A Dynamic Structural Model of Labor Supply and Educational Attainment†

Wayne-Roy Gayle‡
University of Pittsburgh

October 2006

Abstract

This paper develops and estimates a dynamic structural model of labor supply and schooling to investigate the process by which a cohort of males from the NLSY79 accumulate human capital via formal education and labor market participation. 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. In particular, the model explicitly accounts for the simultaneous choice of enrolling in school and working. It also allows for endogenous leisure choices, intertemporal nonseparabilities in preferences, aggregate skill specific productivity shocks, aggregate consumption price effects, and individual heterogeneity. Times spent on schooling, working, and leisure are treated as continuous choice variables. The estimates from the model are then used to conduct simulation exercises to evaluate policies that are aimed at affecting working while enrolled in school and equating school quality across races. The results indicate that these policies may have significantly different effects on different racial groups.

Keywords: Educational Outcome, Labor Supply, non-separability, Dynamic Models.
JEL classification: C14, C23, I20, J24

† The author is grateful to Jean-Francois Richard, Hugo Benitez-Silva, George-Levi Gayle, Robert Miller, Soiliou Namoro, and Holger Sieg for insightful comments and discussions. The author also thanks participants in seminars at Carnegie Mellon University, University of Pittsburgh, SUNY-Stony Brook, and the Econometric Society 2004 summer meetings. All errors are my own.

‡ Economics Department, University of Pittsburgh, 230 S. Bouquet St., Pittsburgh, PA. 15260, phone: (412) 648-5717, fax: (412) 648-1793, e-mail: `wag2@pitt.edu`

1 Introduction

Over the last three decades, there has been an increasing trend of young individuals participating in the US labor market while actively enrolled in school. Young individuals are increasing their incidence of labor market participation, 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 the concern that an intensive amount of working while in school may hinder academic performance and increase drop-out rates, thus jeopardizing future opportunities.[2] 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. Furthermore, working while in school produces immediate work experience and cash that may be used to finance their studies.[3] It is not obvious which of these two opposing effects dominate. It may be that the net effect of these opposing forces varies over different groups of young individuals.

This article develops and estimates a dynamic structural model of schooling and work decisions to investigate the process by which a cohort of young males accumulate human capital over their life cycle. The theoretical model provides a detailed treatment of the economic costs and benefits associated with the schooling and labor supply alternatives faced by individuals. Specifically, the estimated model explicitly accounts for the simultaneous choice of enrollment in school and labor force participation, endogenous leisure choices, intertemporal nonseparabilities in preferences, aggregate skill specific productivity shocks, aggregate consumption price effects, and individual heterogeneity.

In addition to accounting for the simultaneous choice of work and schooling, the model treats hours spent on schooling, working, and leisure as continuous choice variables.[4] This approach is in contrast to other models (see Keane and Wolpin, 1997, and Eckstein and Wolpin (1999) for examples) that treat leisure time as exogenous to the individual, where an increase in labor supply is equivalent to a decrease in time spent on schooling activities if the individual is enrolled in school. In this framework, an individual may optimally choose to sacrifice leisure and increase time spent on both schooling and labor market activities. In this sense the model is one of optimal intra- and inter-temporal allocation of time among schooling, working and leisure. The model also allows for flexible specification of preferences with respect to time allocation. The additional flexibility comes from the specification of intertemporal

[1] A recent documentation of this phenomena is found Bacolod and Hotz (2005).

[2] This apprehension is reflected in the article entitled "Long hours taking toll on youths, studies say," by Paloma McGregor, The Plain Dealer, March 5, 2001.

[3] This opinion was expressed in the article entitled "Teens Find Profit and Loss in Work: Part time jobs bring experience and cash, but can hinder studies," by Jacqueline Salmon, The Washington Post, March 28, 1998.

[4] While some studies model these alternatives as mutually exclusive (Keane and Wolpin, 1994; Cameron and Heckman, 1999), the growing trend is to allow for interior solution to choices where individuals simultaneously participate in the labor market and attend school (see D'Amico, 1984, Ruhm, 1997, Oettinger, 1999, and Eckstein and Wolpin, 1999 for examples)

nonseparabilities in leisure.

Recent studies of the life-cycle models of labor supply have stressed the importance of intertemporally non-separable preferences.[5] Hotz, Kydland, and Sedlacek (1988) found that the assumption of intertemporally separable preferences for leisure is inconsistent with data for prime-age males. Given that hours schooling activities and leisure are related by the time constraint of the individual, such nonseparabilities are also likely to affect their enrollment and study patterns. The estimation results indicate that leisure choices are intertemporal complements. Increases in current hours of leisure increases the future demand of leisure. In other words, an increase in hours of current schooling activities decreases the future marginal disutility of schooling. This evidence of intertemporal complimentarity suggests habit formation by young men.

The primary data used in this study comes from the National Longitudinal Survey of Youth (NLSY79), which is a comprehensive panel data set that follows individuals who were 14 to 21 years of age as at January 1, 1979. The estimation technique implemented is a modified version of the Conditional Choice Probability (CCP) estimator of Hotz and Miller (1993) and Altug and Miller (1998). This estimation technique allows for unobserved individual-specific effects to be arbitrarily correlated with the observed characteristics in the model. The model employs a fixed effects method of controlling for unobserved heterogeneity. Other models of education, such as Eckstein and Wolpin (1999) control for individual-specific effects by way of a random-effects, finite mixture specification. These techniques typically require that the investigator make strong independence assumptions on the relationship between the unobserved covariates, and their observed counterparts. The cost of the flexibility allowed by a fixed effects specification is the resulting incidental parameters problem. We argue, using previous results (Altug and Miller, 1998; Gayle and Miller, 2003) and evidence from the data used in this paper that these biases are likely to be small.

The incidence of working, the number of hours worked, and the number of years that young men spend working while enrolled in school varies across races. Bacolod and Hotz (2005) documents that the number of years working while in high school increased the most for young Hispanic men, followed by young black men. Young black men experienced the largest increase in working while in college. In estimating the parameters of the model, we pay special attention to racial differences in outcomes that are not accounted for by the rich set of observed background variables found in the NLSY79, nor by estimated individual specific effects. The theoretical model provides a natural separation of these unexplained racial variations into preference differences and statistical discrimination (Altonji and Blank, 1999).

