alpha, \beta, \xi \mid \vartheta_j) \quad \#$$

3 The Data

The sample used in the analysis is extracted from 1979 youth cohort of the *The National Longitudinal Survey of Youth* (NLSY). The NLSY is a nationally representative sample of 12,686 Americans who were 14-21 years old as of January 1, 1979. After the initial survey, re-interviews have been conducted in each subsequent year until 1996. In this paper, we restrict our sample to white males who were age 18 or less as of January 1, 1979. We record information on education and on wages for each individual from the time the individual is age 16 up to December 31, 1990. footnote

The original sample contained 3,790 white males. However, we lacked information on family

background variables (such as family income as of 1978 and parents' education) and on AFQT score for 1,161 individuals. footnote The age limit further reduced the sample to 1,299 individuals. Of these 1,299 individuals, it is interesting to note that about 20% had interrupted their schooling attainment for at least one year. This figure, similar to the one reported by Keane and Wolpin (1997), appears to be inconsistent with the optimal stopping assumption and could suggest that some individuals use labor market earnings to finance subsequent education. footnote Before going further, it is important to undertake a deeper investigation of the characteristics of those who do so.

Overall, it has been proven difficult to impute a particular behavioral pattern for those who interrupt school. First, and as pointed out also by Keane and Wolpin (1997), a relatively large fraction of those who leave school non permanently (around 50%) seem to stay out of the labor force totally as their individual records show no wages. The remaining individuals work in the labor market before returning in school but a relatively large fraction work for a small number of weeks. In order to evaluate if the interrupters are systematically different from those who acquire education continuously, we split the individuals between those who had obtain 12 years of schooling or less and those who had obtained 13 years or more and estimated a simple Probit model with an indicator for interruption as the dependent variable. We included eight explanatory variables; parents' education, family income, number of siblings, family composition (nuclear family), area of residence at the age of 16 and AFQT score. The results, not reported to save space, are virtually identical for both groups. For those who have obtained more than 12 years of schooling, all parameter estimates are found insignificant at the 5% level (the lowest p-value was 0.07 for the nuclear family variable while the highest was 0.97 for test scores). For those who obtained 12 years or less, all parameters (except for mother's education) are also estimated without any precision. The lowest p-value was for 0.03 for mother's education and the highest was 0.88 for the number of siblings. Finally, we also regressed education (in log) on household characteristics, test scores and a binary variable indicating if education was acquired continuously or not and found no significant difference between those who interrupted and those who did not. footnote

We take these results as strong evidence that the decision to interrupt schooling (and return later) can be treated as exogenous and uninformative. Empirical evidence shows that it is certainly uncorrelated with family background and, for a larger number of individuals, the decision is most likely dictated by random events such as physical or mental health problems, desire to involve in particular non-market activities or any other events of this type. Given that excluding interrupters would most likely not introduce selection bias problems and to be consistent with our theoretical structure, we decided to exclude these interrupters in the core sample used in this paper, yielding a final sample size of 1,019 white males and focus only on wages observed after individuals have left school. footnote Descriptive statistics for this sample can be found in Table 1A.

The education length variable is the reported highest grade completed as of May 1 of the survey year. Individuals are also asked if they are currently enrolled in school or not. This question allows us to identify those individuals who are still acquiring schooling and therefore to take into account that education length is right-censored for some individuals. The frequencies for various schooling attainments and completions are in Table 1B. There is a large fraction of young individuals who terminate school after 12 years (high school graduation). The next largest frequency is at 16 years and is most likely corresponding to college graduation. The average schooling completed (by 1990) is 12.5 years. footnote

Table 1A- Descriptive Statistics.

Variable	Mean	Std. Dev.	# Individuals
prop. raised in urban areas	0.70	-	1019
father's educ (years)	11.27	3.43	1019
mother's educ (years)	11.34	2.54	1019
family income in 1978	37,685	27,224	1019
number of siblings	3.25	2.15	1019
prop. growing up in nuclear family	0.79	-	1019
prop. raised in southern states	0.28	-	1019
AFQT score	42.43	28.02	1019
education completed (as of May 1990)	12.25	2.45	1019
prop. students (in 1990)	0.01	-	1019
wage 1979 (per hour)	6.61	2.09	86
wage 1980 (per hour)	6.55	2.41	224
wage 1981 (per hour)	6.74	2.64	342
wage 1982 (per hour)	6.93	2.92	496
wage 1983 (per hour)	6.74	2.79	593
wage 1984 (per hour)	7.09	3.23	683
wage 1985 (per hour)	7.65	3.25	719
wage 1986 (per hour)	8.30	3.57	732
wage 1987 (per hour)	9.08	4.16	775
wage 1988 (per hour)	9.79	4.60	830
wage 1989 (per hour)	9.93	4.72	826
wage 1990 (per hour)	10.45	4.86	823

Notes:

Family income and hourly wages are reported in 1990 dollars. Family income is measured as of May 1979 (for 1978). The increasing number of wage observations (until 1988) is explained by the increase in participation rates (schooling completion).

Table 1B- Observed Schooling Attainments in the NLSY (1990).

