A Dynamic Structural Model of Labor Supply and Educational Attainment: the long term effects of study patterns

Wayne-Roy Gayle*
Natalia Khorunzhina†

March 7, 2014

Abstract

This paper develops and estimates a dynamic structural model of labor supply and schooling to investigate how students choose the amount of working, studying and leisure activities while being enrolled in school, and how these choices in the beginning of their carriers affect the accumulation of human capital over the life cycle. The theoretical model provides a detailed treatment of the economic costs, benefits and uncertainties associated with the schooling and labor supply alternatives faced by individuals, including the uncertainty over future returns to education. We find substantial heterogeneity in the returns to education over races and types of education. Our framework also allows for investigating the degree of human capital depreciation (and appreciation) across groups of individuals. We find the rates of knowledge appreciation and depreciation over races and different levels of education are non-trivial. We argue that the hump-shaped life cycle wage profiles may be due to knowledge depreciation rather than decreasing returns to experience.
KEYWORDS: Educational Outcome, Labor Supply, Preference Nonseparability, Dynamic Models.
JEL: I21, J22, J24, C23, C25

*Department of Economics, University of Virginia, Monroe Hall, McCormick Rd, Room 208A, Charlottesville, VA 22903, E-mail: wg4b@virginia.edu. All errors are our own.

†Department of Economics, Copenhagen Business School, Porcelænshaven 16 A, 1, 2000 Frederiksberg, Denmark, E-mail: nk.eco@cbs.dk.

1 Introduction

Over the last three decades, there has been an increasing trend of young people working while being actively enrolled in school. Young individuals are increasing their incidence of work, and the amount of hours worked while enrolled in school.[1] This trend has generated growing interest in the possible immediate and long run effects of working while enrolled in school on educational attainment and future labor market opportunities. On one hand, there is a concern that an intensive amount of working while in school may hinder academic performance and increase drop-out rates, thus jeopardizing future career opportunities. Evidence to this effect, known as the crowd-out hypothesis, is found in D'Amico (1984), Ehrenberg and Sherman (1987), Keane and Wolpin (1997), Eckstein and Wolpin (1999). On the other hand, working while in school may improve a young individual's time organizational skills, sense of responsibility and self esteem, which in turn are traits that may be rewarded in the labor market in the future. Evidence to this effect, referred to as the congruence hypothesis, is found in D'Amico (1984), Ruhm (1997), Oettinger (1999), Stinebrickner and Stinebrickner (2003, 2004).[2] It is not obvious, which of these two opposing effects dominate. Also, the net effect of these opposing forces is likely to vary among young people. However, there seems to be a consensus that both hypotheses are correct in that moderate hours of working while in school improve academic performance and reduces drop-out rates, while an intense amount of hours worked is significantly detrimental to academic performance and educational attainment (see Kalenkoski and Pabilonia, 2010).

Policy makers have responded to the increased incidence and intensity of working while in school by implementing policies aimed at increasing graduation rates and reducing the time taken to graduate. These policy interventions include tuition subsidies and student loans. The evidence of the effectiveness of these policy interventions are mixed. For example, Keane (2002) find that tuition subsidies have modest effects on college drop-out rates and time to completion, while Bound and Turner (2002), Keane and Wolpin (1997), and Eckstein and Wolpin (1999) find significant effects. The effectiveness of these pol-

[1]See Bound et al. (2007), Bacolod and Hotz (2005), and Bound and Turner (2002) for recent documentation of this phenomena.

[2]Additionally, working while in school produces immediate work experience, helps to build a resume, provides networking opportunities and allows for early career experiments. Students may earn cash, which may be used to finance their studies, and acquire valuable budgeting skills.

icy interventions depend on understanding who selects to work moderate versus intense hours while in school. Students self select into the categories of working conditional on observable and unobservable characteristics. Therefore, to understand how these policy interventions affect the distribution of academic performance and drop-out rates, one must understand the process of how different types of students choose the amount of hours to work while in school. Keane and Wolpin (1997) shows the increase in college graduation due to tuition subsidies varies considerably over unobserved types of students. However, the hours worked while in school does not fully capture the time allocation problem students face. Given the choice to work during school, students also face the joint decision of how to allocate time between hours working, hours studying, and leisure activities. These choices are made simultaneously and the choice of time allocation varies over individuals of different characteristics. For example, students with high utility for leisure who obtain tuition subsidies may choose to reduce hours worked, but not increase hours spent studying. In this case, tuition subsidies are not expected to have significant effects on academic performance and college completion rates. Therefore, modeling the simultaneous choice of working and studying is necessary to understanding the effect of tuition subsidies across students on academic attainment, drop-out rates, and time to completion.

Young individuals make education and work decisions based not only on the immediate returns to education and experience, but also based on their expectations of how these choices affect their future employment and income prospects. A key ingredient in the formation of these expectations in the returns they expect from education. In recent years, researchers have recognized the inadequacy of the classical Mincer wage equation to quantify the rate of returns to schooling. Heckman et al. (2008) cite notable extensions to the classical Mincer wage equation that are likely to reduce the biases in estimates of returns to education. Of particular interest is heterogeneity in returns to education. Belzil and Hansen (2002) estimate a structural model of heterogenous returns to schooling, where local returns are estimated using eight spline segments. They strongly reject the hypothesis of constant returns to schooling. They find the returns to education to be lower that what is typically reported in the literature and that it increases smoothly with the level of education. Heckman et al. (2008) also find significant heterogeneity in the returns to schooling. However, Heckman et al. (2008) find larger returns to education that what is typically found in the literature, and that returns to graduating from high school is larger than the returns to graduating from college.

We model returns to education as a first order, discrete support Markov process, where the allow for the distribution to be dependent on observable characteristics, including race and degree level. Therefore, we allow for future returns to education to be unknown to the individual, while the individual knows its distribution conditional on his individual characteristics. This framework allows for addressing issues of heterogeneity in the expected returns to education across groups of individuals. We find that the returns to education for blacks are slightly higher than those for whites (16% compared to 15%), while Hispanics obtain significantly lower returns to education than their black and white counterparts (5%). Similar to the results in Heckman et al. (2008), we find the returns to education are higher for high school graduates than for college graduates across all racial groups, with the average returns for high school graduates being 15% compared to 7% for college graduates. Finally, we find that black and whites with some high school education obtain similar returns to education, whereas Hispanics obtain twice the returns to some high school education compared to high school graduates.

Furthermore, our framework allows for investigating the degree of human capital depreciation (and appreciation) across groups of individuals. We find that, compared to whites, blacks are less likely to experience both knowledge appreciation and depreciation with the knowledge depreciation being less likely relative to the knowledge appreciation. Hispanics are more likely to experience both knowledge appreciation and depreciation, however, the effect of knowledge depreciation is larger relative to the effect of knowledge appreciation. These patterns are consistent with our results of average returns to education for blacks being larger than for whites, and lower for Hispanics.

The rest of the paper is organized as follows. In the next section, we present the model. Section (3) discusses the construction of the sample used in estimation, and Section (**??**) discusses the empirical methodology implemented in estimation of the parameters of interest. Section (5) discusses the empirical findings. Section (**??**) concludes.

2 Model

This section develops the theoretical framework in which we investigate how individuals allocate time between human capital accumulation, labor market participation, and leisure.