The empirical results indicate that, conditional on enrolling, young black males are likely to spend more time on school activities than white males. Young Hispanic males are likely to spend less time on school activities white males. Furthermore, young black and Hispanic males are less likely to be promoted from the grade level than young white males. These young

[5] See Hotz et al., 1988, Eichenbaum et al., 1988, Altug and Miller, 1998, Imai, 2000, and Gayle and Miller, 2003 for examples.

minority males either repeat the grade level or drop out of school during the school year. These racial differences remain significant after the inclusion of the rich set of demographic variables and measures of ability that are found in the NLSY79, as well as measures of unobserved individual specific characteristics. The lower probability of grade promotion for blacks and hispanics is interpreted as statistical discrimination in the school environment. In the paper we argue that this grade promotion probability gap is a measure of the differences in the quality of schools that blacks and Hispanics attend as against the quality of schools that whites attend.

Controlling for racial differences in wages, and the aforementioned racial differences in study patterns and grade promotion propensities, the results indicate that young black and Hispanic males are more likely to enroll in school. This likelihood of enrollment is higher for blacks than for Hispanics. Furthermore, blacks and Hispanics are likely to spend more hours on leisure than their white counterparts, with blacks spending more time than Hispanics. In the framework of the model these differentials in propensity to enroll and consume leisure are interpreted as racial differences in preferences over schooling and leisure. Interestingly, the results indicate that there are no racial differences in the propensity to participate in the labor market.

The model is solved and simulated in order to analyze the effects of various hypothetical policies. The first policy analyzed is one where the government subsidizes students who decide not to participate in the labor market. The simulated results indicate that this policy does very little in affecting the level of education, labor market experience, and wages on young men. The second policy analyzed is one where the school administration adjusts the school curriculum so that young men who enroll necessarily spend more time on school activities. Such a policy can be achieved by increasing the number of hours in school, increasing the number or difficulty of assignments, after school programs, or Saturday (Sunday) classes. The results indicate that such a policy has significant positive effects on whites and Hispanics, but not on blacks. The level of education and wages for whites and Hispanics increase significantly, while their level of experience reduce by a smaller percentage. On the other hand, the level of education of blacks increase only marginally. Their level of experience decreases during school years, and increases for post school years. The net result is that wages for blacks flatten out over the life cycle.

The final simulation exercise analyzes a situation where school quality of blacks and Hispanics are equated to those of whites. The results indicate that this policy has significant positive effects on the level of education and wages of blacks. The effects of this policy on Hispanic are positive but much more modest than that for blacks.

The rest of the paper is organized as follows. In the next section, we present the basic behavioral model. We then discuss the solution of the model in section (3) and describe the first order necessary conditions for optimality that will be used in estimation. Section (4) discusses the construction of the sample used in estimation, and Section (5) discusses the empirical methodology implemented in estimation of the parameters of interest. Section (6) describes the estimation of the consumption function and discusses the empirical findings. Section (7) discusses the estimation of the wage equation and the empirical findings. Section (8) discusses the estimation of the time spent on schooling activities and the transition probabilities. Section (9) presents the methodology used to estimate the conditional choice probabilities and their corresponding derivatives, which are needed to estimate the preference parameters. Section (10) presents the moment conditions and corresponding sample analogs that are used in estimating the preference parameters of the model, as well as discuss the empirical findings of the model. Section (11) presents the method of solving the dynamic programming model and discusses the policy simulations. Section (12) concludes.

2 The Theoretical Model

This section develops the theoretical framework that is used to investigate how individuals allocate time between human capital accumulation, labor market participation, and leisure.

2.1 Environment

The model is set in discrete time $t \in \{0, 1, \cdots, T\}$. We assume that there exists a continuum of individuals on the unit interval [0,1]. Associated with each individual is a K-dimensional vector of exogenous covariates, denoted z_{nt}, which is assumed to be independently distributed over the population with known cumulative distribution function $Q_0(z_{nt+1}|z_{nt})$. In each period, individual $n \in [0, 1]$ is endowed with a fixed amount of time normalized to one. He must choose how to allocate this unit of time between leisure l_{nt}, the time spent on labor market activities h_{nt}, and the time spent on school activities s_{nt}:

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

Define $d^h_{nt} \equiv 1_{\{h_{nt}>0\}}$ and $d^s_{nt} \equiv 1_{\{s_{nt}>0\}}$ where $1_{\{\cdot\}}$ is the indicator function equal to one if the event in parentheses occurs and zero otherwise. There is a single composite consumption good in the economy which is consumed and traded by all individuals. Let c_{nt} denote this composite good.

We assume the model has a Markov structure, in which the individual does not need to remember the full history to solve this problem, but only a summary statistic x_{nt}, belonging to a finite vector space $\mathcal{X}$. In particular, define $(h_{nt-\rho}, \cdots, h_{nt-1})$ as the ρ-dimensional vector of past labor supply outcomes, $(s_{nt-\rho}, \cdots, s_{nt-1})$ as the ρ-dimensional vector of past time spent on schooling activities, S_{nt} as the highest grade completed by individual n as at the beginning of t, and E_{nt} as the total years of labor market experience accumulated by individual n as at the beginning of period t. Define also $(c_{nt-\rho}, \cdots, c_{nt-1})$ to be the ρ-dimensional vector of past consumption. Then the typical observed state vector for individual n at time t is given by the

$(3\rho+k+1)$-dimensional vector[6]

$$x_{nt} \equiv (h_{nt-\rho},\cdots,h_{nt-1},s_{nt-\rho},\cdots,s_{nt-1},S_{nt-\rho+1},\cdots,S_{nt},c_{nt-\rho},\cdots,c_{nt-1},E_{nt-\rho},z'_{nt})' \quad (2.2)$$

Given that individual n has chosen to enroll in school, he may or may not complete that grade level. If he does complete the grade he is currently enrolled in, his level of education increases by one grade. Otherwise, his level of education remains unchanged. The probability that an individual advances a grade level given that he has enrolled in school at the beginning of period t is denoted by $F(x_{nt})$.

2.2 Technology

We assume that the individual has access to a sector specific production technology in each period where, if he works in sector $j=1,\cdots,J$, he produces a quantity of the output $w_{ntj}h_{nt}$. Here, w_{ntj} is marginal product of labor of individual n at time t with skill level j. It is assumed that w_{ntj} is composed of J exogenously determined time specific aggregate skill prices ω_{tj}, an individual specific, time invariant productivity effect, μ_n, and a skill specific function of his stock of human capital, his socio-economic characteristics and other state vectors, $\gamma_j(x_{nt})$:

$$w_{ntj} = \omega_{t,j}\mu_n\gamma_j(x_{nt}), \quad (2.3)$$

Thus $\mu_n\gamma_j(x_{nt})$ is the number of efficiency units of labor supplied by the worker per unit of time in sector j, while $\omega_{t,j}$ is the time specific aggregate price of skill in sector j.

