Intergenerational Transfers and the Rate of Time Preference in a Dynamic Model of Schooling Decisions

Christian Belzil
Concordia University, Dept. of Economics,
Montreal, Canada, H3G 1M8

Jörgen Hansen
IZA, University of Bonn,
Bonn, Germany

April 24, 2000

Abstract

Using a structural dynamic programming model with multi-variate unobserved heterogeneity, we investigate the relative importance of the rate of time preference and household human capital in explaining differences in schooling attainments and wages. We use the structural parameter estimates to evaluate the true intergenerational education correlation and its macroeconomic counterpart; the human capital intergenerational transfers in a growing economy. Given a certain level of school ability and market ability, the variations in schooling attainments and wages imputed to discount rate heterogeneity are found to be systematically higher than the variations imputed to family background variables (between 50% to 80% higher). When school ability and market ability are allowed to be correlated with family human capital (the model specification that reaches the highest likelihood value and fits the data best), the true intergenerational partial correlations between son and father's education (0.22) and between son and mother's education (0.03) are found to be significantly below the correlations found in the data. Various simulations indicate that increasing the level of schooling of the current generation by one year raises schooling attainments of the next generation by 0.1 to 0.3 year.

Key Words: Intergenerational Transfers, Returns to Education, Subjective Discount Rates, Human Capital, Growth, Schooling Decisions.

JEL Classification: J2-J3.

1 Introduction

Individual schooling attainments are one of the key components of the level of human capital in an economy. They are an important determinant of income distribution and are often thought to be one of the key factor explaining the wealth of nations as well as cross-nation differences in economic growth. The recent revival of neo-classical growth models is, indeed, largely based on human capital theory (Lucas, 1988, Barro and Sala-i-Martin, 1995).[1] At the micro level, models of schooling attainments are naturally set in a dynamic and stochastic environment. While enrolled in school, young individuals typically receive parental support. Although parental support is usually unobservable to the econometrician, it is expected to be highly correlated with household human capital and income. The effect of parental background on educational achievements has been well documented in a reduced-form framework (Kane, 1994 and Lazear, 1980) or in a semi-structural framework (Cameron and Heckman, 1998). Within a structural framework, the effects of parental background can take various forms. Households with higher income can transfer more resources to their children and reduce substantially the opportunity cost of schooling. At the same time, innate ability, typically correlated with household human capital, should also have an impact on schooling attainments and wages. As a consequence, labor economists should be skeptical about drawing strong conclusions about the causal effect of parents' human capital from the empirical correlation between various family background variables and individual schooling attainments. Reported correlations between parents' education and children's education could diverge seriously from the "true" intergenerational education correlation.[2]

Although the notion of a true intergenerational education correlation has not raised much interest amongst empirical labor economists, it naturally arises in the dynamic macroeconomic literature concerned with economic growth, overlapping generations and intergenerational transfers. If the true intergenerational education correlation can be inferred from micro data, it can provide a good estimate of the intergenerational human capital transfers that take place at the aggregate level, in a growing economy. For instance, knowledge of the true intergenerational education correlation can allow economists to simulate the effect of an exogenous increase in human capital (accompanied by the resulting growth in income) of the current generation on schooling attainments and labor market productivity of the next generation and help determine if education is also a consequence as well as a cause of economic growth.

Apart from family background variables (measuring household human capital) and individual ability, schooling attainments are strongly affected by the rate of time preference, that is the rate at which one is willing to delay entrance in the labor market in order to raise future wages. As the rate

[1] Although the links between schooling and private wages is well established at the micro level, the relationship between economic growth and education is substantially weaker. This paradox is currently the object of a large amount of work (see Topel, 1999, for a survey).

[2] Indeed, estimating the true intergenerational education correlation has similarities with the problem of estimating the true return to schooling in the presence of ability bias.

of time preference is intrinsically unidentified in reduced-form models, obtaining empirical evidence on the relative importance of individual discount rates and human capital intergenerational transfers requires researchers to use structural dynamic discrete choices econometric techniques. The literature on the estimation of structural dynamic programming problems of schooling decisions is relatively recent. Keane and Wolpin (1997) have used a structural dynamic programming model of schooling and occupational decisions using a cohort of the NLSY while Eckstein and Wolpin (1999) use a dynamic programming model to evaluate the effect of youth employment on academic performance of young Americans. Belzil and Hansen (2000) have used a structural dynamic programming model to estimate the returns to schooling and investigated the relative importance of the "ability bias" and the "discount rate heterogeneity bias".[3]

The main objective of the present paper is to estimate a structural model of schooling decisions in which the separate effects of subjective discount rates, labor market ability and intergenerational transfers on schooling attainments can be identified and which can be used to evaluate the relative importance of discount rate heterogeneity and family background variables. The model, which is estimated in the present paper, is similar to the one discussed in Belzil and Hansen (2000). We pay a particular attention to the following 5 questions.

1. How much of the variations in schooling attainments are due to individual heterogeneity in the rate of time preference as opposed to differences in household human capital (parents' education, family income and the like)?

2. Are expected schooling attainments more elastic with respect to individual differences in discount rates or differences in family human capital?

3. How much of the variations in predicted wages (over the life cycle) can be attributed to discount rate heterogeneity as opposed to heterogeneity in household human capital ?

4. Does the true intergenerational education correlation, predicted by the structural dynamic programming model, differ substantially from the correlation measured in the data?

5. What is the effect of an exogenous increase in the level of education and income of the current generation on schooling attainments of the next generation?

[3]Belzil and Hansen (2000) estimate the return to schooling using flexible methods (without assuming equality between the local and the average return). The function representing the local returns to schooling is assumed to be composed of 9 segments. They find the ability bias to be important; OLS estimates tend to be between 20% to 30% higher than the true return (in a single return framework). The discount rate heterogeneity bias is also found to be important. Contrary to conventional wisdom, they find an inverse relationship between subjective discount rates and the average return to schooling. This is explained by the fact that the local returns are smoothly increasing and it implies that disadvantaged workers experience a lower return to schooling.

As far as we know, none of these questions have ever been investigated. Following Belzil and Hansen (2000), we estimate a finite horizon dynamic programming model which is solved using recursive methods. In our model, schooling decisions affect not only future wages, but also lifetime employment rates as well as employment quality (non-wage benefit). Individuals are endowed with household characteristics (such as parents' education, number of siblings, presence of both biological parents at age 14 and family income), innate ability (ability affecting the utility of attending school and labor market ability) and a rate of time preference (a preference parameter). Both types of ability have a measured component (Armed Forces Qualification Test scores) and an unmeasured component (drawn from a multivariate discrete distribution) and are allowed to be correlated with each other. Given their endowment, individuals decide on the optimal allocation of time between school and the labor market. The model is able to take into account the spurious correlation between market ability and schooling attainments. This is known as the 'ability bias'. As family background variables are strongly correlated with Armed Forces Qualification Test scores (AFQT), our model also captures the potential correlation between ability and household human capital. This prevents us from over-stating the effect of parental background on schooling attainments.

The estimation of our model is computer intensive. In order to estimate of model where the degree of flexibility is high enough to capture the importance of both unobserved ability (in school and in the market) and discount factor heterogeneity, the model must be solved recursively for up to 8 types of individuals. For this reason, we concentrate on a model specification which can be solved in closed-form. The model is implemented on a panel of white males taken from the National Longitudinal Survey of Youth (NLSY). The panel covers a period going from 1979 until 1990. The main results are the following.