The model is set in discrete time. In each period, $t \in \{1, \cdots, T\}$, an generic individual, $i \in \{1, \cdots, n\}$, is endowed with a fixed amount of time, which is normalized to one. He must choose how to allocate this unit of time between leisure, l_{it}, labor market activities, h_{it}, and school activities, s_{it}:

$$l_{it} + h_{it} + s_{it} = 1. \tag{2.1}$$

Define I_{it}^{h} to be equal to one if individual i chooses to work in period t, and zero otherwise. Similarly, define I_{it}^{s} to be equal to one if he decides to enroll in school in period t, and zero otherwise. Therefore, the individual faces four mutually exclusive and exhaustive discrete alternatives, j, in each period: neither work nor attend school ($j = 1$), not work and attend school ($j = 2$), work and not attend school ($j = 3$) and, work and attend school ($j = 4$). Define $d_{itj} = 1$ if alternative j is chosen and zero otherwise. Given alternatives $j = 1, \cdots, 4$ is chosen, the faces the following conditional continuous choices: $\bar{s}_{it1} = \bar{h}_{it1} = 0$ when $j = 1$; $\bar{s}_{it2} > 0$, $\bar{h}_{it2} = 0$ when $j = 2$; $\bar{s}_{it3} = 0$, $\bar{h}_{it3} > 0$ when $j = 3$; and $\bar{s}_{it4} > 0$, $\bar{h}_{it4} > 0$ when $j = 4$. Define $d_{it} = (d_{it1}, \cdots, d_{it4})$, $\bar{s}_{it} = (\bar{s}_{it1}, \cdots, \bar{s}_{it4})$, and $\bar{h}_{it} = (\bar{h}_{it1}, \cdots, \bar{h}_{it4})$. The unconditional continuous choices are: $s_{it} = \bar{s}_{it}' d_{it}$, and $h_{it} = \bar{h}_{it}' d_{it}$. Define $y_{it} = (s_{it}, h_{it})$.

Associated with the individual is an D_z-dimensional vector of observable state variables, z_{it}. If the individual chooses to work in period t, he receives wages $w_{it} = w_t(z_{it}, \mu_{it}, r_{2it})$ and earns an additional year of experience. Here, r_{2it} is the shock to wage levels, and μ_{it} is the shock to the returns to education. If he decides to attend school in period t, he advances a grade level with probability $F_{jt}(z_{it}, y_{it})$, $j = 2, 4$. If the individual enrolls in post- high school, he also pays tuition fees and receive tuition subsidy (we do not model the choice to apply for college tuition).

We do not model savings. Therefore, the individual's choice-specific level of consumption, c_{itj}, is a deterministic function of his state variables, and choices. For $j =, \cdots, J$, let $\varepsilon_{itj}, j = 1, \cdots, J$ be the vector of discrete–choice-specific shock to utility. Let r_{1it} be the shock to the utility of leisure. Define $\varepsilon_{it} = (\varepsilon_{it1}, \cdots, \varepsilon_{it4})$, $r_{it} = (r_{it1}, r_{it2})$, $v_{it} = (\mu_{it}, r_{it})$, and $x_{it} = (z_{it}, v_{it})$. The individual's conditional direct current period utility from choosing alternative j in period t is denoted by $u_{tj}(z_{it}, y_{it}, e_{it}) + \varepsilon_{itj} = u_{tj}(z_{it}, c_{itj}, s_{itj}, h_{itj}, r_{1it}) + \varepsilon_{itj}$.

The individual chooses $\{(d_{it}, y_{it}), t = 1, \cdots, T\}$ to sequentially maximize the expected discounted sum of payoffs

$$E\left\{\sum_{t=1}^{T}\sum_{j=1}^{4}\beta^{t-1}d_{itj}[u_{tj}(x_{it}, y_{it}) + \varepsilon_{itj}]\right\} \tag{2.2}$$

subject to the constraints

$$s_{itj} \geq 0, \tag{2.3}$$

$$h_{itj} \geq 0, \tag{2.4}$$

where $\beta \in (0,1)$is the time discount factor. Let the optimal decision rule in period $t = 1, \cdots, T$ be given by (d^0_{itj}, y^0_{itj}). Assume that ε_{it} is independent of x_{it} and ε_{itj} is independent of ε_{itk} for $k \neq j$. Let the ex-ante value function in period t, $V_t(z_{it})$, be the discounted sum of expected future utilities, before ε_{it} is revealed, given the optimal decision rule:

$$V_t(z_{it}, e_{it}) = E\left\{\sum_{\tau=t}^{T}\sum_{j=1}^{4}\beta^{\tau-t}d^0_{i\tau j}(x_{i\tau}, \varepsilon_{i\tau})[u_{\tau j}(x_{i\tau}, y^0_{i\tau}) + \varepsilon_{i\tau j}]\right\}. \tag{2.5}$$

Then the Bellman equation is given by

$$\begin{aligned} V_t(x_{it}) &= E\left\{\sum_{j=1}^{4} d^0_{itj}(x_{it}, \varepsilon_{it})\left[u_{tj}(x_{it}, y^0_{it}(x)) + \varepsilon_{itj} + \beta E[V_{t+1}(x_{it+1})|x_{it}]\right]\right\} \\ &= E\left\{\sum_{j=1}^{4} d^0_{itj}(x_{it}, \varepsilon_{it})\left[v_{tj}(x_{it}, y^0_{it}(x_{it})) + \varepsilon_{itj}\right]\right\}, \end{aligned} \tag{2.6}$$

where

$$v_{tj}(x_{it}, y^0_{it}(x_{it})) = u_{tj}(x_{it}, y^0_{it}(x_{it})) + \beta\int V_{t+1}(x_{it+1}) f_{jt}(x_{it+1}|x_{it}, y^0_{it}(x_{it}))dx_{it} \tag{2.7}$$

is the choice-specific conditional value function without ε_{itj}, and $f_{jt}(x_{it+1}|x_{it}, y_{it})$ is the probability density function of x_{it+1} given alternative j is chosen in period t. Then the optimal conditional study and work hours satisfy

$$\frac{\partial}{\partial s_{it2}} v_{t2}(x_{it}, y_{it}^0(x_{it})) = 0,$$
$$\frac{\partial}{\partial h_{it3}} v_{t3}(x_{it}, y_{it}^0(x_{it})) = 0,$$
$$\frac{\partial}{\partial s_{it4}} v_{t4}(x_{it}, y_{it}^0(x_{it})) = 0,$$
$$\frac{\partial}{\partial h_{it4}} v_{t4}(x_{it}, y_{it}^0(x_{it})) = 0.$$

Given the individual's optimal continuous study and work hours, his discrete choice of alternative j is optimal if

$$d_{tj}^0(x_{it}, \varepsilon_{it}) = \begin{cases} 1 & \text{if} \quad v_{tj}(x_{it}, y_{it}^0(x_{it})) + \varepsilon_{itj} > v_{tk}(x_{it}, y_{it}^0(x_{it})) + \varepsilon_{itk} \quad \forall k \neq j \\ 0 & \text{otherwise} \end{cases} \tag{2.8}$$

2.1 Alternative representation

In this section, we present the conditional choice probability representation of the alternative-specific conditional value functions developed in Hotz and Miller (1993), Altug and Miller (1998), Arcidiacono and Miller (2011), and Gayle (2013). To explain the method, define the choice-specific conditional choice probability (CCP) of alternative j by

$$p_{tj}(x_{it}) = \int d_{tj}^0(x_{it}, \varepsilon_{it}) g_{\varepsilon} \varepsilon_{it} d\varepsilon_{it}, \tag{2.9}$$

where g_{ε} is the density of ε_{it}. Arcidiacono and Miller (2011) extend the results of Hotz and Miller (1993) and Altug and Miller (1998) to show the period t alternative-specific conditional value function of choice j, which are specified in terms of the period t alternative-specific utility function and the conditional continuation value function given alternative j is chosen in period t, may be equivalently expressed in terms of the period t alternative-specific utility function plus the flow utilities and CCPs for any sequence of future choices. With this generalization of the alternative representation of the alternative-specific conditional value functions, they extend Altug and Miller's definition of finite dependence to the situation where two discrete choice sequences lead to the same distribution of states after a few periods. The consequence of this definition is that if finite dependence holds, then the difference in these alternative-specific value functions can be expressed in terms of the difference in the two period t alternative-specific utility functions and the flow utilities and CCPs for any sequence of a few future choices, thus reducing the representation of the current period CCPs.

While Arcidiacono and Miller's extension of the definition of finite dependence is substantive, the issue of finding the sequence of choices associated associated with two initial choices remain difficult in practice and is largely case-specific. For example, finite dependence as defined by Arcidiacono and Miller is easily satisfied if the sequence of choices associated with two initial choices lead to an absorbing state after a few periods. However, if the transition probabilities are irreducible, then finite dependence as defined by Arcidiacono and Miller is not satisfied in general.