2.3 Choice Set

This model falls within the class of mixed continuous and discrete Markov decision processes. The continuous choice variables in this model are c_{nt},h_{nt},and s_{nt}. If $h_{nt}=0$, individual n does not work at time t. Otherwise, the individual works for the fraction of time $h_{nt}>0$. Likewise if $s_{nt}=0$, individual n does not attend school at time t. Otherwise, the individual studies for the fraction of time $s_{nt}>0$. Define the discrete choice variables for each individual $n\in[0,1]$ at time $t\in\{0,1,\cdots,T\}$:

$$d_{nt0} \equiv \begin{cases} 1 & \text{if } d^h_{nt}=0 \text{ and } d^s_{nt}=0 \\ 0 & \text{otherwise} \end{cases}, \quad (2.4)$$

$$d_{nt1} \equiv \begin{cases} 1 & \text{if } d^h_{nt}=1 \text{ and } d^s_{nt}=0 \\ 0 & \text{otherwise} \end{cases},$$

$$d_{nt2} \equiv \begin{cases} 1 & \text{if } d^h_{nt}=0 \text{ and } d^s_{nt}=1 \\ 0 & \text{otherwise} \end{cases},$$

$$d_{nt3} \equiv \begin{cases} 1 & \text{if } d^h_{nt}=1 \text{ and } d^s_{nt}=1 \\ 0 & \text{otherwise} \end{cases}.$$

[6] To conserve on notation in what follows, we will use x_{nt} to denote any subset of this vector.

2.4 Preferences

Similar to models such as Heckman (1976) and Eckstein and Wolpin (1999), we assume that attending school provides some consumption value to the individual. Learning may be directly valued by the individual, and social interaction within the school environment may provide positive consumption value. However, in this specification, this consumption value of attending school is not confounded with the loss in leisure due to schooling activities since leisure is modelled directly. We specify the contemporaneous utility of attending school as follows:

$$U_{nt1} = u_1(d^s_{nt}, x_{nt}). \tag{2.5}$$

Similarly, we assume that there is a utility associated with labor market participation. We specify this contemporaneous utility of labor force participation as follows:

$$U_{nt2} = u_2(d^h_{nt}, x_{nt}). \tag{2.6}$$

Preferences are assumed to be additive in consumption and leisure, but not separable with respect to leisure over time. The contemporaneous utility of leisure is therefore given by:

$$U_{nt3} = u_3(x_{nt}, l_{nt}). \tag{2.7}$$

The utility of leisure is specified to be dependent on current leisure level and the level of leisure consumed over the last ρ periods.[7]. We assume that u_3 is increasing and concave in l_{nt}. The utility derived from the consumption good in time t is also assumed to be increasing and concave in c_{nt} and is denoted by

$$U_{nt4} = u_4(c_{nt}, z_{nt}). \tag{2.8}$$

We introduce a vector of choice specific utility shifters $(\varepsilon_{nt0}, \cdots, \varepsilon_{nt3})'$, which are assumed to be independent over (n,t) and drawn from a population with a distribution function $Q_1(\varepsilon_{nt0}, \cdots, \varepsilon_{nt3})$. They are interpreted to be choice specific, time-varying characteristics that partially determine the utility associated with the corresponding alternatives and unobserved to the econometrician. Let $\beta \in (0,1)$ denote the common subjective discount factor, and E_0 denote expectation conditional on the information set at date 0. The expected discounted lifetime utility of individual n is given by:

$$E_0\Big\{ \sum_{t=0}^{T} \beta^t \Big[\sum_{k=1}^{4} d_{ntk}(U_{nt1} + U_{nt2} + U_{nt3} + U_{nt4} + \varepsilon_{ntk}) \Big] \Big\}. \tag{2.9}$$

[7] The lags in leisure are not specified explicitly here since it is a subset of the state vector x_{nt} by equation (2.1)

3 The Optimization Problem

The inclusion of an aggregate component in marginal product of labor (2.3), complicates estimation. To make the model empirically tractable, we assume that markets are competitive and complete. Agents are price takers and there are no distortions in the market for the consumption good, labor supply and loans, a common interest rate facing borrowers and lenders, and that a rich set of financial securities exists to hedge against uncertainty. This assumption incorporates uncertainty in a sufficiently simple manner that leads to a tractable econometric model. Competitive and complete capital market assumption was used by Ben-Porath (1967), Blinder and Weiss (1976), Heckman (1976), and Shaw (1989) to analyze life cycle models of human capital accumulation. This assumption was also recently used by Altug and Miller (1990), Altug and Miller (1998), and Gayle and Miller (2003) to estimate life-cycle models of consumption, labor supply and fertility decisions with aggregate shock.

One key restriction that the assumption of competitive and complete markets places on the model is the lack of any binding borrowing constraint. Borrowing constraints are popular considerations in the study of educational choice. It is a widespread postulation that borrowing constraints critically restricts economically disadvantaged individuals from obtaining the level of formal education that they would have attained otherwise. However, the empirical evidence does not support this view. Cameron and Heckman (1999, 1998) conclude that it is the long-term influences of family and environment that account for ethnic and racial disparities in school attendance, and not short term liquidity constraints. Keane (2002) conclude that borrowing constraints have little effect on college attendance decisions. In the light of these and other evidences, we abstract from any considerations of liquidity constraints and thus the assumption of competitive and complete markets presents itself as an appealing approximation.

Under the assumptions of competitive and complete markets, we appeal to the fundamental welfare theorems which allows us to recast the optimization problem as a social planner problem. The objective function of the social planner is the weighted average of the expected discounted utilities of each individual n given in (2.9). The social weight attached to an individual is given by η_n^{-1}. The optimization problem of the social planner is subject to the time allocation constraint for each individual (2.1), as well as the production technology available to each individual as reflected in (2.3). Define L to be the lebesgue measure that integrates over the population. The aggregate feasibility condition is given by:

$$\int_0^1 [c_{nt} + a_{nt} + \pi_{nt} - w_{nt} h_{nt}] \mathrm{d}L(n) \leq 0, \qquad t \in \{0, 1, \cdots, T\}. \tag{3.1}$$

where a_{nt} is the individual savings at time t, or the value of claims to period $t+1$ consumption net of the claims to time t consumption. π_{nt} is the direct schooling expenses incurred by the individual if he chooses to enroll in period t.