Overall, discount factor heterogeneity is found to be a key factor of schooling attainments. Indeed, it is found to be more important than household human capital. For a given level of ability in school and in the market, the variations in individual schooling attainments explained by discount factor heterogeneity are between 50% and 80% higher than the variations explained by family background variables. As a consequence, the variations in predicted wages due to individual heterogeneity in the rate of time preference also exceed the variations due to differences in family background variables by a similar magnitude. When the relative importance of both factors is evaluated by elaticities, mean schooling attainments are found to be 3 to 5 times more elastic (in absolute value) with respect to individual discount rates than with respect to father's education (the family background variable which has the highest elasticity). Not surprisingly, we find that the intergenerational education correlation predicted by the dynamic programming model is sensitive to the inclusion of AFQT scores. This is explained by the fact that AFQT scores are strongly correlated with family background, especially parents' education. In the model specification with test scores (the model specification that reaches the highest likelihood value and fits the data best), the true partial correlation between son and father's schooling (0.22) is around 25% lower than the correlation found in the data. The difference between the true partial correlation between son and mother's education (0.03) and the correlation disclosed by the data (0.23) is even larger. A simulation of the effect of an exogenous increase in the level of schooling of the current generation on schooling attainments of the next generation (the macroeconomic counterpart of the intergenerational education correlation measured from cross-section data) indicates a modest increase ranging between 0.1 year to 0.3 year. It nevertheless illustrates that education can be viewed as one of the consequences as well as one of the causes of economic growth.

The main features of the paper are the following. Section 2 is devoted to the presentation of the theoretical model as well as its econometric specification. Section 3 contains a brief description of the sample used in this paper (NLSY). The main empirical results are discussed in Section 4. The conclusion is in Section 5.

2 The Model

In this section, we outline both the theoretical and the econometric structure of the model. Although the model is almost identical to the one found in Belzil and Hansen (2000), we present it in order to have the paper self-contained. The structure of our model can be summarized as follows. Given their endowments (household human capital, school and market abilities and a rate of time preference), individuals choose between an additional year of schooling and entering the labor market while knowing that they must retire when they are sixty-five years old. Wages grow stochastically with experience. For added realism, we take into account that individuals might experience unemployment over their lifetime and that the human capital accumulation process is not necessarily continuous; that is some individuals can experience temporary interruptions due to various causes such as health problems, travel activities, incarceration or other reasons. This is done by introducing a transition probability to a state in which schooling acquisition is interrupted, with an exogenous return probability driving the length of the interruption. We also work with preferences that allow for the possibility that schooling raises also employment quality (a non-wage benefit of schooling). Using recursive methods, we solve for the value functions in closed-form and set an exact likelihood for the number of years of education obtained by a given individual.

2.1 Theoretical Structure

In this sub-section, we present the theoretical structure of our empirical model. Each individual maximize his expected discounted lifetime utility, according to discount factor $\beta = \frac{1}{1+\rho}$ (where ρ is the subjective discount rate), by choosing the optimal time to enter the labour market. The control variable, d_t, is defined as follows

$d_t = 0$ if an individual, having completed S_t years of schooling, enters the labour market at period t.

$d_t = 1$ if an individual, having completed S_t years of schooling, decides to continue the schooling acquisition process.

- **Assumption 1- The Preferences**

Individual preferences are defined over a vector, Z_t, where $Z_t = (Y_{mt}, Y_{nmt}, e_t)$. The vector Z_t is composed of Y_{mt}, which represents monetary income, Y_{nmt}, which represents non-monetary income and the employment rate, e_t. The per-period (instantaneous) utility function, $U(Z_t)$, is assumed to be given by

$$U(Z_t) = \log(Z_t) = \log\left[Y_{mt}(S_t, d_t) \cdot Y_{nmt}(S_t, d_t) \cdot e_t(S_t, d_t)\right] \tag{1}$$

where S_t is schooling acquired by the beginning of date t. The dependence of Y_{mt} on S_t represents the monetary return to schooling while the dependence of Y_{nmt} and e_t on S_t can be interpreted respectively as the non-monetary return and the employment security return to schooling. It may be justified if, for instance, individuals obtain higher social status in return of being more educated or obtain better working conditions. The dependence of both Y_{mt} and Y_{nmt} on the control variable, d_t, is illustrated as follows;

$$Y_{mt} = \xi_t \; when \; d_t = 1 \; and \; Y_{mt} = w_t \; when \; d_t = 0$$

$$Y_{nmt} = 0 \; when \; d_t = 1 \; and \; Y_{nmt} = w_t^* \; when \; d_t = 0$$

Altogether, these equations illustrate that money income can be earned from working in the labor market (w_t) or from parental transfers (ξ_t) when in school and that non-monetary income (w_t^*) is earned in the labor market . The employment rate, e_t, represents the fraction of the year where income (either monetary or non-monetary income) is received. It is assumed to be equal to 1 when in school.

- **Assumption 2- Schooling Accumulation and the Schooling Interruption Process**

In order to be realistic, we assume that the schooling accumulation process can be interrupted due to exogenous reasons such as health problems, travel, incarceration or others. The interruption process is modeled as an exogenous transition, possibly correlated with individual characteristics. We assume that the schooling accumulation process is interrupted with probability ζ. Once the individual is in the interruption state, the per-period probability of returning to school is equal to ϱ. We assume that individuals know both probabilities. We define a binary variable, I_t, indicating whether the individual, who has not entered the labor market definitively (for whom d_t has not yet been set to 0), is currently in the school interruption state ($I_t = 1$) or is not ($I_t = 0$). It follows that

$$when \; I_t = 1 \; and \; d_t = 1 \; then \; S_{t+1} = S_t$$

$$when \; I_t = 0 \; and \; d_t = 1 \; then \; S_{t+1} = S_t + 1$$

Accumulated schooling at date t, S_t, can then be represented by

$$S_t = \sum_{s=0}^{t} d_t \cdot (1 - I_s)$$

The exogenous transition probabilities are defined follows;

$$\Pr(I_{t+1} = 1 \mid I_t = 0, d_t = 1) = \zeta \; and \; \Pr(I_{t+1} = 0 \mid I_t = 0, d_t = 1) = 1 - \zeta \tag{2}$$

$$\Pr(I_{t+1} = 0 \mid I_t = 1, d_t = 1) = \varrho \; and \; \Pr(I_{t+1} = 1 \mid I_t = 1, d_t = 1) = 1 - \varrho \tag{3}$$

- **Assumption 3- The Utility of Attending School**

The net income while at school, ξ_t, is assumed to be stochastic and reflects the difference between parental transfers and costs such as tuition, books, transportation as well as other psychological costs associated to the disutility of learning. The NLSY does not contain data on parental transfers and, in particular, does not allow a distinction in income received according to the interruption status. As a consequence, we ignore the distinction between income support at school and income support when school is interrupted.[4] The regression function for ξ_t is given by

$$\log \xi_t = E \log \xi_t + \varepsilon_t^{\xi}$$

with

$$\varepsilon_t^{\xi} \sim i.i.d \; N(0, \sigma_{\xi}^2)$$

When an individual leaves school, ξ_t is set to 0.

- **Assumption 4- The Utility of Working**

As the unit of time is a year, we ignore the distinction between the incidence and the duration of unemployment and model employment as the fraction of the year spent employed. This fraction, e_t, is assumed to be a non-stationary stochastic process. This means that, in any year t, the utility of working for a money wage, w_t and non-monetary wage, w_t^*, for a fraction of the year e_t, is simply given by

$$U(e_t, w_t, w_t^*) = \log\left[e_t(S_t) \cdot w_t(S_t) \cdot w_t^*(S_t)\right]$$

We assume that, at the beginning of each period, individuals observe a realization of w_t and decide to enter the labor market or not based on the expected employment rate. In other words, the actual employment rate, e_t, for a given year, is unknown at the beginning of each year.