Schooling	Freq. (%)	# of individuals
6 years	0.4	4
7 years	1.0	10
8 years	4.2	43
9 years	6.7	68
10 years	7.8	79
11 years	9.2	94
12 years	43.1	439
13 years	5.0	51
14 years	4.5	46
15 years	1.5	15
16 years	12.5	127
17 years	1.5	15
18 years	1.5	15
19 years	0.5	5
20 years	0.8	8

4 Empirical Results

The presentation of the results is structured as follows. In section (4.1), we discuss the estimation of three basic model specifications. All three specifications are estimated with AFQT scores (Table 3) and without AFQT scores (Table 2). In the first specification, labor market ability is represented by four types and there is no unobserved heterogeneity affecting the utility of attending school (column 1). In the second specification (column 2), the correlation between labor market ability and ability at school allows us to capture the presence of ability bias. The third specification (column 3) is a restricted version of column 2 in which the covariance between labor market ability and the utility of attending school is forced to be 0. In sub-section 4.2, we analyze whether or not the structural model can predict accurately observed schooling attainments. In sub-section 4.3, we investigate the extent of ability bias by comparing the returns to schooling for various model specifications. In 4.4, we use the structural parameters correlation to simulate how schooling attainments are related to measured ability (AFQT scores) and household characteristics such as mother's education, father's education and family income. Finally, in 4.5, we use the model to simulate the effects of an increase in the return to education and an increase in wage dispersion (variance of the stochastic shock) on educational achievements. This allows us to examine whether individual schooling attainments are more responsive to family background variables or to changes in wage distribution parameters.

4.1 The Structural Parameters

The estimates of the structural parameters for the models without AFQT scores are in Table 2 while those with AFQT scores are in Table 3. In the first column of Table 2 and Table 3, the model is estimated under the assumption that unobserved ability is represented by four types (v_1^w, v_2^w, v_3^w and v_4^w). The estimates which take into account that ability at school and labor market ability are potentially correlated are in column 2 (unrestricted) and in column 3 (restricted) of Table 2 and Table 3.

Turning to the estimates found in column 1 of Table 2, we note that the parameter capturing the degree of relative risk aversion (α) is around 0.62. Note that this indicate a degree of risk aversion much lower than typically found in financial data. footnote The parameter estimate for α is also significantly different than 1 and, along with the value of the likelihood, it indicates a strong rejection of the logarithmic utility function. footnote The estimates of the discount rate is 5% per year; a value certainly close to the long run average risk free rate of interest in the US economy. The return to schooling between 6 to 11 years is estimated to be 0.1244 and we find a 2% premium to high school graduation (at 12 years) while the return to the first 3 years of college, after adding up the splines, is found to be 0.13. The low value of the spline at 16 years indicate no significant return to college graduation. The parameter estimate for experience and its square (0.0733 and -0.0025) indicate concave age earnings profiles.

The estimates surrounding the per period utility of attending school (parental support) show a very strong influence of family background variables. Family income (0.0093), fathers education (0.0658) and mother's education (0.0468) have a positive effect on the utility of attending school. Given family background, the level of parental support appears to decrease with the number of siblings (-0.0612) and appears to be lower for those who have been raised in the south (-0.2348). It is also interesting to note that those who lived in a nuclear family at the age of 14 have a significantly higher value of attending school (0.2015). Whether or not someone has been raised in a rural area does not seem to matter. As expected, the estimates for the 5 splines ($Spline_{6-11}, \dots, Spline_{17-more}$) indicate a strong decline in the probability of continuing school at high school graduation and in the probability to go beyond 4 years of college footnote . As indicated before (see Section 3), these results are consistent with observed schooling attainment frequencies. More details on the goodness of fit of the model can be found in sub-section 4.2

The estimates in column 2 are those obtained when unobserved heterogeneity affecting the

utility of attending school (v^ξ) and unobserved labor market ability (v^w) are allowed to be correlated. In this context, the presence of ability bias arises whenever the covariance between v^ξ and v^w is positive. The estimates for v^ξ (-1.4194 and -1.9144) and v^w (1.4809 and 1.0358), along with the sample proportions (found in the note below Table 2), indicate that the correlation between the utility of attending school and labor market ability is important. The covariance is found to be positive (0.0458) and a likelihood ratio for the null that $\text{cov}(v^\xi, v^w)$ is 0 lead to χ^2_{1df} statistic equal to 58.2 (column 3). There is therefore strong support for the presence of ability bias. Furthermore, after allowing for correlation between v^ξ and v^w , the returns to schooling are found to be much lower (0.0861 between 6 and 11 years and 0.1111 for high school graduation). Further discussions regarding the magnitude of the ability bias and the returns to schooling are delayed to Section 4.3.

It is also interesting to note that, although allowing for ability bias change the returns to schooling significantly, it does not seem to affect the return to experience. The return to experience is around 8% per year and the very robust estimate for experience squared (-0.0036) indicates the presence of concave age earnings profile. One of the most striking effects of allowing for ability bias is the decrease in the effects of family background variables. The effects of father's education drops from 0.0658 to 0.0221 while the effect of mother's education drops from 0.0468 to 0.0163. The effect of family income is also four times smaller (it goes from 0.0093 to 0.0025) while the parameter estimates for siblings goes from -0.0648 to -0.0221. However, both the risk aversion parameter (α) and the discount rate (ρ) appear invariant to the introduction of ability bias; the parameter estimate for α is 0.6502 and the estimate for the discount rate is 0.0421.

Finally, the estimates found in column 3 are those obtained when the covariance between v^ξ and v^w is forced to be 0. There is therefore no ability bias in this model specification. As a consequence, the estimates for the returns to schooling tend to return to higher values (0.1082 between 6 and 11 years) and we still observe a premium to high school graduation (0.0182). The return to experience remain around 8% per year while the degree of risk aversion (0.6123) and the discount rate (0.0458) have not changed much. We also investigated a factor loading version of the model found in column 1 (with 4 ability types) where

$$v^\xi = \zeta \bullet v^w$$

and found positive correlation between v^ξ and v^w . We also found lower returns to schooling than those found in column 1. These estimates are reported in Table 5A.

One of the main advantages of using test scores is that it provides a measure of individual ability (both at school and in the labor market) which can be correlated with other measures of household ability such as parents' education, family income and the like. As a consequence, allowing both the utility of attending school and labor market ability to be correlated with AFQT scores should reduce the effect of family background variables to the point where the measured effect of parents' education and family income on schooling attainments (after controlling for AFQT scores) could be a reliable estimate of the "true causal effects" of family background variables on schooling attainments. In the context of a model in which test scores play a central part, the unobserved heterogeneity terms (v^w, v^ξ) will measure the residual part of individual ability which cannot be explained by test scores (columns 2 and 3).