To overcome this difficulty, Gayle (2013) extend the definition of finite dependence in Arcidiacono and Miller (2011) in a way in which finite dependence is obtained after only one period for any transition probability. Gayle (2013) achieves this property by noting that alternative-specific value functions can be expressed in terms of the period t alternative-specific utility function plus any period-specific weighted sum of the flow utilities and CCPs, so long as the weights sum to one. Gayle (2013) then shows that the weights may always be chosen so that the difference in the alternative-specific value functions associated with two alternatives can be expressed in terms of the difference in the two period t alternative-specific utility functions and the period $t+1$ utilities and CCPs. It is this definition of finite dependence we implement in this paper.

We impose the following restriction on the choice-specific transition probability

$$f_{jt}(x_{it+1}|x_{it}, y_{itj}) = f_{jt}(z_{it+1}|z_{it}, y_{itj})\pi_t(\mu_{it+1}|z_{it}, \mu_{it})g_r(r_{it+1}),$$

where $\pi_t(\mu_{it+1}|z_{it}, \mu_{it})$ is the transition density of shocks to the returns to education, and g_r is the density of r_{it}. Assume that ε_{itj} is distributed i.i.d., type 1 logit. Then Gayle (2013) shows that for $j = 2,3,4$,

$$
\begin{aligned}
& v_{tj}(x_{it}, y_{it}^0(x_{it})) - v_{t1}(x_{it}, y_{it}^0(x_{it})) = u_{tj}(x_{it}, y_{it}^0(x_{it})) - u_{t1}(x_{it}, y_{it}^0(x_{it})) \\
& + \beta \int\int\int \Big([u_{t+1,1}(x_{it+1}, y_{it+1}^0(x_{it+1})) - \ln(p_{t+1,1}(x_{it+1})] a_{j1}(z_{it+1}, y_{it+1}^0) \\
& + [u_{t+1,j}(x_{it+1}, y_{it+1}^0(x_{it+1})) - \ln(p_{t+1,j}(x_{it+1}))] a_{1j}(z_{it+1}, y_{it+1}^0) \Big) g_r(r_{it+1}) \pi_t(\mu_{it+1}|z_{it}, \mu_{it}) \\
& \times [f_{jt}(z_{it+1}|z_{it}, y_{it}^0) - f_{1t}(z_{it+1}|z_{it}, y_{it}^0)] dr_{it+1} d\mu_{it+1} dz_{it+1}, \qquad (2.10)
\end{aligned}
$$

where

$$
\begin{aligned}
& a_{j1}(z_{it+1}) = \\
& \int \frac{f_{jt+2}(z_{it+2}|z_{it+1}, y_{it+1}^0(x_{it+1}))}{f_{1t+2}(z_{it+2}|z_{it+1}, y_{it+1}^0(x_{it+1}) + f_{jt+2}(z_{it+2}|z_{it+1}, y_{it+1}^0(x_{it+1}))} g_z(z_{it+2}) dz_{it+2}, \qquad (2.11)
\end{aligned}
$$

and $a_{1j}(z_{it+1}) = 1 - a_{j1}(z_{it+1})$, where $g_z(z_{it+2})$ is some density function of z_{it+2}. This alternative representation of the difference in the conditional choice-specific value function admits a convenient way to compute the choice specific conditional continuous choices. To see this, consider the optimal study hours given that the individual chooses alternative 2. Because $\partial v_{t1}(x_{it}, y_{it}^0(x_{it}))/\partial s_{it2} = 0$ identically, then $s_{it2}^0(x_{it})$, the optimal conditional study hours associated with alternative 2 given the state vector x_{it} solves

$$
\frac{\partial}{\partial s_{it2}} [v_{t2}(x_{it}, y_{it}^0(x_{it})) - v_{t1}(x_{it}, y_{it}^0(x_{it}))] = 0, \qquad (2.12)
$$

where the LHS of equation (2.12) is given in equation (2.10). The optimal conditional study hours associated with alternative 2 given the observable state vector z_{it} is given by

$$
s_{it2}^0(z_{it}) = \int\int s_{it2}^0(z_{it}, \mu_{it}, r_{it}) g_r(r_{it}) \pi_t(\mu_{it}|z_{it}) dr_{it} d\mu_{it}.
$$

The other optimal conditional choice are obtained similarly. The CCP of alternative $j = 1, \cdots, 4$ given the observable state vector x_{it} is given by

$$
p_{tj}(x_{it}) = \frac{e^{v_{tj}(x_{it}, y_{it}^0(x_{it})) - v_{t1}(x_{it}, y_{it}^0(x_{it}))}}{1 + \sum_{k=2}^4 e^{v_{tk}(x_{it}, y_{it}^0(x_{it})) - v_{t1}(x_{it}, y_{it}^0(x_{it}))}},
$$

And the CCP of alternative $j = 1, \cdots, 4$ given the observable state vector z_{it} is given by

$$p_{tj}(z_{it}) = \int \int p_{itj}(z_{it}, \mu_{it}, r_{it}) g_r(r_{it}) \pi_t(\mu_{it}|z_{it}) dr_{it} d\mu_{it}.$$

2.2 Parametrization of the model

Wage offer

The market wage the individual i of type q receives in period t if he decides to work is assumed to take the form

$$\ln w(z_{it}\mu_q, r_{2it}) = \sum_{k=1}^{7} \tilde{t}^{k-1}\delta_k^w + \sum_{k=8}^{13} z_{itk}^w \delta_k^w + \mu_q G_{it} + r_{2it}, \tag{2.13}$$

Where $\tilde{t} = t/T$, $z_{it}^w = (BLACK, HISPANIC, AFQT, E_{it}, E_{it}^2)$, where E_{it} is the level of labor market experience the individual achieves at time t and G_{it} is the highest grade level achieved by the individual at time t. We assume that observed wages are measured with errors that is independent of the (t, z_{it}^w, G_{it}).

Returns to education

We assume that the joint distribution of (μ_{it+1}, μ_{it}) given z_{it}, $\Pi(\mu_{it+1}, \mu_{it}|z_{it})$ belongs to a family of finite mixture models.

$$\Pi(\mu_{it+1}, \mu_{it}|z_{it}) = (\{\mu_{q'}, \mu_q, \pi q' q(z_{it})\}, q, q' = 1, \cdots, Q).$$

We assume that the number of types of support points of the distribution of returns to education is two, and that given by

$$\pi_{q'q}(z_{it}) = \frac{e^{\tilde{\theta}^{\mu}_{q'q} + \sum_{o=1}^{4} z_{ito}^{\mu} \theta_{qo}^{\mu}}}{\sum_{k'=1}^{2} \sum_{k=1}^{2} e^{\tilde{\theta}^{\mu}_{k'k} + \sum_{o=1}^{4} z_{ito}^{\mu} \theta_{ko}^{\mu}}}. \tag{2.14}$$

where $z_{ito}^{\mu} = (BLACK, HISPANIC, HSG_{it}, COLG_{it})$, where HSG_{it} is equal to one if individual i is a high school graduate, but not a college graduate in period t, and zero otherwise. $COLG_{it}$ is equal to one if the individual is a college graduate in time t. We normalize $\tilde{\theta}_{44}^w = \theta_{21}^w = \cdots = \theta_{24}^{\mu} = 0$. We impose parametric restrictions on the distribution of types because, although there has been significant advances in dynamic structural estimating models with unobserved heterogeneity without function form restrictions on the distribution of types (see Kasahara and Shimotsu (2009) and Arcidiacono and Miller (2011)), identification and estimation the conditional distribution of types given observed characteristics without functional form restrictions remains an unresolved issue.