[4] In the NLSY, we find that more than 82% of the sample has never experienced school interruption.

- **Assumption 5- The Wage and Non-Wage Return to Human Capital**

We assume that the labor market wage, w_t, is stochastic. The log wage regression equation is

$$w_t = \exp(\varphi_0^m + \varphi_s^m(S_t\ ) + \varphi_e^m(Exp_t) + \varepsilon_t^w) \tag{4}$$

where ε_t^w represents a sequence of i.i.d. stochastic shocks and where $\varphi_s^m(.)$ represents the return to education and $\varphi_e^m(.)$ represents the return to experience. At the estimation step, $\varphi_e^m(.)$ is chosen to be quadratic while $\varphi_s^m(.)$ is approximated by spline functions, in which case the marginal (local) return may change with the level of schooling.

The non-wage component, w_t^*, is assumed to be non-stochastic and is given by

$$w_t^* = \exp(\varphi_0^{nm} + \varphi_s^{nm}(S_t\ )) \tag{5}$$

where $\varphi_s^{nm}(.\ )$ represents the non-wage return to schooling.

- **Assumption 6- The Stochastic Process for Employment Security**

To characterize the stochastic process of the employment security variable, e_t, we start from the log inverse employment rate, e_t^*, that is

$$e_t^* = \log(\frac{1}{e_t}) \tag{6}$$

and assume that

$$\log(e_t^*) \sim N(\mu_t, \sigma_e^2) \tag{7}$$

It follows that

$$E\log e_t = -\exp(\mu_t + \frac{1}{2}\sigma_e^2)$$

and

$$Var(\log e_t) = \exp(2\mu_t + \sigma_e^2)\cdot(\exp(\sigma_e^2) - 1)$$

2.2 The Solution

Given assumptions 1 to 6 and a certain level of schooling, S_t, the expected utility of working at wage $\tilde{\mathrm{w}}_t$ and non wage benefit w_t^* is simply

$$EU(Z_t) = E\log\left[e_t(S_t)\cdot w_t(S_t)\cdot w_t^*(S_t))\right]$$

As it is done often in dynamic optimization problems, the solution to the stochastic dynamic problem can be characterized using recursive methods.

Assuming that the problem is solved over a finite horizon of length T, the expected value of entering period T is simply given by

$$EV_T = -\exp(\mu_T + \frac{1}{2}\sigma_e^2) + \varphi_0^m + \varphi_s^m(S_T) + \varphi_e^m(Exp_T) + \varphi_0^{nm} + \varphi_s^{nm}(S_T)$$

which can be used to solve the value functions associated to each possible strategies. The decision to remain in school with $S(t)$ years of schooling already accumulated, $V_t^s(S_t)$, can be expressed as

$$V_t^s(S_t) = \log(\xi_t) + \beta\{\zeta \cdot EV_{t+1}^I(S_{t+1}) + (1-\zeta) \cdot EMax[V_{t+1}^s(S_{t+1}), V_{t+1}^w(S_{t+1})]\}$$

or, more compactly, as

$$V_t^s(S_t) = \log(\xi_t) + \beta E(V_{t+1} \mid d_t = 1) \tag{8}$$

where $V_t^I(S_t)$ denotes the value of interrupting schooling acquisition and 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).

$V_{t+1}^I(S_t)$ can be expressed as follows;

$$\begin{aligned} V_t^I(S_t) &= \log(\xi_t) + \beta\{(1-\varrho) \cdot EV_{t+1}^I(S_{t+1} = S_t) + \\ \varrho \cdot EMax[V_{t+1}^s(S_{t+1} &= S_t), V_{t+1}^w(S_{t+1} = S_t)]\} \end{aligned}$$

where the equality between S_{t+1} and S_t follows from Assumption 2.

The value of stopping schooling (that is entering the labor market) with S_t years of schooling at wage w_t and taking into account the distribution of e_t (because e_t is unknown when w_t is drawn),$V_t^w(S_t)$, is given by

$$V_t^w(S_t) = -\exp(\mu_t + \frac{1}{2}\sigma_e^2) + \log(w_t + \lambda \cdot w_t^*) + \beta E(V_{t+1} \mid d_t = 0)$$

or, more precisely, as

$$V_t^w(S_t) = -\exp(\mu_t + \frac{1}{2}\sigma_e^2) + \varphi_0^m + \varphi_s^m(S_t) + \varphi_e^m(Exp_t) +$$
$$\varphi_0^{nm} + \varphi_s^{nm}(S_t) + \varepsilon_t^w + \beta E(V_{t+1} \mid d_t = 0)$$

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

$$E(V_{t+1} \mid d_t = 0) = \sum_{j=t+1}^{T} \beta^{j-(t+1)}(-\exp(\mu_j + \frac{1}{2}\sigma_e^2) +$$

$$\varphi_0^m + \varphi_s^m(S_j) + \varphi_e^m(Exp_j) + \varphi_0^{nm} + \varphi_s^{nm}(S_j))$$

2.3 Econometric Specification

In order to implement the model empirically, we must make some additional assumptions. First, we only model the decision to acquire schooling beyond 6 years (as virtually every individual has completed at least six years of schooling). Second, we set T (the finite horizon) to 65 years. Given that the model allows 4 distinct states (schooling acquisition, schooling interruption, work and unemployment), we ignore occupation decisions as well as university majors.[5]

- **Assumption 7- Household Characteristics and the Utility of Attending School**

We specify a reduced-form function for ξ_{it} (the net earnings while at school). The function is allowed to depend on household human capital variables as well as individual unobserved and observed ability and is represented by the following equation

$$\log \xi_{it} = X_i'\delta + \varphi(S_{it}) + \tilde{v}_i^{\xi} + \varepsilon_{it}^{\xi} \tag{9}$$

The vector X_i contains the following variables; parents' education (both mother and father), household income, number of siblings, family composition at age 14 and regional controls[6]. 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 lived with both their 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. Yearly family income is measured in \$1,000. The term $\tilde{v}_i^{\xi}$ represents individual heterogeneity (ability) affecting the utility of attending school. It is discussed in more details below. The marginal effect of schooling level on parental transfers, $\psi(S_t)$, is modeled using spline functions. In order to allow the earnings while at school (and, therefore, the utility of attending school) to depend on the level of schooling in a flexible fashion, we use 8 splines at the following grade levels (7 to 9, 10, 11, 12 , 13 to 14, 15 , 16 and 17-more).

- **Assumption 8- The Return to Human Capital**

The log wage received by individual i, at time t, is given by

$$\log w_{it} = \varphi_0^m + \varphi_1^m(S_{it}) + \varphi_2^m.Exp_{it} + \varphi_3^m.Exp_{it}^2 + \tilde{v}_i^w + \varepsilon_{it}^w \tag{10}$$

where $\varphi_1(S_t)$ is estimated with spline techniques (using 9 segments) and where

[5] Occupation choices are also typically ignored in the reduced-form literature. For an empirical model of occupation decisions, see Keane and Wolpin, 1997.

[6] In the NLSY, household income is measured as of 1979.

$$\varepsilon^w_{it} \sim i.i.d\ N(0,\sigma^2_w)$$

represents a purely random innovation to wages paid in the labor market. The term $\tilde{v}^w_i$ plays the role of labor market ability.