First, we concentrate on the estimates found in column 1 of Table 3 and in which unobserved ability is represented by four support points. As expected, the estimates for father's education (0.0367), mother's education (0.0155) and family income (0.0073) are found to be lower than those obtained in Table 2 (column 1). Indeed, the parameter estimate for AFQT scores (0.1901) indicate a strong effect of measured ability on the utility of attending school. Not surprisingly, introducing test scores has reduced the return to schooling; the return for basic education (6 to 11) is now 11% per year but we still find a strong marginal return to graduation (0.1284). The effect of AFQT score (divided by 10) on the log wage is 0.0084 and it indicates that there is approximately a 1% wage premium per decile.

The estimates obtained with the more flexible specification of the covariance between v^w and v^ξ are in column 2. A comparison with the estimates found in column 2 of Table 2 indicate that the

effects of parents' education and family income as well as the returns to schooling are at their lowest when test scores are used. The effect of father's education and mother's education drop to 0.0150 and the effect of family income drops to 0.0021. At the same time, the return to basic education is 0.0556 and there seems to be a 2% premium to high school graduation. The return to experience is very robust at 8% per year and there is now a 3% premium per AFQT score decile. Although test scores affect both the utility of attending school and labor market ability, there is still a positive correlation between v^w and v^ξ (the covariance is 0.0487. The null hypothesis that the covariance between v^w and v^ξ is 0 is again strongly rejected (with a χ^2 statistic of 72.7). Finally, although the degree of risk aversion has changes slightly (0.53), the estimate of the discount factor is robust at 4.6% per year. footnote The results obtained when the model is estimated under the null that the covariance between v^w and v^ξ is 0 (column 3) are consistent with what has been observed so far. The absence of control for ability bias raises the return to schooling (0.0820 for the basic education level) and the effects of family background variables are higher than in column 2 but still lower than in the equivalent column of Table 2 where test scores were excluded.

To summarize, we have found that the fundamental preference parameters (the degree of relative risk aversion and the rate of time preference) are estimated quite robustly. The degree of risk aversion is found to be around 0.6 (a degree of risk aversion lower than for a logarithmic utility function) while the discount rate is invariably estimated between 0.04 and 0.05 (a highly credible value). At the opposite, we saw that estimates of the return to schooling appear to vary significantly with the introduction of a control for ability bias (see Section 4.3). Overall, we find that family background variables (especially parents' education) have a significant effect on the utility of attending school. However, after introducing a measure of ability such as AFQT scores in the wage equation as well as in the utility of attending school, we found that the effects of family background variables are reduced significantly.

Table 2
Structural Estimates (Asymptotic standard errors in parentheses.)

Parental Support	(1)	(2)	(3)
Intercept(v_1^ξ)	-3.5574 (0.708)	-1.4194 (0.336)	-2.1222 (0.637)
Intercept(v_2^ξ)	-	-1.9144 (0.348)	-3.3546 (0.722)
Spline _{6–11}	-0.4163 (0.091)	-0.1663 (0.044)	-0.5042 (0.080)
Spline ₁₂	-5.1822 (0.814)	-2.8815 (0.524)	-2.3414 (0.704)
Spline _{13–15}	8.1393 (1.063)	4.3753 (0.721)	4.3603 (0.912)
Spline ₁₆	-13.767 (1.649)	-6.0126 (0.906)	-9.8846 (1.745)
Spline _{17–more}	14.492 (1.856)	5.9913 (0.994)	10.5478 (2.226)
Fam. Income/1,000	0.0093 (0.001)	0.0025 (0.001)	0.0114 (0.002)
Siblings	-0.0612 (0.014)	-0.0226 (0.005)	-0.0897 (0.018)
Rural	-0.0648 (0.065)	-0.0189 (0.017)	-0.0807 (0.048)
South	-0.2348 (0.061)	-0.0759 (0.022)	-0.3294 (0.076)
Nuclear Family	0.2015 (0.093)	0.0751 (0.025)	0.2870 (0.070)
Father's educ (years)	0.0658 (0.012)	0.0221 (0.005)	0.0868 (0.016)
Mother's educ (years)	0.0468 (0.015)	0.0163 (0.005)	0.0592 (0.017)
Wage Function			
Educ splines _{6–11}	0.1244 (0.005)	0.0861 (0.003)	0.1082 (0.004)
Educ splines ₁₂	0.0188 (0.003)	0.0250 (0.002)	0.0180 (0.002)
Educ splines _{13–15}	-0.0144 (0.003)	-0.0109 (0.003)	-0.0184 (0.003)
Educ splines ₁₆	0.0018 (0.004)	0.0045 (0.004)	0.0019 (0.004)
Educ splines _{17–more}	0.0066 (0.004)	0.0097 (0.003)	0.0094 (0.003)
Experience	0.0858 (0.004)	0.0846 (0.004)	0.0807 (0.0038)
Experience ²	-0.0034 (0.0003)	-0.0036 (0.0003)	-0.0031 (0.0003)
v_1^w	1.4450 (0.039)	1.4809 (0.038)	1.3303 (0.027)
v_2^w	1.1236 (0.039)	1.0358 (0.023)	0.8902 (0.028)
v_3^w	0.8157 (0.043)	-	-
v_4^w	0.4738 (0.054)	-	-
σ_w	0.2951 (0.003)	0.3191 (0.003)	0.3203 (0.003)
Risk Av (α)	0.6268 (0.018)	0.6502 (0.019)	0.6123 (0.023)
Disc rate (ρ)	0.0517 (0.012)	0.0489 (0.013)	0.0458 (0.018)
Cov(v^ξ, v^w)	-	0.0539	0.00
L.R. test Cov(v^ξ, v^w) = 0	-	-	58.2
Mean Log Lik.	-3.9432	-4.2072	-4.2358

Note: In column 1, the sample proportions for v_1^w to v_4^w are **0.40, 0.11, 0.38 and 0.10** respectively. The sample proportions for $(v_1^w, v_1^\xi), (v_2^w, v_1^\xi), (v_1^w, v_2^\xi)$ and (v_2^w, v_2^ξ) are **0.47, 0.0001, 0.009 and 0.52** respectively in column 2 and **0.13, 0.13, 0.36 and 0.38** respectively in

column 3.