Consumption function

Individual i of type q in period t consumes the his earned income in period t if he works (we do not model asset accumulation). If he enrolls in college, he pays college tuition $(TUIT_{it})$ and receives tuition subsidy of $TSUB_{it}$ if he is (exogenously) granted tuition subsidy, which is denoted as $d_{it}^{TSUB} = 1\{TSUB_{it} > 0\}$, where $1\{\cdot\}$ is the indicator function equal to one if the argument inside is true, and zero otherwise. The individual also receives a subsistence level of consumption $\bar{c}$. Therefore

$$c(z_{it}, \mu_q, r_{it}) = I_{it}^h w_{itq} - I_{it}^s d_{it}^{HSG}(TUIT_{it} - d_{it}^{TSUB} TSUB_{it}). \tag{2.15}$$

As is standard, we impose a minimum (sustenance) level of consumption to ensure that consumption is always strictly positive.

Period-specific utility

Individual i of type q in period t has preferences over consumption, leisure, the attending school, and labor market participation, which is represented by the following period specific utility function.

$$\begin{aligned} u(z_{it}, c_{itjq}, l_{itj}, d_{it}^s, d_{it}^h) = {} & z_{it}^c \theta_1^c c_{itjq} + \theta_2^c c_{itjq}^2 + z_{it}^l \theta_1^l l_{itj} + \theta_2^l l_{litj}^2 + \theta^{cl} c_{itjq} l_{itj} \\ & + z_{it}^s \theta^s I_{it}^s + z_{it}^h \theta^h I_{it}^h + \theta^{sh} I_{it}^s I_{it}^h, \end{aligned} \tag{2.16}$$

where $z_{it}^c = (1, AGE_{it}, BLACK_i, HISPANIC_i)$, $z_{it}^l = (1, r_{1it}, AGE_{it}, BLACK_i, HISPANIC_i)$, $z_{it}^s = (1, d_{it-1}^s, AGE_{it}, AGE_{it}^2, BLACK_i, HISPANIC_i)$, and $z_{it}^h = (1, d_{it-1}^h, E_{it}, BLACK_i, HISPANIC_i)$. Notice that we allow for intra-temporal nonseparabilities in preferences between consumption and leisure, enrollment and labor force participation, as well as intertemporal nonseparabilities in enrollment and labor force participation.

Transition probability and structural shocks

Under the restrictions of the model, the predetermined variables (variables that are determined by period $t-1$'s action) are the level of experience G_{it}, E_{it}, and $TSUB_{it}$. We parameterize the grade-transition probability to take the logit form $F_{tj}(z_)$. Because the list of parameters is extensive, we refer to table —-.

Shocks
We assume that $g_r = N(0, \Sigma^r)$, where $\Sigma^r_{11} = \sigma_1^2$, $\Sigma^r_{22} = \sigma_2^2$, and $\Sigma_{12} = \sigma_{12}$.

3 Data

The data is taken from the 1979 youth cohort of the National Longitudinal Survey of Labor Market Experience (NLSY79), a comprehensive panel data set that follows individuals over the period 1979 to 2000, who were 14 to 21 years of age as of January 1, 1979. The data set initially consisted of 12,686 individuals: a representative sample of 6,111 individuals, a supplemental sample of 5,295 Hispanics, non-Hispanic blacks, and economically disadvantaged, non-black, non-Hispanics, and a supplemental sample of 1,280 military youth. Interviews were conducted on an annual basis though 1994, after which they adopted a biennial interview schedule. This study makes use of the first 16 years of interviews, from 1979 to 1994.[3] The data is restricted to include males and to exclude respondents with missing observations on the highest grade level completed that cannot be recovered with high confidence from other data information. A list and description of the variables used in the model is presented in Table 1. Table 2 presents summary statistics of the sample used in this study. Attrition accounts for a loss of approximately 22 percent of the individuals between 1979 and 1994. However, the largest loss occurred between 1990 and 1991, late in the sample period.

4 Estimator

There are two main potential approaches to estimate the parameters of the model; an iterative version of the moment-based approach of Altug and Miller (1998), the iterative likelihood approach of Aguirregabiria and Mira (2002), and the EM likelihood estimator approach of Arcidiacono and Miller (2011). We adopt the iterative version of the moment-

[3] Appendix 1 provides a detailed discussion of the data construction and sample restrictions.

based approach. The reason for our choice because this approach allows for measurement errors in observed wage, hours worked, and study time without imposing parametric restrictions on the distribution of measurement errors.

The moments we construct to estimate the parameters of the model are based on the following residuals.

$$
\begin{aligned}
& d^o_{itj} - p_{itj}(z_{it}), \\
& d^o_{it-1,j} d^o_{itj} - p_{itt-1,j}(z_{it}), \\
& I^h_{it}[w^o_{it} - \sum_{q=1}^{2} w(z_{it}, \mu_q, r_{2it})], \\
& I^s_{it}[s^o_{it} - d_{it2}\bar{s}^0_{it2} + d^{it4}\bar{s}_{it4}], \\
& I^h_{it}[h^o_{it} - d_{it3}\bar{h}^0_{it3} + d^{it4}\bar{h}^0_{it4}], \\
& I^s_{it}[d^{TR}_{it} - F_{2t}(z_{it})^{d_{it2}} F_{4t}(z_{it})^{d_{it4}}],
\end{aligned}
$$

where

$$
p_{itt-1,j}(z_{it}) = \int \left[\int p^0_{it+1,j}(z_{it+1}, s_{q'}) f_{jt}(x_{t+1}|x_t, s_q, r_t) dx_{t+1} \right] p^0_{tj}(x_t, s_q, r_t) g_r(r_t) dr_t. \tag{4.1}
$$

5 Estimation Results

We address several issues while discussing the results obtained from the estimation. The main conclusions from the results are that: (i) there is a strong evidence of intra-temporal complimentarity between the enrollment to school and work decisions; (ii) there are substantial differences in the returns to education as well as in the preferences over schooling and work for three major demographic groups of young males: white, black and hispanics; (iii) there are substantial differences in the patterns of knowledge appreciation and depreciation among three major demographic groups of young males.

Period-specific utility

Period specific utility is composed as a sum of the following components: the utility of consumption, leisure, the psychic value of school enrollment, and the fixed utility of labor market participation. Table 1 reports the estimates of the structural parameters governing the period-specific utility payoff. Column (1) reports components of the utility of consumption, column (2) provides the estimates of the components of the utility of leisure, column (3) gives the estimates that characterize the psychic value of school attendance, and column (4) reports the estimates of the fixed utility of labor force participation.

The estimation results of columns (1) and (2) indicate that period-specific utility is increasing and concave in both consumption and leisure. However, the parameter governing the concavity of utility in leisure is imprecisely estimated. Compared to White males, Black and Hispanic males exhibit lower preferences for consumption and higher preference for leisure. We also find evidence of substitutability of leisure and consumption choices in intratemporal preferences.

Columns (3) and (4) of Table 1 report the parameter estimates for the fixed utility of school enrollment and the utility payoff from labor force participation. The results in column (3) of Table 1 show that most of the coefficients of psychic value of school attendance are negative, which suggests that it is costly to acquire education not only in terms of monetary expenses but also in terms of time and effort spent doing the schoolwork. Several important conclusions emerge from these results. First, we find an evidence of a significant cost of being enrolled in school. In addition, the results suggest on strong intertemporal substitutability in preferences over the fixed cost of school enrollment. This result provides an evidence on a statistically significant psychic cost of a continued enrollment in school. Next, we find intertemporal complimentarity in preferences over labor force participation. This finding suggests that continued employment is preferred to breaks in labor supply. Further, we find significant intra-temporal complimentarity in preferences over school enrollment and labor force participation decisions. Observationally equivalent individuals (with the same optimal conditional work and study hours) find enrollment in school more valuable if they are participating in the labor market. Finally, according to the results reported in column (4) of Table 1, we find that the consumption value of labor force participation is decreasing in the level of labor market experience.

We find that the fixed utility of school enrollment and the utility payoff from labor force participation differ significantly over demographic characteristics of young males. Relative to their white counterparts, black males exhibit higher preferences over employment and enrollment to school, while Hispanic males exhibit lower preferences over these two economic decisions. This implies that after controlling for racial differences in wages, hours worked, time spent of schooling, and school quality, black males are more likely to enroll in school and be engaged in labor market participation, while Hispanic males are less likely to enroll and to be employed than their white counterparts. We find that the psychic payoff of school attendance decreases with age and enrollment to school becomes prohibitively costly for older males (the coefficients of age and age squared in column (3) are negative significant). Coupled with the finding on substantial psychic cost of a continued enrollment in school, these results capture the decreasing rate of enrollment in school for higher levels of education and older individuals.