The employment rate, e_{it}, is also allowed to depend on accumulated human capital (S_{it} and Exp_{it}) and AFQT scores so that

$$E(\log e^*_{it}) = \mu_{it} = \kappa_0 + \kappa_1 \cdot S_{it} + \kappa_2 \cdot Exp_{it}\ + \kappa_3 \cdot AFQT_i$$

where κ_0 is an intercept term, κ_1 represents the employment security return to schooling, κ_2 represents the employment security return to experience and κ_3 is the effect of observed ability on employment rates. The probability of experiencing unemployment is assumed to be independent from the stochastic shock affecting wages .

- **Assumption 9- Observed and Unobserved Heterogeneity**

We assume that individual ability has a measured component (measured by AFQT scores) and an unmeasured component. Each individual is endowed with a pair ($\tilde{v}^w_i$,$\tilde{v}^\xi_i$) such that

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

where v^w_i and v^ξ_i represent the unmeasured component of individual ability. The stochastic specification is chosen such that we can allow unobserved labor market ability (v^w) to be correlated simultaneously with ability at school (v^ξ) and subjective discount rates (ρ). To do so, we assume that v^w_i, v^ξ and ρ_i are jointly distributed, with CDF H(v^w_i, v^ξ_i ρ_i), which can be approximated by a tri-variate discrete distribution. The distribution function H(.) is constructed such that each marginal distribution contains 2 points of support. The probabilities are expressed as follows:

$$Pr(v^\xi = v^\xi_1, v^w = v^w_1, \rho = \rho_1) = p_1$$

$$Pr(v^\xi = v^\xi_2, v^w = v^w_1, \rho = \rho_1) = p_2$$

$$Pr(v^\xi = v^\xi_1, v^w = v^w_2, \rho = \rho_1) = p_3$$

$$Pr(v^\xi = v^\xi_2, v^w = v^w_2, \rho = \rho_1) = p_4$$

$$Pr(v^\xi = v^\xi_1, v^w = v^w_1, \rho = \rho_2) = p_5$$

$$Pr(v^\xi = v^\xi_2, v^w = v^w_1, \rho = \rho_2) = p_6$$

$$Pr(\upsilon^{\xi} = \upsilon_1^{\xi}, \upsilon^{w} = \upsilon_2^{w}, \rho = \rho_2) = p_7$$

$$Pr(\upsilon^{\xi} = \upsilon_2^{\xi}, \upsilon^{w} = \upsilon_2^{w}, \rho = \rho_2) = p_8$$

where $\upsilon_1^w > \upsilon_2^w$, $\upsilon_1^\xi > \upsilon_2^\xi$ and $\rho_1 > \rho_2$ and with the restriction that

$$\sum_{j=1}^{J} p_j = 1 \; and \; p_j \geq 0 \; j = 1, 2...J(where J = 8) \tag{11}$$

Estimating H(.) flexibly therefore requires the estimation of 6 support points and 7 free probability parameters. The probabilities are estimated using logistic transforms, that is

$$p_i = \frac{\exp(q_i)}{\sum_{j=1}^{8} \exp(q_j)}$$

for i=1,2,..,8 and normalize q_8 to 0. Altogether, and depending on the model specification, the empirical implementation of the model requires the estimation of up to 55 parameters.

2.4 The Likelihood Function

Constructing the likelihood function is relatively straightforward. To do so, we use the definitions of d_t and I_t, along with the assumption that both the interruption and the return probabilities are exogenous. This allows us to represent all transition probabilities needed to write down the likelihood function. These transition probabilities can be split into a set of exogenous probabilities, which characterizes the process of school interruption, and endogenous probabilities, characterizing the decision to leave school permanently or to continue in school. Altogether, they represent all possible destinations.

The transition probabilities defining the choice between interrupting school permanently (start working) and obtain an additional year of schooling, are given by,

$$\Pr(d_{t+1} = 0 \mid I_t = 0, d_t = 1) = (1 - \zeta) \cdot \Pr(V_t^w(S_t) \geq V_t^s(S_t)) \tag{12}$$

$$\Pr(d_{t+1} = 1, I_{t+1} = 0 \mid I_t = 0, d_t = 1) = (1 - \zeta) \cdot \Pr(V_t^w(S_t) < V_t^s(S_t)) \tag{13}$$

$$\Pr(d_{t+1} = 1, I_{t+1} = 1 \mid I_t = 0, d_t = 1) = \zeta \tag{14}$$

$$\Pr(d_{t+1} = 0 \mid I_t = 1, d_t = 1) = \varrho \cdot \Pr(V_t^w(S_t) \geq V_t^s(S_t)) \tag{15}$$

$$\Pr(d_{t+1} = 1, I_{t+1} = 0 \mid I_t = 1, d_t = 1) = \varrho \cdot \Pr(V_t^w(S_t) < V_t^s(S_t)) \tag{16}$$

$$\Pr(d_{t+1}=1, I_{t+1}=1 \mid I_t=1, d_t=1) = 1-\varrho \tag{17}$$

where

$$\Pr(V_t^w(S_t) \geq V_t^s(S_t)) = \Pr(\varepsilon_t^w - \varepsilon_t^\xi \geq M(.))$$

and where

$$\begin{aligned} M(.) &= E\log(\xi_t) + \beta E(V_{t+1} \mid d_t = 1) + \exp(\mu_t + \frac{1}{2}\sigma_e^2) - \\ &\quad [\varphi_0^m + \varphi_s^m(S_t\ ) + \varphi_e^m(Exp_t) + \lambda(\cdot\varphi_0^{nm} + \varphi_s^{nm}(S_t\ ))] \end{aligned}$$

Both (14) and (17) represent a probability of leaving school permanently in $t+1$ (depending on I_t) while equations (15) and (18) represent a probability of staying in school and acquire an additional year of human capital (depending on I_t). Equations (16) and (19) represent the exogenous probability of being attracted to the interruption state, depending on whether or not the individual was already in the interruption state. The likelihood function is constructed from data on schooling attainments as well as the allocation of time between years spent in school ($I_t = 0, d_t = 1$) and years during which school was interrupted ($I_{t+1} = 1, d_t = 1$) and on employment histories (wage/unemployment) observed when schooling acquisition is terminated (until 1990). The construction of the likelihood function requires to evaluate the following probabilities;

- the probability of having spent at most τ years in school (including years of interruption), $Pr[(d_{i,0} = 1, I_0), (d_{i,1} = 1, I_1)....(d_{i,\tau} = 1, I_\tau)]$
- the probability of entering the labor market, in year $\tau+1$, at observed wage $w_{i,\tau+1}$, $P(d_{i,\tau+1} = 0, w_{i,\tau+1})$
- the density of observed wages and employment rates from $\tau+2$ until 1990, $Pr\ (\{\tilde{w}_{i,\tau+2}\}..\{\tilde{w}_{i,1990}\})$

The probability of stopping school after τ years, is given by

$$\Pr[(d_{i,0} = 1, I_0), (d_{i,1} = 1, I_1)....(d_{i,\tau} = 1, I_\tau)] = \left\{\prod_{s=0}^{\tau} P(d_{i,s} = 1, I_s)\right\} \tag{18}$$

where individual contributions, $P(d_{i,s} = 1, I_s)$, will vary according to whether or not the individual is experiencing an interruption or not in period s and on the state occupied in the previous period (I_{s-1}).