The Pareto optimal allocations are found by maximizing

$$E_0\Big\{\int_0^1\sum_{t=0}^{T}\beta^t\eta_n^{-1}\Big[\sum_{k=1}^{4}d_{ntk}(U_{nt1}+U_{nt2}+U_{nt3}+U_{nt4}+\varepsilon_{ntk})\Big]\mathrm{d}L(n)\Big\}, \tag{3.2}$$

subject to (3.1) and (2.1) with respect to sequences for consumption, schooling, and labor supply $\{c_{nt}, s_{nt}, h_{nt}\}_{t=0}^{T}$ for all individuals $n \in [0,1]$.

3.1 Optimal consumption

Define $\beta^t\lambda_t$ as the Lagrange multiplier associated with the aggregate feasibility constraint in equation (3.1). Given the assumption of an interior solution for consumption allocation, the set of necessary conditions characterizing optimal consumption allocation are given by

$$\frac{\partial u_3(c_{nt}, x_{nt})}{\partial c_{nt}} = \eta_n\lambda_t, \tag{3.3}$$

for all $n \in [0,1]$ and $t \in \{0,\cdots,T\}$. Under the assumption of contemporaneous separability of consumption from education and labor supply choices, (3.3) can be used to solve for individuals' Frisch demand functions which determines optimal consumption allocation in terms of the time-varying characteristics x_{nt} and the shadow value of consumption $\eta_n\lambda_t$. Assume that the utility derived from consumption takes on the following augmented CRRA specification:

$$u_3(c_{nt}, x_{nt}) = g(x_{nt})\frac{c_{nt}^{\alpha}}{\alpha}. \tag{3.4}$$

Then condition (3.3) takes the form

$$g(x_{nt})c_{nt}^{\alpha-1} = \eta_n\lambda_t. \tag{3.5}$$

Multiplying (3.5) by $\alpha^{-1}c_{nt}$ gives the following alternative representation of the indirect contemporaneous utility derived from consumption:

$$u_3(c_{nt}, x_{nt}) = \frac{\eta_n\lambda_t}{\alpha}c_{nt}. \tag{3.6}$$

The empirical strategy comprises of estimating the parameters of the utility function u_3 from (3.3) and (3.4) to obtain estimates of the individual specific weights η_n as well as the Lagrange multiplier λ_t. These estimates are then substituted in (3.6), which is in turn substituted into the social planner's objective function (3.2).

Under the assumption that none of the consumption good is wasted at the optimal allocation, the first order necessary condition with respect to the the lagrange multiplier $\beta^t\lambda_t$ gives the optimal consumption allocation for each individual

$$c_{nt} = w_{nt}h_{nt} - a_{nt} - \pi_{nt}. \tag{3.7}$$

3.2 Optimal schooling and labor supply

Characterizing the optimal labor supply, leisure and schooling decision is more complicated. The optimal schooling and work allocations are confounded by the constraint imposed by (2.1). In particular, in any period, increasing both schooling and labor supplied by individual n necessarily leads to a decline in the level of leisure enjoyed by that individual. Consequently, the optimal allocation of labor supply, education and leisure cannot be separately solved for as in the case of optimal consumption allocation. Following Altug and Miller (1998), the conditional valuation functions associated with the discrete choices on individual n in period t is defined as:

$$V_{ntj} + \varepsilon_{ntj} \equiv \max_{\{s_{nr}, h_{nr}\}_{r=t}^{T}} E_t \left\{ \begin{array}{l} \sum_{r=t}^{T} \beta^{r-t} [\sum_{k=0}^{3} d_{nrk} (U_{nr0} + U_{nr1} \\ +\alpha^{-1} \eta_n \lambda_r (w_{nt} h_{nt} - a_{nt} - \pi_{nt}) + \varepsilon_{nrk}) | d_{ntj} = 1] \end{array} \right\}. \tag{3.8}$$

Let d^0_{ntj} be the socially optimal decision by individual n in period t. The term $V_{ntj} + \varepsilon_{ntj}$ denotes the social value from individual n choosing alternative j at time t. Accordingly, individual n's choice of alternative j at time t is optimal if

$$d^0_{ntj} = \begin{cases} 1, & \text{if } V_{ntj} + \varepsilon_{ntj} > V_{ntk} + \varepsilon_{ntk} \quad \forall\, k \neq j \\ 0, & \text{otherwise} \end{cases}. \tag{3.9}$$

Let h^0_{nt} and s^0_{nt} be the optimal interior choice of labor supply and study time. Given that it is socially optimal for individual n to work in time t, h^0_{nt} must satisfy

$$\frac{\partial V_{ntj}}{\partial h_{nt}} = 0, \qquad \text{for } j = 1, 3. \tag{3.10}$$

Likewise, given that it is socially optimal for individual n to enroll in time t, s^0_{nt} must satisfy

$$\frac{\partial V_{ntj}}{\partial s_{nt}} = 0, \qquad \text{for } j = 2, 3. \tag{3.11}$$

In order to express the conditional valuation function recursively, define p_{ntj} to be the probability of individual n choosing option j in period t conditional on the information set available to him in period t

$$p_{ntj} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{V_{ntj} - V_{nt0} + \varepsilon_{ntj}} \cdots \int_{-\infty}^{V_{ntj} - V_{nt3} + \varepsilon_{ntj}} \mathrm{d}Q_1(\varepsilon_{nt0}, \cdots, \varepsilon_{nt3}). \tag{3.12}$$

The information set available to individual n at period t is composed of the observed state vector x_{nt}, and the unobserved individual specific and aggregate shocks to productivity and consumption. Define this state vector as $\Psi_{nt} \equiv (x'_{nt}, \mu_n, \eta_n, \lambda_t, \omega_{t1}, \cdots, \omega_{tJ})'$. Define also $\mathcal{A}^i_{nt}$ to be the set of all possible realizations of the state vector for individual n at i periods after t given the realization of the state vector Ψ_{nt} at period t. Correspondingly, let $F_j(\Psi^{(i)}_{nt} | \Psi_{nt})$ is the probability that the state vector of individual n in period $t+i$ is $\Psi^{(i)}_{nt}$, given that his state vector in period t is Ψ_{nt} and he chooses alternative j in period t. Then from equation (3.9), the conditional probability that alternative j is chosen by n in period t in equation (3.12) has the following alternative representation

$$p_{ntj} \equiv p_j(\Psi_{nt}) \equiv E[d^0_{ntj}|\Psi_{nt}], \tag{3.13}$$

and Hotz and Miller (1993) prove the existence of a mapping $\varphi_k : [0,1] \to \Re$ such that

$$\varphi_k(p_k(\Psi_{nt})) = E[\varepsilon_{ntk}|\Psi_{nt}, d^0_{ntk} = 1], \qquad k \in 0,\cdots,3. \tag{3.14}$$