Returns to education

In recent years, researchers have recognized the inadequacy of the classical Mincer wage equation to quantify the rate of returns to schooling. Heckman et al. (2008) cite notable extensions to the classical Mincer wage equation that are likely to reduce the biases in estimates of returns to education. These include heterogeneity in returns to education, direct and psychic cost of schooling, nonseparability between experience and schooling, loss of working life with additional years of schooling, and disentangling marginal and average returns to schooling.[4] Other important factors that may affect estimates of the returns to education include: the endogeneity of schooling and work experience choices, uncertainty about the returns to education, uncertainty about the completed level of education, and uncertainty receiving tuition subsidies. Our model incorporates all of these important factors.

Table 2 presents a summary of the distribution of the heterogenous returns to schooling. We estimate the returns to education as a first-order Markov process conditioned of race and education groups. Therefore we allow for observed and unobserved heterogeneity in returns to education, as well as conditional correlation in the returns to education over time. This extends the analysis of heterogeneity by Heckman et al. (2008), in that, in their framework, the returns to education are assumed to be known to each individual. Without a loss of generality, we assume that there are two types of individuals in the population: high and low. Next, we make one step forward to allow for transition between types over time.

[4] Heckman et al. (2008) also points out to the importance of income tax treatment as a factor to reduce the biases in estimates of returns to education.

Table 1: Period-Specific Utility

Utility of:	Consumption (1) CONSUMPTION ×	Leisure (2) LEISURE ×	School Attendance (3) ENROLLMENT ×	Labor Force Participation (4) EMPLOYMENT ×
CONSTANT	30.6582 (1.4348)	6.9123 (0.3535)	−2.7890 (0.8050)	−1.1557 (0.8484)
LEISURE	−2.7657 (0.7632)	−1.0129 (0.7336)		
CONSUMPTION	−0.7786 (0.0094)			
ENROLLMENT				6.1612 (0.8432)
LAGGED ENROLLMENT			−47.4873 (0.2991)	
LAGGED EMPLOYMENT				7.0367 (0.9474)
BLACK	−6.8667 (0.0457)	5.8782 (0.5600)	7.6191 (0.5420)	4.4343 (0.5420)
HISPANIC	−11.9096 (0.0207)	5.9113 (0.4411)	−3.8937 (1.0048)	−8.2882 (0.8287)
AGE	−13.5433 (0.6247)	−3.9506 (0.2699)	−2.3407 (0.8115)	
SQUARED AGE			−27.3080 (0.5317)	
WORK EXPERIENCE				−2.9118 (1.4961)
r_{1it}		−7.4104 (0.6518)		
σ^r_{12}		0.6555 (0.4846)		

One advantage of this specification is the ability to account for the dynamics in depreciation and appreciation of knowledge over time (for a general discussion, see McFadden, 2008).

Panel A of Table 2 reports the average returns to education by race and education groups. Our estimated average returns to education are significantly higher than those typically obtained using the classical Mincerian regression model. For example, using data from the NLSY79, Neal and Johnson (1996) report estimates of the returns to education around 6%, and Card (1999) reports estimates ranging between 7% and 8.5%. However, our results are comparable to the resent findings in returns to education literature, for example to the results reported in Heckman et al. (2008). Similar to Heckman et al. (2008), we find that, on average, the returns to education for blacks are slightly higher than those for whites (16% compared to 15%). Hispanics obtain significantly lower returns to education than their black and white counterparts (5%). Next, we find the returns to education are significantly higher for high school graduates than for college graduates across all racial groups, with the average returns for high school graduates being 15% compared to 7% for college graduates. Indeed, Heckman et al. (2008) also find that returns to for college graduates typically align with those estimated using the classical Mincerian regression model. In contrast to Heckman et al. (2008), we find that black and whites with some high school education obtain essentially the same returns to education, though slightly larger for whites. This pattern reverses for Hispanics, who obtain over twice the returns to some high school education compared to being high school graduates.

Our results on larger returns to high school completion relative to the returns to college completion are not uncommon due to recent findings in Heckman et al. (2008), nevertheless they are still striking. These findings are even more puzzling in light of the evidence on declining high school graduation rates documented in Heckman and LaFontaine (2010) and greater unemployment spells among high school graduates compared to college graduates (see Carnevale et al., 2011a). Whereas the source of heterogeneity in returns to education over races in the same education cohort could be various, one of the plausible explanation may be the racial differences in selecting into occupations that grant higher returns to education. Data presented in Carnevale et al. (2011b) are consistent with this story and show that, among blacks and hispanics, there is a larger share of blacks with college majors that result in higher earnings.

Panel B of Table 2 reports the returns to education for each type (high and low) as well as a matrix of unconditional types transition probabilities. The results imply that, whereas 13% of young males, which constitute low type, received 0.2% rate of return to education, 63% of young males, which constitute high type, receive 21% rate of return. Further, we find that whereas 3.5% of young males benefit from knowledge appreciation, by moving from low type to high type, 20% of them experience knowledge depreciation, signified by their moving from high type to low type. This pattern of transition implies that knowledge depreciation accounts for a significant portion of the curvature in wage profile, and not just the concavity of log wages in labor market experience.[5]

Panel C of Table 2 reports the marginal effects of race and education on the estimates of the type transition matrix. The omitted group is white males with less than average AFQT scores and some high school education. The marginal effects on the type transition matrix sheds light on the source of variation in the average returns to education discussed above. Whereas the probability of moving from low type to high type characterizes knowledge appreciation, the probability of moving from high type to low type represents knowledge depreciation. We report and discuss the direction of the marginal effects, where a positive estimate indicates on greater probability of appreciation or depreciation. We also discuss the relative importance of the effects of knowledge appreciation and depreciation, where the particularly important findings emerge from comparisons of the transition probabilities between different types.

We find that, compared to whites, blacks are less likely to experience both knowledge appreciation and depreciation. However, if we compare the magnitudes of the marginal effects of knowledge appreciation or depreciation for blacks, they are less likely to experience knowledge depreciation relative to the possible benefit from knowledge appreciation. On the other hand, Hispanics are more likely to experience both knowledge appreciation and depreciation. However, the effect of knowledge depreciation is quantitatively substantially larger relative the the magnitude of the effect of knowledge appreciation. These patterns explain why the average returns to education for blacks are larger than for whites, and why they are substantially lower for hispanics.

[5]Whereas the study of factors determining knowledge depreciation is not in the scope of the current study, one of the plausible determinants of knowledge depreciation discussed in the literature are the career interruptions due to out of work spells. Recent study of Li (2013) argues that occupational choices may be driven by the considerations about possible job interruptions, which may be even stronger for occupations with frequent knowledge-updating requirements, such as the sciences, engineering, and finance.

The results in Panel C of Table 2 indicate that high school graduates are less likely to experience both knowledge depreciation and appreciation. The magnitude of the marginal effect of knowledge appreciation is larger (less negative), indicating on lower chances for knowledge depreciation. College graduates are highly likely to experience knowledge depreciation and less likely to benefit from knowledge appreciation.[6] These patterns shed light on why the returns to high school education are substantially larger than the returns to college education.

The rates of knowledge appreciation and depreciation for differen levels of education could be explained by the more specialized nature of the college education (see Kinsler and Pavan, 2012). A more basic and universally applicable high school education is not a subject to frequent updates in the curriculum, therefore it is not likely to loose its value quickly. At the same time appreciation of high school education is also not likely without taking further training. Job productivity for college graduates depends, among other factors, on whether individuals take jobs related to their fields of study (Kinsler and Pavan, 2012) and whether they choose occupations, which require frequent updating of knowledge (Li, 2013). Returns to college education may be sensitive to the compatibility between individual's field of study and the level of specificity of skills required by job (see Kinsler and Pavan, 2012). In particular, workers who select into jobs not related to their chosen college major are likely to undercut their individual returns to education. In dynamics, the increasing disagreement between the specificity of skills and the choice of pursued careers may exacerbate and result in greater depreciation of college education.