The probability of entering the labor market at wage $w_{i,\tau+1}$, $P(d_{i,\tau+1} = 0, w_{i,\tau+1})$, can be factored as

$$\Pr(d_{i,\tau+1} = 0, w_{i,\tau+1}) = \Pr(d_{i,\tau+1} = 0 \mid w_{i,\tau+1}) \cdot \Pr(w_{i,\tau+1}) \tag{19}$$

where $P(d_{i,\tau+1} = 0 \mid w_{i,\tau+1})$ is a normal conditional probability. Finally, using the fact that the random shocks affecting the employment process and the wage process are mutually independent and are both i.i.d., the contribution to the likelihood for labor market histories observed from $\tau_i + 2$ until 1990 is given by[7]

$$Pr\left(\{\tilde{w}_{i,\tau+2}\}..\{\tilde{w}_{i,1990}\}\right) = Pr\left(\{w_{i,\tau+2} \cdot e_{i,\tau+2}\} \cdot .... \Pr\{w_{i,1990} \cdot e_{i,\tau+2}\}\right) = \prod_{s=\tau+2}^{1990} \Pr(\tilde{w}_{i,s}) \quad (20)$$

and it can easily be evaluated using the following distributional assumptions;

$$\Pr(\log w_{i,s}) = \frac{1}{\sigma_w}\phi\left(\frac{\log w_{i,s} - (\varphi_0^m + \varphi_s^m(S_{i,s}\ ) + \varphi_e^m(Exp_{i,s}) + v_i^w)}{\sigma_w}\right)$$

$$\Pr(\log e^*_{i,s}) = \frac{1}{\sigma_e}\phi\left(\frac{\log e^*_{i,s} - (\kappa_0 + \kappa_1 \cdot S_{is} + \kappa_2 \cdot Exp_{is})}{\sigma_e}\right)$$

Using (20), (21) and (22) the likelihood function, for a given individual (conditional on unobserved heterogeneity), is given by

$$L_i(. \mid \upsilon^{\xi}, \upsilon^{w}, \rho) = \left\{\prod_{s=0}^{\tau_i} P(d_{i,s} = 1, I_s)\right\} \cdot P(d_{i,\tau+1} = 0, w_{i,\tau+1}) \cdot \prod_{\tau_i+2}^{1990} \Pr(\tilde{w}_{i,s}) \quad (21)$$

Maximizing the log likelihood function simply requires to write the individual contributions (23) conditional on every possible combinations $\vartheta_j = \left(\upsilon^{\xi}, \upsilon^{w}, \rho\right)_j$ and taking a weighted average of all contributions according to the $p's$. The unconditional contribution to the log likelihood, for individual i, is therefore given by

$$\log L_i = \log \sum_{j=1}^{J} p_j \cdot L_i(. \mid \vartheta_j) \quad (22)$$

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 20 or less as of January 1, 1979. We record information on education, wages and on employment rates for each individual from the time the individual is age 16 up to December 31, 1990.[8]

The original sample contained 3,790 white males. However, we lacked information on family background variables (such as family income as of

[7] For those who are still in school at the survey time or those for whom wages are missing, the contribution to the likelihood is easily adjusted.

[8] The reason for not including information beyond 1990 is that the wage data do not appear reliable for these more recent waves.

1978 and parents' education) and on AFQT score for 1,161.[9] The age limit and missing information regarding actual work experience further reduced the sample to 1,710.

Before discussing descriptive statistics, it is important to describe the construction of some important variables. In particular, both the schooling attainment variable and the experience variable deserve some discussions. First, 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. It also helps us to identify those individuals who have interrupted schooling. Overall, young individuals tend to acquire education without interruption. In our sample, only 306 individuals have experienced at least one interruption (Table A2). This represents only 18% of our sample and it is along the lines of results reported in Keane and Wolpin (1997). As well, we note that interruptions tend to be short. Almost half of the individuals (45 %) who experienced an interruption, returned to school within one year while 73% returned within 3 years.

Second, unlike many studies set in a reduced-form which use potential experience (age -education- 5), we use data on actual experience. Actual experience accumulated by period t (exp_t) is constructed as follows

$$\exp_t = \sum_{s=\tau_{i+1}}^{t} e_{is} \ for \ t \geq \tau_{i+1}$$

where e_{it} denotes the fraction of the year worked by a given individual. The availability of data on actual employment rates allows use to estimate the employment security return to schooling.

Descriptive statistics for the sample used in the estimation can be found in Table A1 (in Appendix). The frequencies for various schooling attainments and completions are in Table A2 (in Appendix). 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 12.8 years.[10] The low incidence of interruptions explains the low average number of interruptions per individual (0.22) and the very low average interruption duration (0.43 year).

4 Empirical Results

In Section 4.1, we discuss the main structural parameter estimates. In Section 4.2, we investigate the relative importance of household human capital and subjective discount rates in explaining individual variations in schooling attainments. In Section 4.3, we look at the relative importance of household human capital and subjective discount rates on lifetime wages. In Section

[9] We lost about 17% of the sample due to missing information regarding family income and about about 6% due to missing information regarding parents' education.

[10] The average schooling attainment at age 16 is between 10 and 11 years.

4.4, we evaluate the true intergenerational education correlation and its macroeconomic counterpart; the impact of human capital based economic growth on schooling attainments of the next generation.

4.1 Parameter estimates

In this section, we present structural estimates obtained from 2 different model specifications. In the first one (Model 1), ability has only an unmeasured component, assumed to be represented by a multi-variate discrete distribution. It is therefore orthogonal to parents' human capital by construction. In the second specification (Model 2), ability has a measured component (AFQT scores) as well as an unmeasured component. It is therefore a more general specification and it can capture the correlation between individual ability and parents' human capital. To facilitate presentation of the results, we split the estimates in 2 tables. The effects of household human capital and school ability on the utility of attending school, the correlations between both types of ability and schooling attainments and the distribution of subjective discount rates are in Table 1. The effects of education, ability and experience on wages, employment rates and non-wage benefit are in Table 2. Finally, the ability of the model to fit data on schooling attainments is illustrated in Table 3.

The parameter estimates for the effects of family human capital variables on the utility of attending school, found in Table 1, indicate clearly that, other things equal, those raised by parents' who have more schooling and higher income, those who were living with both parents at age 14 (nuclear family indicator) and those raised in families with smaller number of children (siblings) tend to have a higher utility of attending school. This is true in Model 1 as well as in Model 2. However, a comparison between those estimates obtained without AFQT (Model 1) and those with AFQT (Model 2) indicates that introducing AFQT scores typically reduces the effect of parents' human capital. The estimates for father's education (going from 0.0316 to 0.0226) and mother's education (going from 0.0240 to 0.0136) have been substantially lowered. This is also true about the nuclear family indicator (going from 0.0734 to 0.0385). The only exception is the effect of family income which has increased slightly from 0.0022 to 0.0026. The correlations computed in Table 1 indicate a clear positive correlation between school ability and market ability (0.26 in Model 1 and 0.40 in Model 2) and a high correlation between market ability and schooling attainments (0.24) in Model 2. As discussed in Belzil and Hansen (2000), this is a clear indication of an ability bias. The main characteristics of the distribution of subjective discount rates are also found in Table 1. Like other parameters, subjective discount rates are affected by the presence of AFQT scores. In Model 1, the high discount rate (0.1635) and the low discount rate (0.1043) translate into an average rate of 0.1307 in the population. In Model 2, the high (0.1132) and low (0.0755) discount rates average to a value of 0.0919. Introducing test scores has also reduced the standard deviation of individual discount rates (0.0294 in Model 1 against 0.0187 in Model 2). Finally, both in Model 1 and Model 2, the interruption probability is found to be 0.0770 (significant at the 1% level).