Therefore, the conditional valuation function has the following recursive representation:

$$\begin{aligned} V_{ntj} &= \max_{h_{nt}>0}\Big\{U_{nt0} + U_{nt1} + \alpha^{-1}\eta_n\lambda_t(w_{nt}h_{nt} - a_{nt} - \pi_{nt}) \\ &+\beta\Big[\textstyle\sum_{\Psi^{(1)}_{nt}\in\mathcal{A}^1_{nt}}\Big[\sum_{k=0}^3 p_{nt+1,k}(V_{nt+1,k} + \varphi_k(p_k(\Psi^{(1)}_{nt})))\Big]F_j(\Psi^{(1)}_{nt}|\Psi_{nt})\Big]|d_{ntj} = 1\Big\}. \end{aligned} \tag{3.15}$$

Finally, the optimality conditions for interior solution to labor supply h^0_{nt} (3.10) and study time s^0_{nt} (3.11) are given by

$$\begin{aligned} \frac{\partial U_{nt1}}{\partial h_{nt}} + \frac{\eta_n\lambda_t}{\alpha}w_{nt} = \; & -\beta\Big\{\textstyle\sum_{\Psi^{(1)}_{nt}\in\mathcal{A}^1_{nt}}\Big[\sum_{k=0}^3\Big[p_{ntk+1}\frac{\partial(V_{nt+1,k}+\varphi_k(p_k(\Psi^{(1)}_{nt})))}{\partial h_{nt}} \\ & +\frac{\partial p_{nt+1,k}}{\partial h_{nt}}(V_{nt+1,k} + \varphi_k(p_k(\Psi^{(1)}_{nt})))\Big]F_j(\Psi^{(1)}_{nt}|\Psi_{nt}) \\ & +\textstyle\sum_{k=0}^3 p_{nt+1,k}(V_{nt+1,k} + \varphi_k(p_k(\Psi^{(1)}_{nt})))\frac{\partial F_j(\Psi^{(1)}_{nt}|\Psi_{nt})}{\partial h_{nt}}\Big]|d_{ntj} = 1\Big\}, \text{ and,} \end{aligned} \tag{3.16}$$

$$\begin{aligned} \frac{\partial U_{nt1}}{\partial s_{nt}} = \; & -\beta\Big\{\textstyle\sum_{\Psi^{(1)}_{nt}\in\mathcal{A}^1_{nt}}\Big[\sum_{k=0}^3\Big[p_{ntk+1}\frac{\partial(V_{nt+1,k}+\varphi_k(p_k(\Psi^{(1)}_{nt})))}{\partial s_{nt}} \\ & +\frac{\partial p_{nt+1,k}}{\partial s_{nt}}(V_{nt+1,k} + \varphi_k(p_k(\Psi^{(1)}_{nt})))\Big]F_j(\Psi^{(1)}_{nt}|\Psi_{nt}) \\ & +\textstyle\sum_{k=0}^3 p_{nt+1,k}(V_{nt+1,k} + \varphi_k(p_k(\Psi^{(1)}_{nt})))\frac{\partial F_j(\Psi^{(1)}_{nt}|\Psi_{nt})}{\partial s_{nt}}\Big]|d_{ntj} = 1\Big\}, \end{aligned} \tag{3.17}$$

for $j = 1,3$, and $j = 2,3$ respectively. The first condition in (3.16) says that the net current benefit from an additional hour of work is equal to the present discounted value of future utility costs of that additional hour. The current marginal utility from an additional hour of work is equal to the net of the utility cost of leisure forgone, and the consumption value of the additional goods and services produced. The future value of an additional hour of work is decomposed into three main components. The first term on the RHS captures the direct effect of an increase in hours worked on future productivity and future utitily. Future utility is directly affected because of the assumption that current and future leisure are intertemporally nonseparable. Future productivity is affected by the assumption that current labor force participation enhances human capital, which is reflected in higher future marginal productivity of labor. The second term on the RHS captures the indirect effect on future utility by current hours worked through its effect on future probability of employment. The third term on the RHS accounts for the indirect effect of current hours worked on future utility through its effect on the transition probability. The probability of being promoted a grade level given that the individual is currently enrolled is assumed to be dependent on hours worked.

4 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.[8] 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.

5 Estimation Method

The empirical analysis employs a multi-stage version of the conditional choice probability (CCP) estimator developed in Hotz and Miller (1993) and extended by Altug and Miller (1998) and Gayle and Miller (2003). We outline the estimation strategy of each stage in turn. The parameters of the model can be estimated from the optimality conditions derived in section (3). First, there is contemporaneous separability between consumption and labor supply in the utility function. Given that consumption is measured with error and that the measurement error is uncorrelated with the information set of the individual, the consumption function can be estimated separately from the equations characterizing optimal discrete choice to provide first stage estimates of the of the shadow price of consumption. Similarly, assuming that observed wages are noisy measures of the marginal product of labor, where the measurement error is assumed to be independent of the information set of the individual over time, the parameters of the marginal product of labor can be estimated separately from the other parameters of the model.

Examination of equations (3.15) and (3.10) in section (3) suggest that estimation of the conditional choice probabilities p_{knt} and their derivatives with respect to hours worked h_{nt} and study time s_{nt} are required. These quantities are estimated nonparametrically and substituted into the necessary conditions for optimal choice and hours allocation. The technique employed here also requires that the transition probabilities be estimated. The remaining parameters of the model are estimated by nonlinear GMM, where the moment conditions are formed as

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

sample analogs of equations (3.9), (3.16) and (3.17). Since the first stage regressions are of interest in their own right, we discuss them in separate sections.