Table 3- Structural Estimates with Test Scores (Asymptotic standard errors in parentheses)

Parental Support	(1)	(2)	(3)
Intercept(v_1^ξ)	-2.4067 (0.472)	0.3061 (0.122)	0.5258 (0.036)
Intercept(v_2^ξ)	-	-0.1622 (0.122)	-0.4717 (0.039)
Spline _{6–11}	-0.3938 (0.067)	-0.1700 (0.045)	-0.4189 (0.026)
Spline ₁₂	-5.0205 (1.477)	-2.1008 (0.091)	-2.0367 (0.032)
Spline _{13–15}	7.4274 (2.026)	3.0831 (0.126)	3.1408 (0.022)
Spline ₁₆	-10.0084 (2.591)	-4.8438 (0.140)	-4.9212 (0.030)
Spline _{17-more}	10.174 (3.069)	5.0539 (0.236)	4.9999 (0.021)
Fam. Income/1,000	0.0073 (0.001)	0.0021 (0.0005)	0.0080 (0.001)
Siblings	-0.0368 (0.012)	-0.0156 (0.005)	-0.0472 (0.012)
Rural	-0.0773 (0.038)	-0.0283 (0.020)	-0.1284 (0.009)
South	-0.1215 (0.055)	-0.0505 (0.028)	-0.1507 (0.023)
Nuclear Family	0.1240 (0.045)	0.0504 (0.023)	0.2282 (0.029)
Father's educ (years)	0.0367 (0.008)	0.0150 (0.004)	0.0421 (0.009)
Mother's educ (years)	0.0155 (0.016)	0.0150 (0.004)	0.0138 (0.012)
Test score (AFQT/10)	0.1901 (0.023)	0.0996 (0.009)	0.2217 (0.016)
Wage Function			
Educ splines _{6–11}	0.1103 (0.004)	0.0556 (0.004)	0.0820 (0.004)
Educ splines ₁₂	0.0181 (0.003)	0.0216 (0.003)	0.0138 (0.003)
Educ splines _{13–15}	-0.0083 (0.003)	-0.0054 (0.003)	-0.0069 (0.002)
Educ splines ₁₆	0.0015 (0.004)	0.0020 (0.003)	0.0002 (0.003)
Educ splines _{17-more}	0.0063 (0.003)	0.0126 (0.003)	0.0111 (0.003)
Test score (AFQT/10)	0.0084 (0.002)	0.0309 (0.003)	0.0219 (0.003)
Experience	0.0881 (0.004)	0.0865 (0.004)	0.0851 (0.004)
Experience ²	-0.0037 (0.003)	-0.0038 (0.0003)	-0.0036 (0.0003)
v_1^w	1.4836 (0.032)	1.5389 (0.027)	1.3982 (0.027)
v_2^w	1.1680 (0.033)	1.1089 (0.027)	0.9736 (0.028)
v_3^w	0.8641 (0.033)	-	-
v_4^w	0.5282 (0.039)	-	-
σ_w	0.2945 (0.003)	0.3167 (0.003)	0.3183 (0.003)
Risk Av (α)	0.6160 (0.021)	0.5278 (0.028)	0.4963 (0.025)
Disc rate (ρ)	0.0502 (0.016)	0.0456 (0.017)	0.0585 (0.013)
Cov(v^ξ , v^w)	-	0.0487	0.00
L.R. test Cov(v^ξ , v^w) = 0	-	-	72.7
Mean Log Lik.	-3.7629	-4.0021	-4.0378

Note: In column 1, the sample proportions for v_1^w to v_4^w are 0.38, 0.41, 0.09 and 0.12 respectively. The sample proportions for (v_1^w, v_1^ξ) , (v_2^w, v_1^ξ) , (v_1^w, v_2^ξ) and (v_2^w, v_2^ξ) are 0.45, 0.001, 0.01 and 0.54 respectively in column 2 and 0.23, 0.25, 0.24 and 0.27 respectively in column 3.

4.2 Can a Structural Model fit Observed Schooling Attainments ?

At this stage, it is natural to ask whether or not our model is able to explain observed schooling attainments accurately. As seen from Table 1B, there is a large fraction of young individuals who terminate school after 12 years (high school graduation). The next largest frequency is at 16 years and is most likely corresponding to college graduation. These features seem to indicate that both the returns to schooling and the utility of continuing school beyond a given level may vary with the level of schooling. For instance, regulations of the minimum school enrollment age (16 years in most states), large increase in tuition costs and fees beyond high school, decline in parental support, or wage premia for those graduating from high school could explain the large frequency at 12 years.

The predicted frequencies (for various model specifications) are found in Table 4 and indicate that our approach is indeed able to predict schooling attainments quite accurately. This is true at low levels of schooling (6 to 10 years) as well as at higher levels (17 years or more). For all model specifications reported in Table 4, a simple χ^2 test statistic (with 4 degrees of freedom) typically fails to reject the null hypothesis that the model is properly specified. The critical value (at the 5% level) is 9.49. In particular, we note that model specifications allowing for the correlation between the utility of attending school and the intercept of the wage equation (controlling for ability bias) are found to fit data on schooling attainments much better than model specifications which ignore ability bias.