The estimates found in Table 2 summarize the return to human capital under various forms. First, the negative estimates for the effects of schooling (-0.1347 in Model 1 and -0.0999 in Model 2) and experience (-0.0098 in Model 1 and -0.0990 in Model 2) on the mean log inverse employment rates indicate that both education and experience increase employment rates (reduce unemployment).[11] Note that the effect of schooling on employment is reduced significantly by the introduction of AFQT scores and that AFQT scores are also found to increase employment rates. The non-wage benefit to schooling is estimated to be 0.0204 in Model 1 and 0.0208 in Model 2.

However, the most notable results are at the level of the wage equation. The estimates of the wage equation reveal that assuming constant marginal returns to schooling is a serious mistake. The high level of significance of the parameter estimates for the spline functions (see the note below Table 1) indicate that a model with constant marginal (local) returns would be strongly rejected. Overall, the estimates indicate that the local returns to schooling are small at low levels and increase significantly in grade 12 (10% per year in Model 1 and 7.7% in Model 2), in grade 14 (15% in Model 1 and 10% in Model 2) and in grade 16 (13% in Model 1 and 10% in Model 2). As pointed out by Belzil and Hansen (2000), the relatively higher estimates obtained in a specification such as Model 1 are explained by the incapacity of a model, where ability is orthogonal to parents background, to capture a significant correlation between labor market ability and schooling attainments. Indeed, Model 2 achieves a much better likelihood value than Model 1. We note, however, that both in Model 1 and Model 2, the average return to schooling over the entire range (grade 7 to grade 18) is much below the returns to schooling reported in the literature. This is due to the mis-specification of the function representing the local returns to schooling.[12]

As estimating a structural dynamic programming model requires stronger assumptions than those needed in a reduced-form model, it is important to investigate the capacity of our model to fit data on schooling attainments. The predicted schooling attainment frequencies are found in Table 3. Both models are able to predict schooling attainments quite accurately, although the model that incorporates AFQT scores (Model 2) fits the data best and also achieves a higher value of the likelihood. For both model specifications reported in Table 3, a simple χ^2 test statistic (with 4 degrees of freedom) fails to reject the null hypothesis that the model is properly specified at the 1% level. The critical value (at the 1% level) is 13.3.

[11] Belzil and Hansen (2000) show that the log normal specification of the log inverse employment rate fits the data quite well.

[12] Belzil and Hansen (2000) show that restricting the local returns to be constant will result in an overstatement of the average return of the order of 0.02 and that this overstatement is present whether or not the ability bias is controlled for. For more detailed discussions, see Belzil and Hansen (2000).

Table 1
The Effects of Household Human Capital on the Utility of Attending School

	Model 1 Without AFQT Param (p value)	Model 2 With AFQT Param (p value)
Utility in School		
intercept v_1^{ξ}	0.2478 (0.05)	0.0871 (0.10)
Intercept v_2^{ξ}	-0.5901 (0.05)	-0.9324 (0.01)
Father's Educ	0.0316 (0.01)	0.0226 (0.03)
Mother's Educ	0.0240 (0.03)	0.0136 (0.04)
Family Income/1000	0.0022 (0.02)	0.0026 (0.03)
Nuclear Family	0.0734 (0.03)	0.0385 (0.04)
Siblings	-0.0273 (0.03)	-0.0166 (0.04)
AFQT Scores	-	0.1150 (0.00)
Interruption Prob.	0.0770 (0.01)	0.0770 (0.01)
Discount Rates		
ρ_1	0.1635 (0.01)	0.1132 (0.01)
ρ_2	0.1043 (0.01)	0.0755 (0.01)
Proportion (ρ_1)	0.4452 (0.02)	0.4348 (0.01)
Proportion (ρ_2)	0.5548 (0.01)	0.5652 (0.01)
Correlations		
$Corr(\tilde{v}_i^{\xi}, \tilde{v}_i^{w})$	0.2609 (0.01)	0.4038 (0.01)
$Corr(\tilde{v}_i^{\xi}, S_{i,90})$	0.0226 (0.35)	0.3250 (0.01)
$Corr(\tilde{v}_i^{w}, S_{i,90})$	0.0090 (0.71)	0.2392 (0.01)
Mean Log likelihood	-14.2343	-14.0317

Note: P Values are in parentheses. In Model 1, the average discount rate is 0.1307 and the standard deviation is 0.0294. In Model 2, the average discount rate is 0.0919 and the standard deviation is 0.0187. Discount rates are estimated as exponential transforms and standard errors are computed using the delta method.

Table 2
The Return to Human Capital

		Model 1 without AFQT Param (P value)	Model 2 with AFQT Param (P value)
Non-wage benefit			
	Schooling	0.0204 (0.02)	0.0208 (0.01)
Employment			
	Schooling	-0.1347 (0.01)	-0.0999 (0.01)
	experience	-0.0098 (0.01)	-0.0090 (0.02)
	AFQT	-	-0.0462 (0.03)
Wages			
	int1υ_1^w	1.7200 (0.01)	1.6815 (0.01)
	int2υ_1^w	1.2623 (0.01)	1.2328 (0.01)
	Experience	0.0963 (0.01)	0.0949 (0.01)
	experience^2	-0.0040 (0.03)	-0.0090 (0.02)
	AFQT	-	0.0288 (0.01)
	Schooling		
	7-10 years	0.0147	0.0145
	11 years	0.0314	0.0180
	12 years	0.1042	0.0770
	13 years	0.0449	0.0036
	14 years	0.1515	0.1011
	15 years	0.0505	0.0363
	16 years	0.1284	0.1029
	17 years	0.1016	0.0728
	18-more	0.0706	0.0625
	Average Return	0.0618	0.0443

Note: P Values are in parentheses. The local returns are computed from the spline parameter estimates. In Model 1, the spline estimates are 0.0147 (7-10), 0.0167 (11), 0.0728 (12), -0.0593 (13), 0.1066 (14), -0.1010 (15), 0.0779 (16), -0.0268 (17) and 0.0310 (18-more). They are all significant at 5%. In Model 2, the corresponding spline estimates are 0.0145, 0.0035, 0.0590, -0.0734, 0.0985, -0.0648, 0.0666, -0.0301 and -0.0103. Except for the 18-more segment, they are all significant at 5%.

Table 3
Model Fit: Actual vs Predicted Schooling Attainments

	Pred. (%)	Pred. (%)	Actual (%)
	Model 1	Model 2	-
Schooling:			
6-10 years	0.143	0.146	0.146
11-12 years	0.465	0.459	0.472
13-14 years	0.142	0.144	0.147
15-16 years	0.167	0.169	0.158
17 or more	0.083	0.082	0.078
$\chi^2_{4d.f.}$	2.949	1.579	

Note: The critical value, at a 1% level, is 13.3

4.2 The Relative Importance of Discount Rates, Parents' Human Capital and Ability in Explaining Individual Schooling Attainments.

One of the most striking advantages of estimating a structural dynamic programming model, is the possibility to identify a fundamental preference parameter such as the subjective discount rate. At this stage, a question naturally arises. What is the relative importance of subjective discount rates and household human capital in explaining individual schooling attainments? To answer this question, we have evaluated the variations in predicted values explained by various sources (Table 4) and computed elasticities of mean schooling attainments with respect to the same variables (Table 5). The variations in predicted schooling attainments, explained by a particular variable or parameter, are computed as follows: $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(PRED(S_i) - E(PRED(S))^2}$ where $PRED(S_i)$ is the predicted schooling attainment of individual i obtained when we fix all attributes other than the one of interest at their respective sample average (or mode) and where $E(PRED(S)$ denotes the sample average of the predicted schooling attainments (when every variable and parameter are allowed to vary). The relative dispersion observed for various sources (such as parents' human capital or discount factor heterogeneity) will provide a good indication of the relative importance of each factor.