6 Consumption

Estimation of the marginal utility of consumption requires further parametrization of the utility of consumption given by equation (3.4). We assume that $g(x_{nt})$ has the following parametrization:

$$g(z_{nt}) = \exp(x'_{nt}B_1), \tag{6.1}$$

The first order necessary conditions for optimal consumption allocation are then given by:

$$\exp(x'_{nt}B_1)c_{nt}^{\alpha-1} = \eta_n\lambda_t. \tag{6.2}$$

The necessary conditions (6.2) and (3.7) provide the key equations for the estimation of the shadow value of consumption λ_t and the individual specific effect η_n. Taking the natural log of equation (6.2) and rearranging results in the following equation

$$ln(c_{nt}) = (1-\alpha)^{-1}x'_{nt}B_1 - (1-\alpha)^{-1}\ln(\eta_n) - (1-\alpha)^{-1}\ln(\lambda_t). \tag{6.3}$$

Assuming that observed consumption $\tilde{c}_{nt}$ is measured with error so that $\tilde{c}_{nt} = c_{nt}e^{\nu_{nt}}$, where c_{nt} is the true level of consumption, and $E[\nu_{nt}|x_{nt}, \eta_n, \lambda_t] = 0$. Let Δ denote the first-difference operator. Taking first difference of equation (6.3) and rearranging, we have that

$$\Delta\nu_{nt} \quad = \quad \Delta\ln(\tilde{c}_{nt}) - (1-\alpha)^{-1}\Delta x'_{nt}B_1 + (1-\alpha)^{-1}\Delta\ln(\lambda_t). \tag{6.4}$$

Equation (6.4) is estimated by the efficient GMM. The estimated results in Table 3 indicate that consumption increases with the size of the family, average family income, and the average age of the family. Consumption decreases with the level of unemployment local to the residence of the individual. Table 3 also suggests that for a given level of education, consumption is increasing and concave in the age of the individual. For a given age of the individual, consumption is decreasing and convex in the level of education.

The first panel of Table 5 reports the estimated log change in aggregate prices with the corresponding standard errors. The graph along with the 95% confidence interval are also presented in Figure 1. These figures show that the changes in aggregate prices are estimated precisely. The figure also show that there are significant variation in the time effects. The simple F-test reject the restriction that $(1-\alpha)^{-1}\ln(\lambda_2) = \cdots = (1-\alpha)^{-1}\ln(\lambda_T)$ at the 99% confidence level.

7 Wages

Assume that the time varying component of the individuals productivity function has the representation:

$$\gamma_j(x_{nt}) \equiv \exp(x'_{nt}B_{2j}). \tag{7.1}$$

Observed wages are assumed to be noisy measures of the marginal productivity of labor, where the multiplicative error term is assumed to be conditionally independent over individuals, the covariates in the wage equation, and the labor supply decision

$$\tilde{w}_{ntj} = \omega_{tj}\mu_n \exp(x'_{nt}B_{2j}) \exp(\varepsilon_{nt}). \tag{7.2}$$

The individual specific effects captures absolute advantage of the individual in the labor market (Willis, 1986). Assume that human capital comes in two types, an unskilled type ($j = 1$) and a skilled type ($j = 2$). The skilled group is defined as having at least 16 years of formal education. All occupations in the economy are sorted across these groups according to the level of education required to carry out the task. Workers are assumed to be perfect substitutes within, but not across efficiency units. Since the model is in the panel data framework, we do not need to assume that schooling and employment choices are independent of the individual's ability as captured by the individual specific effect. This is in contrast to the model proposed in Willis (1986). The absence of this restriction serves to eliminate the problem of sample selection caused by ability bias.

Another key consideration in the estimation of equation (7.2) is whether there is the need to estimate separate models for the different racial groups. The results of Neal and Johnson (1996) and Altonji and Blank (1999) indicate that the large majority of the wage gap between races in the NLSY is due to differences in measures of abilities (AFQT scores) and family background (parents education). Since these measures are time invariant, a suitable transformation of a single wage equation provides accurate estimate in the pooled data.

Taking logs of both sides of equation (7.2) and taking first difference gives the following equation:

$$\Delta\varepsilon_{nt} = \Delta\ln(\tilde{w}_{ntj}) - \Delta\ln(\omega_{tj}) - \Delta x'_{nt}B_{2j} \tag{7.3}$$

Define e_{nt1} to be equal one if individual n is belongs to efficiency unit 1 in period t. Likewise, define e_{nt2} to be equal one if individual n is belongs to efficiency unit 2 in period t. Equation (7.3) is estimated by the efficient GMM. The skill specific coefficients are obtained by interacting the explanatory variables with these indicator variables for each skill group. The skill specific aggregate effects are also obtained by interacting the time dummies with these indicator variables.

The estimated results for the wage equation are reported in Table 4. The positive coefficients on lagged hours indicate that there are positive returns to on the job training. Also, the effect of past hours worked on current wages decline with further lags. The declining magnitude and significance of lagged hours worked is consistent with the conjecture of depreciation in human capital. The returns to on the job training are higher for skilled workers than for unskilled workers. At 2000 hours per year, the wage elasticity of the first lagged hours is 0.04 for low skilled workers and 0.06 for high skilled workers. However, the wage elasticity of the second lagged hours is 0.01 and 0.02. These qualitative results are in line with those found in Miller and Sanders (1997), Altug and Miller (1998) and Gayle and Miller (2003).

The coefficients on the education and experience variables are all estimated highly precisely, with the exception of education squared for low skilled workers.[9] The coefficient of squared education is positive and significant at the 1% level for the high skilled group, indicating nonlinearity in marginal returns to education. We find that the coefficient on the interaction term between education and experience is positive for low skilled workers and negative for high skilled workers, both significant at the 1% level. This suggests that in terms of the productivity of young males, formal education and labor market experience are compliments in the low skilled sector, and substitutes in the high skilled sector.

The flexibility of the specification of the wage equation also allows for some heterogeneity in the returns to education. It allows for comparative advantage with respect to human capital in the labor market to be manifested through differences in patterns of schooling and employment. At first glance marginal return to education for both the skilled and unskilled sector seem very low. Indeed, the calculation would produce a marginal rate of return of 0.024 for low skilled workers and 0.069 for high skilled workers of age 30 in the sample. Table 4 of Card (1999) lists the estimated marginal returns to education found in a number of studies. The marginal returns to education found here are lower than these other estimates. However, these other studies do not account for growth in skill specific aggregate wages. When the average growth in log aggregate wages in included in the calculation, the estimated marginal return to log wages increases to 0.044 for low skilled workers and 0.217 for high skilled workers of age 30. The estimated marginal returns to education in table 4 of Card (1999) all fall within the range.

The last two panels of Table 5 report the estimated changes in unskilled and skilled piece rates. These series are also plotted in Figures 2 and 3 along with their 95% confidence bands. The changes in unskilled piece rates ar less precisely estimated than the changes in skilled piece rates. Two separate hypothesis tests are performed. The first is an F-test of the restriction of equality of all the aggregate effects $\Delta\ln(\omega_{21}) = \cdots = \Delta\ln(\omega_{T1}) = \Delta\ln(\omega_{22}) = \cdots = \Delta\ln(\omega_{T2})$. The second is an F-test of the restriction of a single set of time varying aggregate effects $\Delta\ln(\omega_{21}) = \Delta\ln(\omega_{22}), \cdots, \Delta\ln(\omega_{T1}) = \Delta\ln(\omega_{T2})$. Both restrictions are rejected at the 99% level.