The capacity of the model to fit observed schooling frequencies is, however, not the only criterion by which the empirical relevance of our model can be judged. Within a dynamic framework, our estimates of the subjective discount factor can also be used to evaluate the accuracy of our model. Although a review of the literature on empirical dynamic discrete choices would indicate that estimates of the subjective discount rates tend to fluctuate largely across types of studies (job search, retirement, human capital accumulation), our estimates are most credible. One of our most robust finding is a subjective discount rate averaging 4% to 5% per year which, as pointed out earlier, is very close to average risk free interest rates computed from historical time series (Kocherlakota, 1996).

Finally, in order to evaluate how our model fits data on wages (as opposed to OLS estimates), we used expected and observed wages (as of 1990) to evaluate the fraction of observed wages explained by our model as opposed to the fraction explained by OLS estimates (frequently used in the reduced-form literature). On average, our various model specifications explained between 24% and 25% of the variations in observed wages; a value comparable to what is typically observed in the reduced-form literature (see Card, 1998). footnote This means that our dynamic programming model is capable to explain accurately observed schooling attainments (something virtually impossible in a reduced-form framework) while fitting wage data as well as reduced-form models. footnote

Table 4
Observed vs Predicted Schooling Attainments

Frequencies of various Schooling attainments

Table 2 (1) Table 2 (2) Table 3 (1) Table 3 (2)

	Pred. (%)	Pred. (%)	Pred. (%)	Pred. (%)	Actual (%)
Control for Ability bias	no	yes	no	yes	
Schooling					
6-10 years	0.1998	0.2062	0.2066	0.2056	0.2002
11-12 years	0.4992	0.5175	0.4988	0.5188	0.5231
13-14 years	0.1141	0.0887	0.1065	0.0873	0.0952
15-16 years	0.1286	0.1407	0.1332	0.1415	0.1354
17-more	0.0481	0.0469	0.0549	0.0468	0.0422
Good. of fit					
χ^2_{4df} statistic	5.90	1.31	6.06	1.47	

Note: The critical value for a χ^2_{4df} . (at a 5% level) is equal to 9.49.

4.3 Ability Bias and the Return to Education

Although theoretical models of schooling decisions are naturally set in a dynamic and stochastic framework, the empirical literature is largely dominated by linear wage equations in which educational attainment is sometimes instrumented out. Labor economists have, up to now, ignored most surrounding dimensions of schooling decisions and focussed on simple wage regression frameworks in order to estimate the return to schooling. It has however been recognized for a long time that ordinary least squares (OLS) estimates can seriously over estimate the true return to schooling. As a consequence, several economists have tried to use instrumental variables or institutional changes in order to obtain more reliable estimates of the effect of an additional year of education. Although many authors have reported that the estimates of the returns to schooling obtained by instrumental variable methods tend to exceed OLS estimates, many estimates are meant to measure a local return to schooling (say around high school graduation or around minimum school leaving age) and it is often difficult to generalize the results to higher and lower levels of schooling. As pointed out by Staiger and Stock (1997), estimates of the returns to schooling obtained using instrumental methods are often faced with the “weak instrument” problem and reported standard errors may be misleading. footnote

In the present context, and in the case where test scores are not included, ability bias arises whenever the covariance between v^w and v^ξ is positive and the consequence of ability bias can easily be investigated by comparing the returns to schooling obtained from models estimated under the null hypothesis ($\text{cov}(v^w, v^\xi) = 0$) and those obtained from models estimated under the alternative. When test scores (measured ability) are included in the analysis and affect wages as well as the utility of attending school, ability bias can arise either from measured ability or unobserved ability.

A summary of all estimates of the returns to schooling are found in Table 5A (without test scores) and Table 5B (with test scores). A comparison between estimates of column 1 and column 2 reveals a clear presence of ability bias. Overall, and depending on the level of schooling, estimates obtained under the alternative hypothesis can be up to 49% lower than those obtained under the null (no ability bias). For instance, in the first column of Table 5A, which refers to the structural estimates found in the first column of Table 2, the estimates for basic education (6 to 11 years) drops from 0.1244 to 0.0890 and the return to high school graduation drops from 0.1432 to 0.1133 when ability bias is taken into account. A comparison between estimates of column 3 and column 4 also indicate a clear presence of ability bias. Overall, the estimates reveal that, except for a 2% to 3% premium to high school graduation (when compared to the return to basic education), the returns to schooling are relatively constant with schooling levels. footnote

As noted before, introducing measured ability through test scores has reduced the returns to schooling for all model specifications. For instance, a comparison between column 1 and column 3 of Table 5B with their comparable columns in Table 5A, reveal that the introduction of AFQT scores has reduced the return to schooling by 30%. Indeed, introducing test score has an effect comparable to allowing for ability bias through unobserved ability. The most general results are however found in column 2 and column 4 of Table 5B. In both cases, the utility of attending school and labor market wages are potentially correlated through both measured ability and unmeasured ability. The results are quite spectacular. The returns to schooling are around 6% and 7% per year (again with a 2% premium to high school graduation) and, despite the introduction of test scores, there is still a clear negative correlation between unmeasured ability affecting the utility of attending school and unmeasured ability affecting wages.

Overall, estimates of the return to basic schooling obtained after controlling for ability bias are around 30% lower than those obtained when ignoring it. For the return to high school graduation, estimates controlling for ability bias are around 25% below those obtained when ability bias is ignored. Finally, it should be noted that a return to schooling slightly higher than subjective discount rates is certainly not inconsistent with average schooling attainments. Given that the minimum age at which students can leave school is 16 years, at which time students have acquired

an average of 11 years of schooling in the NLSY, an average schooling attainment of 12.3 years (as of 1990) would be difficult to explain if the true returns to schooling are between 12% and 15% per year (as claimed in the reduced-form literature).