The measures of dispersion are presented in Table 4. In Model 1, the total variation imputed to all characteristics (1.4460), to parents' education and income (1.3626) and parents' education only (1.0986) are all much below the variation imputed to discount rate heterogeneity (2.3356). For instance, the measure of variation in individual schooling attainments imputed to heterogeneity in discount rates is around 80% larger than the variation imputed to all family attributes (parents, education and income, nuclear family status and siblings). Only school ability, with a measure of variation equal to

2.2790, appears to be more important. Although the measure of variations imputed to discount rate heterogeneity has dropped when we introduced AFQT scores (in Model 2), its value (0.6164) is around 50% higher than the variability imputed to all family characteristics (0.4461) and the variability imputed to parents' education and income (0.4111). Again, with AFQT scores, the most important source of variation is school ability.

When the relative importance of various factors are evaluated in terms of elasticities (computed in Table 5), the ranking is identical. The elasticities of mean schooling attainments with respect to potential discount rates (-0.3631 in Model 1 and -0.1404 in Model 2) are always much larger (in absolute value) than elasticities with respect to measurable characteristics such as education and income. No surprisingly, among observable characteristics, the highest elasticity is with respect to father's education (0.1065 in Model 1) and (0.0270 in Model 2). The elasticities with respect to mother's education (0.0788 and 0.0180) also tend to exceed the income elasticities. These results are consistent with those reported by Cameron and Heckman (1998).

4.3 The Relative Importance of Discount Rates, Parents' Human Capital and Ability in Explaining Lifetime Wages

As seen in Section 4.2, individual unobserved heterogeneity in the rate of time preference is found to account for a larger fraction of the variation in schooling attainments than do family human capital variables. As a consequence, individual variations in discount rates will also have an impact on lifetime wages. To illustrate this, we have performed an exercise similar to the one presented in the first 2 columns of Table 4 and computed a measure of variation in predicted wages that can be imputed to family background variables, discount factor heterogeneity and ability. Without loss of generality, we used predicted entry wages. The results are in column 3 and 4 of Table 4. As expected, the relative share of variations in explained wages due to individual discount rates and family background variables are in agreement with what we have seen when we analyzed variations in predicted schooling attainments. While ability in the market is the prime factors explaining predicted wages, heterogeneity in the rate of time preference accounts for a larger variation in predicted wages than other family characteristics. In Model 1, the variation in predicted wages (0.0934) is also around 80% higher than the variation explained by all family background variables (0.0578). In Model 2, the variation explained by the rate of time preference (0.0245) is approximately 50% higher than the variation explained by family background variables (0.0178). Finally, and not surprisingly, it should be noted that, in both Model 1 and Model 2, labor market ability accounts for the largest share of the variation in explained wages.

To illustrate the effect of discount factor heterogeneity on lifetime wages (log), we also computed predicted wages at various points in time (as well as the sum of all lifetime wages), for individuals who differ only in terms of their rate of time preference (Table 6). The homogeneity with respect to all other variables and parameters implies that all individuals receive the same amount of parental transfers. As can be seen from the predicted wages in Table 6, despite higher entry wages, those endowed with lower discount rates will not always realize higher lifetime wages. For instance, in Model 1, those with the lower discount rate (0.10) will give up 3 years of potential experience when compared to those with the high discount rate (0.16). The relatively low returns (local) to schooling will allow individuals endowed with a higher discount rate (type 1) to achieve higher lifetime wages. In Model 2, the difference in discount rates (0.07 for type 2 and 0.11 for type 1) translates to a difference of 1 year in schooling attainment (14 years for type 2 against 13 years for type 1). The relatively high return to completion of grade 14 (around 10%) is sufficient to ensure that the lifetime earnings of those who have completed grade 14 exceeds the lifetime earnings of those who stopped in grade 13.[13]

[13]As far as we know, in the literature devoted to income distribution, differences in the rate of time preference are rarely advanced as a potential factor of income/wage inequality.

Table 4
Sources of Variations in Schooling Attainments

	Model 1	Model 2	Model 1	Model 2
AFQT scores	**No**	**Yes**	**No**	**yes**
	Variations in pred. schooling		**Variations in pred. wages**	
Variable/Parameter				
Preference				
Discount Rates	2.3356	0.6164	0.0934	0.0245
Family background				
Parents' education	1.0986	0.2429	0.0439	0.0097
Parents'educ and income	1.3626	0.4111	0.0545	0.0164
All Family variables	1.4460	0.4461	0.0578	0.0178
Ability				
School Ability	2.2790	2.9125	0.0909	0.1165
Market Ability	0.1857	0.5441	0.2474	0.2455

Note: variations in predicted schooling attainments are computed as follows:

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(PRED(S_i) - E(PRED(S))^2}$$

where $PRED(S_i)$ is computed at the sample average of all variables or parameters other than the one of interest and where $E(PRED(S)$ denotes the sample average of the predicted schooling attainments. Variations in predicted wages (log) are computed similarly:

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(PRED(\log w_i) - E(PRED(\log w))^2}$$

$PRED(\log w_i)$ is computed at the sample average of all variables other than the one of interest and $E(PRED(\log w))$ denotes the sample average of all predicted wages.

Table 5
Elasticities of Schooling Attainments with Respect to Discount Rates,
Parents' Education and Family Income

	Model 1 **No AFQT**	**Model 2** **With AFQT**
$\frac{\%\Delta E(Educ)}{\%\Delta discountrate}$	-0.3631	-0.1404
$\frac{\%\Delta E(Educ)}{\%\Delta father's.educ}$	0.1065	0.0270
$\frac{\%\Delta E(Educ)}{\%\Delta mother's.educ}$	0.0788	0.0180
$\frac{\%\Delta E(Educ)}{\%\Delta family\ income}$	0.0196	0.0085

Note: Elasticities are computed at the sample averages of all other variables and parameters.

Table 6
Lifetime Wages and the Rate of Time Preference

	Model 1 **No AFQT**		**Model 2** **With AFQT**	
	Predicted Wages			
	Type 1	**Type 2**	**Type 1**	**Type 2**
	$\rho_1 = 0.16$	$\rho_2 = 0.10$	$\rho_1 = 0.11$	$\rho_2 = 0.08$
Year				
1984	1.70	-	1.72	-
1985	1.79	-	1.82	1.83
1986	1.88	-	1.90	1.92
1987	1.97	2.03	1.99	2.01
1990	2.25	2.31	2.27	2.28
1995	2.71	2.77	2.73	2.74
2000	3.18	3.23	3.18	3.19
2010	4.10	4.15	4.09	4.10
2020	5.02	5.07	5.00	5.01
2030	5.94	6.00	5.91	5.93
Lifetime Wages	179.61	176.59	179.51	180.11

Note: Predicted wages are computed for individuals who are endowed with same household characteristics (set at the sample average) and same level of ability. Differences in wages are therefore explained by differences in schooling attributable only to differences in discount rates.

4.4 What is the True Intergenerational Education Correlation ?

The sensitivity of the effects of parents' human capital on schooling attainments to the inclusion of a measure of ability such as AFQT suggests the possibility that the intergenerational education correlation measured in the data is spurious. It may reflect the fact that parents with high ability tend to have children who also have high ability. If so, the true intergenerational education correlation may differ substantially from the correlation found in the data. To investigate the issue, we used the structural parameters of Model 1 and Model 2 to generate data on schooling attainments, letting vary all observable attributes or parameters that are subject to heterogeneity. Then, we computed the partial correlations between predicted schooling attainments and parents' education and income and compared those with the partial correlations between actual schooling attainments and parents' education and income.