[9]Because most individuals in the sample have no breaks in schooling until they have completed their total level, identification of level of schooling in a first difference model is fragile at best and is excluded from the specification. We exclude the level of experience for the same reason.

7.1 Individual-specific Effects

To estimate preference parameters of the model we need to estimate the individual specific effects η_n and μ_n. They are estimated from the residuals in the log-linear versions of consumption and wage equations (6.3) and (7.3) respectively. These estimators are subject to small sample bias when T is small. However, Hotz et al. (1988) provide Monte Carlo evidence that the small sample bias caused by using such fixed-effects estimates in computing the remaining parameters of interest are quite small for moderate to large sample sizes. Altug and Miller (1998) and Gayle and Miller (2003) estimate the parameters of their structural model under two assumptions on the fixed effects. The first is the traditional definition. The second assumes that fixed effects can be written as functionals of observed covariates. Under the second assumption, consistency of the other parameters of the model is achieved. In their studies, the resulting estimates of the structural parameters were very similar, and lead to the same conclusions. This also indicates that the bias induced by employing estimates of the traditional fixed effects is quite small in these models. The estimates of μ_n and η_n are calculated from samples where $T_1 = 15$ and $T_2 = 12$ respectively. Hahn et al. (2001) suggests that these sample sizes are actually large, implying that the bias of these estimates are expected to be small.

The fixed effects estimators of $\ln(\mu_n)$ and $\ln(\eta_n)$ are obtained as simple time averages of the estimated residuals of the consumption and wage equations.

8 Study Patterns and the Probability of Grade Promotion.

8.1 Study Patterns

In 1981, the NLSY79 collected information on the patterns of school activities of the respondents that are enrolled in school. In particular, the NLSY79 asked the respondent about the amount of hours they spent in school during the week before the interview date. They also asked whether or not the time they reported is typical or not, and if no, to report the typical hours spent in school. The respondents also reported the number of hours they spent studying outside of school during the week before the interview date. These responses are used to construct yearly measures of the time spent by individuals on school activities. We show that one can get reliable estimates of time spent on schooling activities from this data. We call this time spent on schooling activities study time. Clearly this includes not only the time the individual spends actually studying, but also time the student allocates to activities related to school, both during regular hours of school and outside of school.

Assume that the study time of an individual n in period t is an exponential function of observed demographic characteristics and literacy indicators of the individual, as well as unobserved individual-specific characteristics ,

$$s_{nt} \equiv \exp(x'_{nt}B_3).$$

Assume further that observed study time is a noisy measure on the true study patterns of the individual, where the measurement error is assumed to be independent of the regressors.

$$\tilde{s}_{nt} \equiv \exp(x'_{nt}B_3)\exp(\varepsilon_{nt}). \tag{8.1}$$

Under these assumptions, we can consistently estimate the study time of individuals enrolled in school using OLS on the log-linearized version of equation (8.1).

To estimate the preference parameters of the model, we need a consistent estimate of study time given that an individual has enrolled in school. Thus the issue of sample selection bias does not affect the estimation of equation (8.1). Another consideration is the fact that individuals were questioned about their study patterns for only one week prior to the interview period. If the interview is taken at a time where there are generally academic deadlines such as exams, then the reported time spent studying may be overstated. However, interviews were administered to different individuals at different times of the year. This makes plausible the assumption that on average, one does not expect to observe over nor under reporting of study time in the data.

Table 6 reports the regression of the time spent on school activities. The number of observations in estimation is 2253. All variables included in the specification are significant at the 5% level. The F-statistic for the model is 20.47, and the Adjusted R^2 is 11.24%. These statistics show that the instruments do well, both individually and as a group, in capturing variation in log study time. In particular there is no problem of weak instruments in this estimation of study time. This issue of weak instruments is important since the predicted values of study time serve as first stage estimates in all the estimators that follow.

The results in Table 6 show that lagged enrollment decisions are positively associated with study time, with further lags becoming less important. The size of the coefficients indicate also that lagged enrollments decisions are also quite relevant in explaining current study time. Lagged hours of work are negatively correlated with current study time, with diminishing impact for further lags. The magnitude of these effects are also considerable. Individuals with higher AFQT scores spend more time on school activities. Since the AFQT test was administered in 1980 and the data on schooling activities were collected in 1981, there is no issue of feedback effects of current study time on AFQT scores. The results also indicate that the time spent on schooling activities is approximately 11% higher for blacks and 10% lower for hispanics compared to time spent by whites. These differences are quite large, working out to be approximately 154 more hours per year for blacks and 140 less hours per year for hispanics at an average of 1400 hours, approximately what is in the sample.

8.2 The 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 it own right, and is also a key ingredient in the final stage estimation. Assume that this probability takes the logit form:

$$F(x_{nt}) \equiv (1-d_{nt}^{h})\frac{\exp(x_{nt}'B_{41})}{1+\exp(x_{nt}'B_{41})}+d_{nt}^{h}\frac{\exp(x_{nt}'B_{42})}{1+\exp(x_{nt}'B_{42})}. \tag{8.2}$$

Similar to the study time regression. What is needed for consistent estimates of the preference parameters of the model is a consistent estimate of the probability of grade promotion given enrollment. Estimation of equation (8.2) provides us with this. In principle, if the enrollment decision is correlated with the error term defining equation (8.2), then the coefficient estimates obtained would be biased and inconsistent and not conducive to direct interpretation. However, the inclusion of AFQT in the regression should at least mitigate the level of biased induced by regressing only on the subset of individuals that choose to enroll.

Another issue is the choice of separate regressions for the set of students who choose to work while enrolled in school and the set who choose not to work while enrolled in school. This main reason for this specification is to improve the flexibility of the resulting estimated transition probabilities. However, if the decision to work is correlated with the error term that defines equation (8.2), then the coefficient estimates are expected to be biased and inconsistent. The inclusion of our measure of labor market ability, the estimated fixed effects from the wage regression are included to reduce the bias of the estimated coefficients. At the very least however, the coefficients in equation (8.2) can certainly be interpreted for the relevant groups of individuals.

A third issue involves the appropriateness of including current period decision variables in equation (8.2). The theoretical model assumes that the individual makes his schooling and employment decisions $(d_{nt}^{s}, s_{nt}, d_{nt}^{h}, h_{nt})$ at the beginning of each period conditioned on the information set available to him at that point in time. The grade promotion probability function is known by the individual, and he has control over it in so far as he has control over the decision variables. However, the uncertainty is not resolved until the beginning of the following year. The timing of the model thus makes the period t decision variables predetermined in equation (8.2).