Table 5A
Ability Bias and the Returns to Schooling without Test scores

	(1)	(2)	(3)	(4)
Tab 2 (1) Tab 2 (foot.) Tab. 2, col 3 Tab. 2, col 2				

Control for				
Ability bias	no	yes	no	yes
AQFT scores	no	no	no	no
	returns	returns	returns	returns
Levels				
6-11 years	0.1244	0.0890	0.1082	0.0861
12 years	0.1432	0.1133	0.1262	0.1111
13-15 years	0.1288	0.0961	0.1078	0.1002
16 years	0.1306	0.1089	0.1097	0.1047
17-more	0.1372	0.1149	0.1191	0.1144
Discount rate	0.0517	0.0429	0.0458	0.0489

Table 5B
Ability Bias and the Returns to Schooling with Test scores

Tab 3 (1) Tab 3 (foot.) Tab. 3, col 3 Tab. 3, col 2

Control for			
Ability bias	no	yes	no
AQFT scores	yes	yes	yes
	returns	returns	returns
Levels			
6-11 years	0.1103	0.0566	0.0820
12 years	0.1284	0.0767	0.0958
13-15 years	0.1201	0.0683	0.0889
16 years	0.1216	0.0758	0.0891
17-more	0.1279	0.0839	0.1002
Discount rate	0.0517	0.0534	0.0458
			0.0489

4.4 The Effect of Household Characteristics and Ability on Schooling Attainments

The positive relationship between individual schooling attainments and parents' education is well known. For instance, in the NLSY, father's and mother's education explain around 23% of the observed variation in schooling. However, despite the strong correlation between schooling attainments and parents' education, the true causal relationship between parents' and children's education is hard to establish. Both parents' education and family income are likely to be correlated with household's ability which, in turn, are likely to be correlated with individual ability. footnote The availability of a measure of ability such as AFQT scores is however promising. After conditioning on test scores and unobserved ability, the residual effects of parents' education, family income and the like are likely to measure the true causal effect of parents' human capital and income.

In order to obtain a unit-free measure of the effects of father's education, mother's education, family income and measured ability on schooling attainments, we simulated the effects of an increase in each of these variables within a range of value covering the 20th centile to the 80th centile and computed elasticities at several values over the relevant range. The estimates are in Table 6A (father's education), Table 6B (mother's education), Table 6C (family income) and Table 6D (for AFQT scores). The resulting elasticities can therefore be used to infer which of family background variables have the highest influence on schooling attainments.

The simulations indicate that schooling attainments are relatively elastic with respect to measured ability (test scores). However, schooling attainments are quite inelastic with respect to household characteristics. Among household characteristics, the highest elasticities are for father's education (around 0.03 and the median value of father's education) and mother's education (averaging 0.02 at the median). It is interesting to note that schooling attainments are particularly inelastic with respect to household income (below 0.02 at median family income).

These results suggest some important policy implications. First, it appears that policies aimed at increasing schooling attainments, which are based financial subsidies, are likely to be ineffective. For instance, the amount of subsidies needed (increase in household income) to increase schooling attainments by 1 year would have to be enormous. As an illustration, after controlling for observed and unobserved ability and considering an individual living in a household with the median income in the sample, inducing the individual to obtain an additional year of schooling beyond high school would require a level of family income 3 to 4 times the sample median. footnote Second, although parents' education (especially father's education) appear to be more important than family income, the relatively low values for the computed elasticities (0.03 for father's education) indicate that, after controlling for individual observed and unobserved ability, the "true intergenerational education correlation" is relatively small. As a comparison, the elasticities indicate that, after controlling for ability and family income, an individual raised with parents who obtained 13 years of schooling would obtain around 0.05 year more than a similar individual raised by parents who have both obtained 12 years of schooling.

4.5 Simulating the effects of a Change in the Wage Distribution

After having investigated the elasticities of schooling attainments with respect to household/family characteristics, it is natural to analyze how education choices are influenced by economic variables external to the family environment. In particular, it would be interesting to ask how schooling attainments vary when the return to schooling varies and when wage dispersion (represented by the variance of the stochastic shock affecting wages) also varies. In Table 7A, we report a change in schooling attainments following a 25% increase in the return to high school education, a 25% increase in the return to college and a 25% increase in the variance of wages (σ_w). To obtain a better comparison between family background and economic parameters, we also

computed elasticities for various model specifications. These elasticities are in Table 7B.

Turning to Table 7A, we see that, not surprisingly, increasing the return to high school education (including the return to graduation) does not increase average schooling attainments by much. This is clearly a reflection that the average schooling attainments are close to 12 years. However, increasing the returns to college by 25% raises mean schooling attainments by 4 years. Finally, increasing wage dispersion (increasing σ_w), seems to reduce schooling attainments by half a year. The larger sensitivity of schooling attainments to changes in the returns to college education is well illustrated in Table 7B. The elasticities are above 1 (1.3 on average) and are much larger than those obtained for high school education (between 0.11 and 0.03) and those obtained for wage dispersion, which fluctuate between -0.25 and -0.02. It is particularly interesting to note that the elasticities of mean schooling attainments with respect to the returns to college education are quite robust and do not fluctuate much across model specifications.