The partial correlations measured in the data are found in the last column of Table 7. The partial correlation between actual schooling and father's schooling is around 0.28, while the partial correlation with mother's education is 0.23 and the partial correlation with respect to family income is 0.1654. In the model where ability has only an unobserved component (Model 1), the partial correlation between mean schooling and father's schooling, computed from simulated schooling attainments, is found to be 0.288 and is very close to the partial correlation found in the data (0.278). For mother's education, the partial correlation emerging from Model 1 (0.231) is stronger than the partial correlation measured in the data (0.16). However, when allowing for a measure of ability correlated with family background (in Model 2), both correlations are found to be much weaker. The partial correlation between son and father's education (0.226) is 25% lower than the correlation found in the data. For mother's education, the drop is even more spectacular; the correlation dropping from 0.23 to 0.03. To a large extent, this result is a reflection of the stronger correlation between AFQT scores and mother's education than father's education. Finally, unlike for parents' education, introducing AFQT scores has strengthened the effects of family income. The correlation goes from 0.167 (in Model 1) to 0.195 (in Model 2) and exceeds the partial correlation found in the data (0.165). Overall, our results indicate that the true intergenerational education correlation, predicted from our structural dynamic programming model, is much lower than what is found in the data.

As argued before, the intergenerational education correlation, inferred from cross-section data, has a macroeconomic counterpart. Precisely, knowledge of the true intergenerational education correlation can be used to measure the average increase in the level of schooling of the next generation explained by an exogenous increase in the level of human capital of the current generation. While the effect of human capital and education on growth is one of the central questions addressed by empirical macroeconomists, it is also important to investigate how household human capital affects schooling attainments of the next generation. Although, following Lucas (1988), more general theoretical models have been introduced, which involve overlapping generations and human capital transfers across generations, very few of them have been tested empirically.[14] In our model, the intergenerational human capital transmission mechanism is relatively simple. An increase in education increases household income of the current generation. This increase in schooling and income will, in turn, increase the utility of attending school and reduce the opportunity cost of schooling for the next generation. In other words, education may be seen as a by-product of economic growth. To investigate this issue, we have performed the following experiment. We increased exogenously the level of schooling of both the father and mother by one year and imputed a 5% increase in family income (a relatively credible average return to schooling given our estimates). Then, we computed the average increase in schooling attainments of the next generation. This has been done for both model specifications. The results are found in Table 8. On average, increasing both parents' education by one year will raise schooling attainment of the next generation by 0.1 (Model 2) to 0.3 (Model 1) year. This increase may be seen as relatively small. However, there exist no well-defined benchmark result which is equivalent in the macroeconomic literature devoted to economic growth. As a consequence, our result cannot be evaluated easily. Although this increase appears relatively modest, it nevertheless illustrates some persisting effects of human capital growth and indicate that an increase in the average level of education is, at the same time, one of the consequences as well as one of the causes of economic growth.

[14] See Barro and Sala-i-Martin, 1995, for a survey.

Table 7
The Partial Correlations between Individual Schooling Attainments and Family Human Capital

	Model 1	Model 2	OLS
AFQT	**No**	**Yes**	**No**
	Partial Correlations	**Partial Correlations**	**Partial Correlations**
Father's educ	0.2878 (0.01)	0.2256 (0.01)	0.2786 (0.01)
Mother's educ	0.2314(0.02)	0.0327 (0.24)	0.1603 (0.01)
Family Income	0.1673 (0.02)	0.1954 (0.01)	0.1654 (0.03)

Note: P Values are in parentheses.

Table 8
Human Capital, Growth and Intergenerational Transfers: The effect of an Exogenous Increase in Parents' Education on Schooling Attainments of the Next Generation

	Model 1 Without AFQT	**Model 2 With AFQT**
Δ EDUCATION OF NEXT GENERATION	0.3 year	0.1 year

Note: Simulations are computed at the sample averages of all other variables and parameters.

5 Conclusion

We have estimated a structural dynamic programming model of schooling decisions where individual heterogeneity (observed as well as unobserved) has several dimensions; ability in school, ability in the labor market, household human capital and subjective discount rates. The econometric specification of the model is quite general. Parents' human capital, observed market ability and observed school ability are allowed to be correlated while, at the same time, unobserved ability in school, unobserved ability in the market and subjective discount rates are drawn from a multi-variate discrete distribution. The structure of the model has allowed us to investigate the relative importance of the rate of time preference and household human capital in explaining cross sectional differences in schooling attainments and wages. It also enabled us to investigate the link between the intergenerational education correlation (typically measured from cross-section data) and the effect of human capital based economic growth on schooling attainments of the next generation (the aggregate counterpart to the intergenerational education correlation).

Overall, we find that differences in the individual rate of time preference are more important than differences in household human capital in explaining individual differences in schooling attainments and wages. In the model specification where ability is made of an unobserved component only, the variations explained by discount factor heterogeneity are around 80% higher than variations explained by family background variables. With measured ability, the variations due to discount factor heterogeneity are 50% higher than the variations explained by household human capital variables. However, whether or not AFQT scores are included in the analysis, ability in school remains the principal factor explaining individual variations in schooling attainment. When the relative importance of both factors is evaluated by elaticities, mean schooling attainments are found to be 3 to 5 times more elastic (in absolute value) with respect to individual discount rates than with respect to father's education (the family background variable which has the highest elasticity). While ability in the labor market appears to be the prime factor explaining predicted wages, the variation in predicted wages due to heterogeneity in the rate of time preference also exceeds the variation due to differences in family background variables. Modelling schooling decisions in a framework where agents are forward looking is therefore fundamental.

Not surprisingly, we find that the intergenerational education correlation predicted by the dynamic programming model is sensitive to the inclusion of AFQT scores. In the model with test scores (the model specification that reaches the highest likelihood value and fits the data best), the true partial correlation between son and father's schooling (0.22) is around 25% lower than the correlation found in the data. The difference between the true partial correlation between son and mother's education (0.03) and the correlation disclosed by the data (0.23) is even larger. Simulations indicated that an exogenous increase of 1 year in the level of schooling of the current generation will increase schooling attainments of the next generation by 0.1 to 0.3 year. While there is no benchmark result in the growth literature based on cross-country growth rate comparisons, our findings illustrate the link between the intergenerational education correlation (typically found in cross-section data) and the intergenerational effects of human capital growth at the aggregate level. Our findings also illustrate the need for economic models in which education is a consequence as well as a cause of economic growth.

References

[1] Barro, Robert and Sala-i-Martin (1995) *Economic Growth*, New-York: McGraw Hill.

[2] Belzil, Christian and Hansen, Jörgen (2000) "Subjective Discount Rates, Unobserved Ability and the Return to Schooling" IZA Working paper, University of Bonn.

[3] 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.

[4] Eckstein, Zvi and Kenneth Wolpin (1999) "Youth Employment and Academic Performance in High School" IZA Discussion Paper no. 18, forthcoming in *Econometrica*

[5] 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.

[6] Keane, Michael P. and Wolpin, Kenneth (1997) "The Career Decisions of Young Men" *Journal of Political Economy, 105 (3)*, 473-522.

[7] Lazear, Edward (1980) "Family Background and Optimal Schooling Decisions" *Review of Economics and Statistics, 62 (1)*, 42-51.

[8] Lucas, Robert (1988) "On the Mechanics of Economic Development" *Journal of Monetary Economics 22:3-42*

[9] Rust, John (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Analysis of Harold Zurcher" *Econometrica, 55 (5)*, 999-1033.

[10] Topel, Robert (1999) "Labor Markets and Economic Growth", Working Paper, University of Chicago.