The evidence on knowledge depreciation is represented empirically in the study of Weber (2013), who estimates human capital depreciation rate for different groups of workers in Switzerland. He finds that human capital depreciation appears to be related to the type of skills possessed by a worker. He finds that workers having vocational (specialized) studies suffer from a larger human capital depreciation than workers having traditional studies. He concludes that traditional studies, providing general skills, better protect workers against depreciation than vocational studies, which provide specific skills.

Wage offer

[6] Neuman and Weiss (1995) estimate Mincerian equation under the assumption that, on average, vintage effects of are more important for the more highly educated people. Our results provide an encouraging evidence to support this assumption.

Table 3 reports the estimates of the parameters of log-wage equation. The estimation results show that the wage profile admits an expected hump shaped pattern over time, with the coefficients of time trend being precisely estimated for each time polynomial. Demographic and socio-economic variables add an individual characterization to the common time trend with coefficients on them having an expected sign. The results of Neal and Johnson (1996) and Altonji and Blank (1999) indicate that, among other factors, the substantial part of the wage gap between races in the NLSY is due to the differences in measured abilities. Indeed, the estimation results show that the ability level, measured by the Armed Force Qualification Test (AFQT) has a significant influence on wages.

Probability of Grade Promotion

An individual who decides to enroll in a particular grade level may or may not be promoted from the grade. This probability of promotion is of interest in its own right, and is also a key ingredient in the estimation. We parameterize the grade-transition probability as logit. Under the restrictions of the model, the predetermined variables (variables that are determined by period $t-1$'s action) are completed grade G_{it}, tuition subsidy $TSUB_{it}$, and the level of experience E_{it}.

Table 4 reports the result of the logit regression of the probability of completing a grade. The results in Table 4 indicate that the probability of grade level promotion is increasing and concave in time spent on schooling activities. At the same time, this probability is decreasing and convex in hours spent in the labor market. The results in Table 4 indicate that current labor market participation is positively correlated with the probability of grade promotion. This provides evidence for the congruence hypotheses as per D'Amico (1984).

We find that blacks have a greater probability of being promoted a grade level than their white counterparts. Hispanics also have a higher probability of being promoted than their white counterparts. This result is in contrast to the classical drop-out story of minorities.

References

Aguirregabiria, V. and P. Mira (2002): "Swapping The Nested Fixed Point Algorithm: A Class of Estimators for Distrete Markov Decision Models," *Econometrica*, 70, 1519–1543.

Altonji, J. G. and R. M. Blank (1999): *Race and Gender in the Labor Market*, Elsevier Science Publishers.

Altug, S. and R. A. Miller (1998): "Effect of Work Experience of Female Wages and Labour Supply," *The Review of Economic Studies*, 65, 45–85.

Arcidiacono, P. and R. A. Miller (2011): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Econometrica*, 79, 1823–1867.

Bacolod, M. and J. Hotz (2005): "Cohort Changes in the Transition from School to Work: What Changed and What Consequences did it Have for Wages?" *Unpublished Manuscript:* $http : //www.econ.ucla.edu/hotz/working_{p}apers/cohort.pdf$.

Belzil, C. and J. Hansen (2002): "Unovserved Ability and the Return to Schooling," *Unpublished Manuscript, Concordia University*.

Bound, J., M. Lovenheim, and S. Turner (2007): "Understanding the Decrease in College Completion Rates and the Increased Time to th Baccalaureate Degree," *Population Studies Center Report 07-626*.

Bound, J. and S. Turner (2002): "Going to War and Going to College: Did World-War II and the G.I. Bill Increase Educational Attainment for Returning Veterans?" *Journal of Labor Economics*, 20, 784–815.

Card, D. (1999): "The Causal Effect of Education on Earnings," *Handbook of Labor Economics*, 3, 1801–1863.

Carnevale, A. P., S. Rose, and B. Cheah (2011a): "The College Payoff: Education, Occupation, and Lifetime Earnings," *Georgetown University, Center on Education and the Workforce*.

Carnevale, A. P., J. Strohl, and M. Melton (2011b): "What's It Worth?: The Economic Value of College Majors," *Georgetown University, Center on Education and the Workforce*.

D'Amico, R. (1984): "Does Employment During High School Impair Academic Progress?" *Sociology of Education*, 57, 152–164.

ECKSTEIN, Z. AND K. I. WOLPIN (1999): "Why Youths Drop Out of High School: The Impact Of Preferences, Opportunities, and Abilities," *Econometrica*, 67, 1295–1339.

EHRENBERG, R. G. AND D. R. SHERMAN (1987): "Employment While in College, Academic Achievement, and Postcollege Outcomes," *Journal of Human Resources*, 23, 1–23.

GAYLE, W.-R. (2013): "CCP Estimation of Dynamic Discrete/Continuous Choice Models with Generalized Finite Dependence and Correlated Unobserved Heterogeneity," *MIMEO: University of Virginia*.

HECKMAN, J. J. AND P. A. LAFONTAINE (2010): "The American High School Graduation Rate: Trends and Levels," *Review of Economics and Statistics*, 92, 244–262.

HECKMAN, J. J., L. J. LOCHNER, AND P. E. TODD (2008): "Earnings Functions and Rates of Return," *Journal of Human Capital*, 2, 1–31.

HOTZ, J. V. AND R. A. MILLER (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529.

IMAI, S. (2000): "Intertemporal Labor Supply and Human Capital Accumulation," *Unpublished Manuscript, Pennsylvania State University*.

KALENKOSKI, C. AND S. PABILONIA (2010): "Parental transfers, student achievement, and the labor supply of college students," *Journal of Population Economics*, 23, 469–496.

KASAHARA, H. AND K. SHIMOTSU (2009): "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices," *Econometrica*, 77, 135–175.

KEANE, M. P. (2002): "Financial Aid, Borrowing Constraints and College Attendance: Evidence from Structural Estimates," *American Economic Review (Papers and Proceedings)*, 92, 293–297.

KEANE, M. P. AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *The Journal of Political Economy*, 105, 473–522.

——— (2001): "The effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 42, 1051–1103.

KINSLER, J. AND R. PAVAN (2012): "The Specificity of General Human Capital: Evidence from College Major Choice," Tech. rep.

LI, H.-H. (2013): "The Effects of Human Capital Depreciation on Occupational Gender Segregation," *MIMEO, University of Wisconsin-Madison*.

McFadden, D. (2008): “Human Capital Accumulation and Depreciation,” *Applied Economic Perspectives and Policy*, 30, 379–385.

Neal, D. A. and W. R. Johnson (1996): “The role of premarket factors i black-white wage differences,” *Journal of Political Economics*, 105, 869–895.

Neuman, S. and A. Weiss (1995): “On the effects of schooling vintage on experience-earnings profiles: Theory and evidence,” *European Economic Review*, 39, 943–955.

Oettinger, G. S. (1999): “Does High School Employment Affect High School Academic Performance?” *Industrial and Labor Relations Review*, 53, 136–151.

Ruhm, C. J. (1997): “Is High School Employment Consumption of Investment?” *Journal of Labor Economics*, 15, 735–776.

Stinebrickner, T. R. and R. Stinebrickner (2003): “Working During School and Academic Performance,” *Journal of Labor Economics*, 21, 473–491.

——— (2004): “Time-Use and College Outcomes,” *Journal of Econometrics*, 121, 243–269.

Weber, S. (2013): “Human Capital Depreciation and Education Level,” *International Journal of Manpower*, forthcoming.