At this stage, we are now in position to evaluate the relative importance of household characteristics as opposed to wage distribution parameters (such as the return to college education) in explaining schooling attainments. Although the ranking of elasticities appear to be invariant to the model specification, it is easier to do so for a particular model specification. For this purpose, we choose the model specification which takes into account ability bias in a flexible way and in which test scores are included (Table 3, column 3). By far, the return to college education has the strongest impact on schooling attainments. Indeed, of all variables considered, it is the only elasticity which exceeds 1.0 (it is around 1.3). After the returns to college education, the highest elasticities are found for measured ability (AFQT scores elasticities average around 0.09) and for wage dispersion (-0.05). Schooling attainment elasticities with respect to household human capital, such as father's education (around 0.03) and mother's education (averaging 0.02), are found to be smaller than elasticities with respect to changes in the wage distribution. Finally, it is interesting to note that family income elasticities (below 0.02 at median family income) are slightly below mother's education and appear to be the lowest elasticity footnote .

Table 6A
Illustrating the Effect of Father's Education on Schooling Attainments

	Table 2 (2)	Table 2 (3)	Table 3 (2)	Table 3 (3)
Ability bias control	yes	no	yes	no
AFQT control	no	no	yes	yes
	$\frac{\Delta\%Educ}{\Delta\%father's\ educ}$	$\frac{\Delta\%Educ}{\Delta\%father's\ educ}$	$\frac{\Delta\%Educ}{\Delta\%father's\ educ}$	$\frac{\Delta\%Educ}{\Delta\%father's\ educ}$
Father's educ				
6 to 7 years	0.0894	0.0547	0.0399	0.0198
10 to 11 years	0.1369	0.0736	0.0720	0.0270
12 to 13 years	0.1574	0.0998	0.0683	0.0392
15 to 16 years	0.2362	0.1819	0.1076	0.0485

Table 6B
Illustrating the Effect of Mother's Education on Schooling Attainments

	Table 2 (2)	Table 2 (3)	Table 3 (2)	Table 3 (3)
Ability bias control	yes	no	yes	no
AFQT control	no	no	yes	yes
	$\frac{\Delta\%Educ}{\Delta\%mother's\ educ}$	$\frac{\Delta\%Educ}{\Delta\%mother's\ educ}$	$\frac{\Delta\%Educ}{\Delta\%mother's\ educ}$	$\frac{\Delta\%Educ}{\Delta\%mother's\ educ}$
Mother's educ				
6 to 7 years	0.0680	0.0476	0.0264	0.0066
10 to 11 years	0.1001	0.0552	0.0359	0.0091
12 to 13 years	0.1183	0.0599	0.0391	0.0196
15 to 16 years	0.1555	0.1473	0.0605	0.0222

Note: Simulations are performed while keeping the remaining characteristics to their sample average.

Table 6C
Illustrating the Effect of Family Income on Schooling Attainments

	Table 2 (2)	Table 2 (3)	Table 3 (2)	Table 3 (3)
Ability Bias control	yes	no	yes	no
AFQT control	no	no	yes	yes
	$\frac{\Delta\% Educ}{\Delta\% fam.income}$	$\frac{\Delta\% Educ}{\Delta\% fam.income}$	$\frac{\Delta\% Educ}{\Delta\% fam.income}$	$\frac{\Delta\% Educ}{\Delta\% fam.income}$

Family Income

15,000 to 17,000	0.0191	0.0141	0.0127	0.0074
26,000 to 28,000	0.0349	0.0241	0.0193	0.0128
34,000 to 42,000	0.0478	0.0313	0.0278	0.0167
41,000 to 52,000	0.0481	0.0383	0.0342	0.0212
55,000 to 57,000	0.0842	0.0677	0.0442	0.0357

Table 6D
Illustrating the Effect of Ability on Schooling Attainments

	Table 2 (2)	Table 2 (3)	Table 3 (2)	Table 3 (3)
Ability bias control	yes	no	yes	no
	$\frac{\Delta\% Educ}{\Delta\% Test score}$			

AFQT/10

2.0 to 3.0	-	-	0.0436	0.0360
4.0 to 5.0	-	-	0.0969	0.0497
5.0 to 6.0	-	-	0.1469	0.0900
6.0 to 7.0	-	-	0.2283	0.1645
8.0 to 9.0			0.3684	0.6662

Note: Simulations are performed while keeping the remaining characteristics to their sample average.

Table 7A
Changes in mean Schooling following a 25% increase in Wage Distribution Parameters

	Table 2 (2)	Table 2 (3)	Table 3 (2)	Table 3(3)
Ab. bias Control	yes	no	yes	no
AFQT control	no	no	yes	yes
Parameter				
High School	0.22 year	0.34 year	0.09 year	0.10 year
College	4.12 years	4.30 years	3.89 years	3.84 years
Variance (σ_w)	-0.76 year	-0.29 year	-0.38 year	-0.05 year

Table 7B
Mean schooling Elasticities with respect to Wage Distribution Parameters

	Table 2 (2)	Table 2 (3)	Table 3 (2)	Table 3(3)
Ab. bias control	yes	no	yes	no
AFQT control	no	no	yes	yes
	$\frac{\Delta\% Educ}{\Delta\% param.}$	$\frac{\Delta\% Educ}{\Delta\% param.}$	$\frac{\Delta\% Educ}{\Delta\% param.}$	$\frac{\Delta\% Educ}{\Delta\% param.}$
Parameter				
High School	0.0727	0.1136	0.0294	0.0327
College	1.3631	1.4369	1.2702	1.2569
Variance (σ_w)	-0.2500	-0.0969	-0.1241	-0.0164

Note: The mean schooling attainments are computed for a representative agent.

5 Conclusion

We have estimated a dynamic expected utility model of schooling decisions in which individual preferences are represented by a power utility function. Unlike most dynamic discrete choice models found in the literature, our model allows for the estimation of the rate of time preference and the degree of risk aversion as separate parameters. Our model was used to investigate two main issues; the importance of ability bias in estimating the returns to schooling and the elasticities of expected schooling attainments with respect to changes in the wage distribution (the return to education and wage dispersion) and a change in economic variables internal to the household (family income and parents' education).

We find a degree of risk aversion much smaller than usually found in the finance literature; our estimates of the degree of relative risk aversion is around 0.6. We find a subjective annual discount rate between 4% and 5%. The model fits the data quite well and is able to explain pronounced frequencies around high school and college graduation. Overall, and depending on the level of schooling, estimates obtained under the alternative hypothesis can be as much as 40% lower than those obtained under the null (no ability bias). The most reliable estimates of the returns to schooling are around 6% and 7% per year (with a 2% premium to high school graduation) and, despite the introduction of test scores, there is still a clear positive correlation between unmeasured ability affecting the utility of attending school and unmeasured ability affecting wages.

Various simulations indicate that schooling attainments are particularly elastic with respect to the return to college education (elasticities are around 1.3) and relatively elastic with respect to measured ability (test scores). However, schooling attainments are quite inelastic with respect to household characteristics. Among household characteristics, the highest elasticities are for father's education (around 0.03 and the median value of father's education) and mother's education (averaging 0.02 at the median). It is interesting to note that schooling attainments are particularly inelastic with respect to household income (below 0.02 at median family income). The true correlation between schooling attainments and parents' education (after conditioning on observed and unobserved ability) is therefore found to be quite low.

Finally, it should be noted that our estimates of the returns to education and of subjective discount rates are able to reconcile long run average risk free interest rates observed in US data (expected to be close to subjective discount rates) and observed schooling attainments. As noted before, observed completed schooling attainments (between 12 and 13 years) and average schooling attainments at age 16 (slightly below 11 years) are hardly compatible with returns to education between 10% and 15% (as suggested by various instrument variable techniques) and with subjective discount rates anywhere near the long run average risk free interest rates.

At this stage, several avenues for future research come to mind. First, it is realistic to presume that individual ability affects not only the intercept term of the wage equation but also the return to education itself. It would therefore be interesting to determine the relative importance of test scores and unobserved ability in explaining population heterogeneity in the returns to education. Second, given that population heterogeneity in labor market ability and ability at school is found to be important, is discounting heterogeneity a key factor explaining schooling attainments? Finally, it would be interesting to compare the effects of household characteristics, such as parents' education and income, found in a liberal economy (such as the US) with those found for a Welfare State (such as Sweden) in which education is highly subsidized. These are avenues for research which we are likely to undertake.

bibitem Angrist, Joshua and Krueger, Alan (1991) "Does Compulsory School Attendance Affect Schooling and Earnings" *Quarterly Journal of Economics*, 106: 979-1014

bibitem Ashenfelter, Orley and Krueger, Alan (1994) "Estimates of the Economic Return to Schooling from a New Sample of Twins" *The American Economic Review*, 84 (5), 1157-1173.

bibitem Belzil, Christian and Hansen, Jörgen (1998) "Subjective Discount Rates, Unobserved Ability, Family background and the Return to Schooling" CRDE Working paper 2597,

CRDE and University of Montreal.

- bibitem** Ben-Porath, Yoram (1967) "The Production of Human Capital and the Life Cycle of Earnings" *Journal of Political Economy*, 75 (4): 352-65.
- bibitem** Blackburn, McKinley and Neumark, David (1993) "Omitted Ability Bias and the Increase in the Return to Schooling" *Journal of Labor Economics*, 11 (3), 521-543.
- bibitem** Cameron, Stephen and Heckman, James (1998) "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males" *Journal of Political Economy*, 106 (2), 262-333.
- bibitem** Card, David (1998) "The Causal Effect of Education on Earnings" Working Paper, University of California at Berkeley.
- bibitem** Eckstein, Zvi and Kenneth Wolpin (1989) "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey" *Journal of Human Resources*, 24 (4), 562-598.
- bibitem** Harmon, Colm and Walker, Ian (1995) "Estimates of the Economic return to Schooling for the United Kingdom" *American Economic Review*, 85 (5), 1278-1286.
- bibitem** Heckman, James and Singer, Burton (1984) "A method for Minimizing the impact of distributional assumptions in econometric models for duration data" *Econometrica*, 52 (2), 271-320.
- bibitem** Kane, Thomas (1994) "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education" *Journal of Political Economy*, 102, 878-911.
- bibitem** Keane, Michael P. and Wolpin, Kenneth (1997) "The Career Decisions of Young Men" *Journal of Political Economy*, 105 (3), 473-522.
- bibitem** Kocherlakota, Narayana (1996) "The Equity Premium: It's Still a Puzzle" *Journal of Economic Literature* 24 (1), 42-71.
- bibitem** Lazear, Edward (1980) "Family Background and Optimal Schooling Decisions" *Review of Economics and Statistics*, 62 (1), 42-51.
- bibitem** Rust, John (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Analysis of Harold Zurcher" *Econometrica*, 55 (5), 999-1033.
- bibitem** Staiger, Douglas and James H. Stock (1997), "Instrumental Variables Regression with Weak Instruments" *Econometrica*, 65: 557-586.
- bibitem** Stokey, Nancy and Lucas, Robert E. (1989) Recursive Methods in Economic Dynamics. Harvard University Press
- bibitem** Taber, Christopher (1998) "The Rising College Premium in the Eighties: Return to College or Return to Ability", Working Paper, Northwestern University.
- bibitem** Taubman, Paul (1976) "The determinants of Earnings: Genetics, Family and Other environments: A Study of White Male twins" *American Economic Review*, 66 (5), 858-870.
- bibitem** van den Berg, Gerard, Lindeboom, Marten and Ridder, Geert (1994) "Attrition in Longitudinal Panel data and the Empirical Analysis of Dynamic Labour Market Behaviour", *Journal of Applied Econometrics*, 9, 421-435.
- bibitem** Wolpin, Kenneth (1987) "Estimating a Structural Search Model: The Transition from School to Work" *Econometrica*, 55 (4), 801-817.