A Data and Sample Construction

The data is taken from the 1979 youth cohort of the National Longitudinal Survey of Labor Market Experience (NLSY79), a comprehensive panel data set that follows individuals over the period 1979 to 2000, who were 14 to 21 years of age as of January 1, 1979. The data set initially consisted of 12,686 individuals: a representative sample of 6,111 individuals, a supplemental sample of 5,295 Hispanics, non-Hispanic blacks, and economically disadvantaged, non-black, non-Hispanics, and a supplemental sample of 1,280 military youth. Interviews were conducted on an annual basis though 1994, after which they adopted a biennial interview schedule. This study makes use of the first 16 years of interviews, from 1979 to 1994. By 1990, the NLSY79 experienced attrition of 2,250 sample members, of which 1,097 were from the supplemental sample of military youth. I discuss briefly the construction of some of the key variables used in estimation

Employment
The NLSY79 collects detailed work history data for individuals in the sample. The work history data includes beginning and ending dates for all of 5 possible jobs, a maximum of

5 possible gaps in employment with each of the 5 possible jobs, the usual hours worked per day or per week on each job, and the hourly rate of pay on each job. The biggest complication in calculating hours worked is the fact that it must be calculated for the relevant year, which is the school year in this case. Since the actual weeks that comprise the school year vary from state to state, the dates chosen for the school year are somewhat arbitrary. Following Eckstein and Wolpin (1999), the year for those not attending school starts at October 1^{st} in year t and ends September 30^{st} of year t+1. For those attending school the school year instead ends at June 30 of year t+1. Weeks employed is then calculated based on these calendar dates. Hours worked per week or per day and hourly rate of pay is reported retrospectively back to the previous interview date. These variables were also adjusted to the above specified calendar dates. From these, we then construct hours worked for the relevant years, as well as average hourly rate of pay and an employment rate variable, which is the fraction of the relevant year in which the respondent was actively employed.

Education
The NLSY79 also collects information on the respondents' education. In particular, the NLSY79 collects , among others, enrollment status, highest grade level completed, current grade level, and degree held. The primary variables used in the paper are highest grade completed and enrollment status. In 1981, the NLSY collected information on the patterns of school activities of the respondents that are enrolled in school. In particular, the NLSY asked these respondent about the amount of hours they spent in school during the week before the interview date. They asked whether or not the time the reported is typical or not, and if no, to report the typical hours spent in school. The NLSY also asked the respondents to report the number of hours they spent studying outside of school during the week before the interview date. The response to these questions are used in the paper to estimate the study pattern of individuals enrolled in school.

There are a number of missing observations on highest grade completed. Many of these missing observations could be recovered from the information provided by enrollment status and highest grade completed in other years by the respondent. Since the model relies very much on the data on highest grade completed, we decide not to impute those years that are not recoverable with very high confidence.

The model construction and estimation requires data on the cost of schooling for an individual who decides to enroll in school. The yearly in-state tuition and required fees for four-year institutions and two-year institutions are taken from the NCES web site. Also, to identify the the aggregate shocks in wages and consumption, all nominal variables have to normalized to the same base year. To do this, the CPI is taken from the BLS web site, and converted to have a base year of 1981.

Asset holdings
Beginning in 1985, the NLSY79 began collecting comprehensive information on the asset

holdings of the respondents. This information was collected annually up to and including 1994, except for the year 1991 where asset data is missing. The best way to deal with these missing observations on asset holdings depends on exactly how the data will be used in estimation. In the case of Keane and Wolpin (2001) and Imai (2000), asset holding itself plays a central role in their model. Their method of imputation was therefore to model and asset holdings as normally distributed, and the estimate the mean and variance, from which they impute the missing years. In my case however, I require savings balance to impute total family consumption. For years in which the data is available, this is simply the difference between the Asset holding from one year to the next. For the years in which the data is missing, I take savings balance to be zero. For the early years of the cohort, net savings is relatively small and centered around zero. This suggests that the bias induced by this imputation is small. Furthermore, in estimating the consumption equation, savings is one the right hand side of the equation. The consistency of parameter estimates in the case where the left hand side variable is measured with a mean zero error is well documented in classical econometric textbooks. Finally, if there were large biases introduced by this imputation, they would show up in the estimated aggregate prices, These is no unusual visible discrete change in estimated aggregate prices for these periods. All these reasons lead me to believe that such imputations results in minimal biases in the parameters of interest.

Consumption
The NLSY79 does not collect data on individual consumption. However, the unique advantage of this data set that it collects detailed information on individual asset holding. To estimate the parameters in the above equation, family consumption is imputed from family income, family savings, four year schooling costs, and two year schooling costs. The way this is done is a follows. Subtracting family savings is taken from family income gives an estimate of the total resources available to the family in that year, net of savings. If the individual goes to high school, then his cost of schooling is assumed to be 0. If he goes to a two-year college, his cost of schooling is the two-year tuition cost, and if he goes to a four-year college, his cost of schooling is the four-year tuition cost. The individual's cost of schooling is subtracted from his individual resources. The yearly averages of the imputed consumption is given in Table 2.

Demographics
Demographic and family background variables collected by the NLSY79 and used in this study include age, race, mother's education, Father's education, family income, and year of experience working. Experience is calculated from the employment history section of the data set, which gives complete employment status for each year. Missing observations in family income are imputed by first using a three year moving average smoothing technique, followed by regressing family income on other covariates, some of which not listed here, and using the predicted income for the cases in which family income is missing. The resulting distribution of imputed family income match the distribution of actual (observed)

family remarkably well.

Sample Restriction
As stated above, the data employed in this paper span the years of 1979 though 1994. The model specified in section (**??**) does not include the decision to enter the military, and thus as the first restriction on the data we drop all males who enter the military in 1979. This restriction reduces the sample size to 11406. As stated above, we drop respondents for cases where missing observations in highest grade completed cannot be recovered with very high confidence. This reduces the sample to 7814 respondents. This is clearly are somewhat severe restriction on the data, and it may pay to invest is less restrictive imputation rules. This however is not pursued here. In the literature, female members are treated differently from male sample members. The choice set of a female is generally considered larger than that of a male. The additional decisions usually included in the choice set for women are marriage decisions and fertility decisions. To avoid these additional complications, the data is restricted to include males only. This results in a sample size of 3916 male respondents. The summary statistics and all estimations make use of this sample.

Table 2: Summary of the distribution of the returns to education

Panel A: Average returns to education

	AVERAGE	WHITE	BLACK	HISPANIC
AVERAGE	0.1405 (0.0906)	0.1537 (0.0993)	0.1617 (0.1046)	0.0582 (0.0363)
SOME HIGH SCHOOL	0.1475 (0.0952)	0.1560 (0.1009)	0.1679 (0.1087)	0.1059 (0.0678)
HIGH SCHOOL GRAD	0.1546 (0.0999)	0.1796 (0.0116)	0.1683 (0.1089)	0.0381 (0.0230)
COLLEGE GRAD	0.0733 (0.0463)	0.0766 (0.0485)	0.0935 (0.0596)	0.0103 (0.0048)

Panel B: Type Transition Probabilities

	Type LOW	Type HIGH
Type LOW	0.1309	0.0356
Type HIGH	0.2009	0.6327
Support	0.0024 (0.0013)	0.2090 (0.1405)

Panel C: Marginal effects on type transition probabilities

	LOW - LOW	LOW - HIGH	HIGH - LOW	HIGH - HIGH
BLACK	−0.0715 (0.0097)	−0.1667 (0.0363)	−0.3392 (0.0079)	0.5573 (0.0404)
HISPANIC	0.3096 (0.0260)	0.0997 (0.0171)	0.5340 (0.0057)	−0.9432 (0.0380)
HIGH SCHOOL GRAD	0.2758 (0.0155)	−0.1849 (0.0117)	−0.4292 (0.0109)	0.3383 (0.0193)
COLLEGE GRAD	0.4552 (0.0367)	−0.0464 (0.0244)	0.6282 (0.0028)	−1.0369 (0.0526)
HIGH AFQT	−0.0598 (0.0246)	−0.0110 (0.0162)	0.2898 (0.0018)	−0.2190 (0.0351)

Standard errors in parenthesis.

Table 3: The Wage Equation

Variable	Estimate	Std. Err.
Intercept and Time Dummies		
CONSTANT	-0.0128	0.7446
t	-2.7984	0.7368
t^2	27.0378	0.3027
t^3	-98.1293	0.1199
t^4	167.7391	0.0431
t^5	-132.0163	0.0148
t^6	40.6886	0.0050
Demographics and Socio-Economic Variables		
BLACK	0.5596	0.1425
HISPANIC	0.2436	0.5032
AFQT	5.0e-9	1.5e-9
EXPERIENCE	0.0531	0.0305
SQUARED EXPERIENCE	-0.0022	0.7745

Table 4: Probability of Grade Promotion

Variable	Estimate	Std. Err.
CONSTANT	2.6024	0.3238
Time Use Variables		
STUDY TIME	3.9949	1.2624
STUDY TIME SQUARED	-4.5421	3.5925
HOURS WORKED	-6.6799	1.1755
HOURS WORKED SQUARED	12.0961	3.6024
Enrollment and Participation Variables		
LAGGED ENROLLMENT	-1.9988	0.3724
EMPLOYMENT	0.3120	0.0929
Socio-Economic Variables		
AGE$\times$ EDUCATION	-0.0009	0.0001
BLACK	0.9135	0.0992
HISPANIC	0.1978	0.0954
AFQT	0.5848	0.0271

Table 5: List and Description of Variables Used

Employment, Financial	
d^s_{nt}	Indicator variable equal to 1 if individual n enrolls in year t
d^s_{nt}	Indicator variable equal to 1 if individual n works in year t
s_{nt}	Fraction of time spent on school activities in year t
h_{nt}	Fraction of time spent working in year t
S_{nt}	Completed level of education
E_{nt}	Level of experience
AGE_{nt}	Age at year t
$WHITE$	Indicator variable equal to 1 if White and 0 otherwise
$BLACK$	Indicator variable equal to 1 if Black and 0 otherwise
$HISPANIC$	Indicator variable equal to 1 if Hispanic and 0 otherwise
FAM_INC_{nt}	level of family income at year t
FAM_SIZE_{nt}	size of n's household at year t
FAM_AGE_{nt}	average age of n's household at year t
$SIBLINGS$	number of siblings of n as at age 14
US_BORN	indicator variable equal to 1 if n was born in the US
$AFQT$	The Armed Force Qualification Test score for individual n
$ASSETS$	Level of asset holdings by the household of n in year t
$UNEMP$	Level of the unemployment rate local to n in year t
$RURAL$	Indicator variable equal to 1 if n lives in a rural area in year t
$TUITION$	Level of college tuition that individual n is subject to in year t

Table 6: Summary Statistics

Year	1979	1980	1981	1982	1983	1984	1985	1986
Observations	3749	3512	3595	3575	3594	3549	3504	3413
d_0	0.0205	0.0529	0.1115	0.1325	0.1719	0.1541	0.1435	0.1300
d_1	0.0381	0.1452	0.2842	0.4215	0.5158	0.6198	0.6889	0.7380
d_2	0.5644	0.3809	0.2439	0.1367	0.0951	0.0617	0.0345	0.0240
d_3	0.3769	0.4208	0.3602	0.3090	0.2170	0.1642	0.1329	0.1078
d^s	0.9413	0.8018	0.6041	0.4458	0.3121	0.2259	0.1675	0.1318
s	1436.5	1354.6	1276.0	1203.3	1149.7	1139.3	1114.6	1077.3
S	9.7967	10.730	11.335	11.842	12.198	12.416	12.578	12.708
d^h	0.4150	0.5660	0.6445	0.7306	0.7328	0.7841	0.8219	0.8458
h	710.90	972.82	1080.5	1159.8	1310.0	1477.6	1577.7	1694.5
E	1.2107	1.6136	2.1655	2.8036	3.5166	4.2310	4.9877	5.8025
w^1	4.3872	4.1601	4.3383	4.6541	4.8560	5.1220	5.5749	6.0788
AGE	16.743	17.653	18.695	19.697	20.706	21.699	22.690	23.688
$WHITE$	0.5727	0.5769	0.5713	0.5757	0.5759	0.5711	0.5736	0.5722
$BLACK$	0.2625	0.2640	0.2651	0.2626	0.2613	0.2646	0.2606	0.2625
$HISPANIC$	0.2648	0.1592	0.1635	0.1617	0.1627	0.1643	0.1658	0.1653
FAM_INC^1	17647	19086	20011	21168	21398	21785	23577	25319
FAM_SIZE	4.8434	4.5948	4.3171	3.9625	3.7045	3.3722	3.1726	2.9856
FAM_AGE	26.225	26.823	26.978	26.665	26.699	26.653	26.538	26.175
$SIBLINGS$	3.6220	3.5899	3.6069	3.6204	3.6165	3.6238	3.6204	3.6024
US_BORN	0.9306	0.9328	0.9310	0.9311	0.9315	0.9323	0.9326	0.9326
$AFQT$	42.024	43.186	42.793	42.835	42.774	42.606	42.545	42.565
$ASSETS^1$						4141.2	4278.8	4998.8
$UNEMP$	2.5646	2.8476	3.1652	3.7848	4.1978	3.4356	3.2919	3.1693
$RURAL$	0.2125	0.20871	0.1997	0.1932	0.1830	0.1718	0.1680	0.1614
$TUITION^1$	813.19	793.04	809.79	865.54	916.18	960.77	1029.0	1087.4

[1] In 1981 dollars

Table 7: Summary Statistics (Contd.)

Year	1987	1988	1989	1990	1991	1992	1993	1994
Observations	3338	3357	3389	3328	2931	2936	2937	2896
d_0	0.1207	0.0965	0.0994	0.0943	0.1044	0.1226	0.1113	0.1142
d_1	0.8001	0.8394	0.8574	0.8647	0.8614	0.8474	0.8593	0.8649
d_2	0.0155	0.0071	0.0023	0.0006	0	0	0	0
d_3	0.0635	0.0568	0.0407	0.0402	0.0341	0.0299	0.0292	0.0207
d^s	0.0790	0.0640	0.0430	0.0408	0.0341	0.0299	0.0292	0.0207
s	1043.2	977.74	970.54	962.93	976.60	1006.7	1118.6	1128.3
S	12.833	12.890	12.917	12.962	13.050	13.049	13.073	13.08
d^h	0.8636	0.8963	0.8982	0.9050	0.8955	0.8773	0.8886	0.8857
h	1836.4	2016.8	2078.7	2025.0	2072.1	2126.6	2076.2	2111.7
E	6.6363	7.4566	8.2912	9.1908	10.022	10.853	11.676	12.548
w^1	7.0968	7.6098	7.6038	8.0964	7.7159	7.8402	8.2973	8.4466
AGE	24.680	25.684	26.686	27.687	28.624	29.620	30.621	31.611
$WHITE$	0.5733	0.5737	0.5716	0.5736	0.5165	0.5150	0.5138	0.5162
$BLACK$	0.2657	0.2654	0.2653	0.2644	0.2972	0.2973	0.3006	0.2987
$HISPANIC$	0.1609	0.1609	0.1632	0.1620	0.1863	0.1877	0.1856	0.1851
FAM_INC^1	26572	29047	46666	34705	36938	59830	41624	43778
FAM_SIZE	2.8406	2.7768	2.7722	2.7641	2.8161	2.8692	2.9240	2.9229
FAM_AGE	26.154	25.624	25.707	25.814	26.108	26.231	24.292	24.610
$SIBLINGS$	3.6096	3.6136	3.6208	3.6283	3.6349	3.6294	3.6275	3.6339
US_BORN	0.9340	0.9368	0.9350	0.9353	0.9344	0.9335	0.9342	0.9350
$AFQT$	42.789	42.565	42.270	42.422	42.089	41.905	41.869	41.965
$ASSETS^1$	7107.8	7132.9	20246	10064	11688	13922	13488	12195
$UNEMP$	2.9331	2.6094	2.3865	2.4002	2.9512	3.1757	3	2.9499
$RURAL$	0.1791	0.1805	0.1844	0.1850	0.1641	0.1665	0.1722	0.1833
$TUITION^1$	1153.1	1170.5	1181.5	1234.6	1351.1	1404.9	1490.2	1504.5

[1] In 1981 dollars