Table 7 reports the result of the logit regression of the probability of completing a grade and Table 7.1 reports the corresponding average derivatives. The standard errors reported are corrected for the inclusion of predicted study time. Computation of the corrected standard errors is complicated by the nonlinear specification of the study time function and the probability of grade transformation. The details are presented in Appendix 2 for completeness. The number of observations used in estimation for the two groups ($d_{nt}^{h}=0$, and $d_{nt}^{h}=1$) are 2216 and 5606, the Likelihood ratio statistics are 400.65 and 1350.78, and the Pseudo R^2's are 15% and 17%. Furthermore, all coefficients except for the constant term are significant at the 10% level, and slope parameters, except for 2 are significant at the 5% level. Note that some variables are dropped from estimation in either groups because of their low precision and statistical irrelevance.

The results in Table 7 indicate that lagged labor market participation decisions are positively correlated with the probability of grade promotion. This provides evidence for the congruence hypotheses. However, the effect is a lagged effect, and the interpretation varies slightly from that proposed by D'Amico (1984). The decision to participate in the labor market in either of the last two periods increases the current probability of grade promotion by approximately 5%. The full model will have to simulated to see exactly how large this effect turns out to be on completed education. However, at this stage it is clear that a 5% increase in the probability of completing a grade level is a significant magnitude.

We find that blacks have a lower probability of being promoted a grade level than their white counterparts. For the group that works, hispanics also have a lower probability of being promoted than their white counterparts. This result is not simply the classical drop-out story of minorities. The interpretation of these coefficients are that: given two males, one black and the other white, with the same abilities (as measured by AFQT scores and the estimated fixed effects), the same hours studied, the same hours worked, and in the same grade level, along with other conditioned covariates, the black male has a significantly lower probability of being promoted from that grade level. To understand what may be driving this result, one must also look at what in not included in the regression, that is, what factors are not controlled for and may be correlated with race. The primary excluded factor in the regression would be the quality of the schools attended. It is well known that the quality of schools attended by blacks are on average lower than those attended by their white counterparts. I argue therefore that the negative coefficient of blacks in the grade advancement regression captures the lower schooling opportunities and qualities available to these racial groups. The quality of schooling is typically measured by, among other factors, the level of funding that school receives, class size, in particular the student-teacher ratio, and the socio-economics conditions of the community surrounding the school. The available data does not contain information on these measures of school quality. However, if one is only interested in the difference in schooling opportunities across races, as this study is, and not to identify the sources of these differences, then the estimated regression is sufficient.

The results in table 7 also indicate that the probability of grade level promotion is increasing in time spent on schooling activities for both groups, and concave for the group that works. Conversely, this probability is decreasing and convex in hours spent in the labor market. Students in grades 11 and 12 have a larger probability of being promoted than college students.

9 Conditional Choice Probabilities

Estimation of conditions characterizing labor supply and schooling decisions also requires that estimates of the conditional choice probabilities defined in equation (3.12). Inclusion of the individual-specific effect, and time-specific effects as explanatory variables allows us to treat the sample as pooled cross-section and time series data that is independently distributed over individual and time. This implies straightforward nonparametric estimation of (3.13).

To estimate the preference parameters, we also need to estimate the conditional choice probabilities conditional on all the states that remain feasible. This is done by taking advantage of the finite state dependence of the model. In particular, we need to estimate the probability that individual n chooses alternative j in period $t+i$ conditional on observing state k in that period $p_j(\Psi_{ntk}^{(i)})$. We achieve this by estimating the probability that an observationally equivalent individual chooses alternative j in the current period conditional on observing the state k in the current period. The validity of this method depend on the inclusion of the individual-specific effects and the time-specific effects in these regressions. These auxiliary CCP's are estimated using nonparametric techniques. The technical details of these estimators are outlined in Appendix C.

Table 7 presents the means and standard deviations of these estimated probabilities and the required derivatives. The sample average of the CCP's are equal to the sample average of their corresponding indicator functions with 4 decimal places. This indicates that the bias in these estimates are small. The relative magnitudes of the conditional state probabilities are also plausible. The probability that an individual chooses home production given that he enrolled in school last period and did not get promoted the grade level is larger than the probability of choosing home production if he was promoted.

The average derivatives of the conditional state probabilities are also empirically plausible. An additional hour of work in the past reduces the probability that the individual will choose home production in the current period. An additional hour of school activity in the past increases the probability of choosing home production in the current period if the individual did not get promoted the grade level. On the other hand, an additional hour of school activity in the past decreases the probability of choosing home production in the current period if the individual was promoted the grade level.

10 Schooling, Participation, and Hours

10.1 The moment conditions.

Estimation of the remaining parameters of the model makes use of an alternative representation of the conditional valuation function derived in Hotz and Miller (1993). This requires that parametric restrictions be placed on the utility functions. Let the components of the utility of schooling, labor supply in equation (2.6), and utility of leisure in equation (2.7) take the form

$$u_1(x_{nt}, d^s_{nt}) = d^s_{nt} x'_{nt} B_5, \tag{10.1}$$

$$u_2(x_{nt}, d^h_{nt}) = d^h_{nt} x'_{nt} B_6, \tag{10.2}$$

$$u_3(x_{nt}, g_{nt}) = l_{nt} z'_{nt} B_7 + \sum_{i=0}^{\rho} \delta_i l_{nt} l_{nt-i}. \tag{10.3}$$

The utility of leisure is assumed to be quadratic. Economic theory suggests that the utility of leisure is concave in leisure, $\delta_0 < 0$. The parameters $\delta_i, i = 1, \cdots, \rho$ capture intertemporal nonseparabilities in the preference for leisure. For $i > 0$, $\delta_i < 0$ implies that current leisure and leisure lagged i periods are intertemporal substitutes. On the other hand, $\delta_i > 0$ implies that current leisure and leisure lagged i periods are intertemporal complements.

Define $\theta \equiv (B'_5, B'_6, B'_7, \delta_0, \cdots, \delta_\rho, \alpha)'$, $\gamma \equiv (B'_1, \cdots, B'_4)'$, $P \equiv (P_{nt0}, \cdots, P_{nt3})'$. Let F denote the set of conditional state probabilities and their relevant derivatives and let $\Theta \equiv (\theta', \gamma', P', F')'$. Define also $l^{(0)}_{nt} \equiv 1$, $l^{(1)}_{nt} \equiv 1 - h_{nt}$, $l^{(2)}_{nt} \equiv 1 - s_{nt}$, and $l^{(3)}_{nt} \equiv 1 - h_{nt} - s_{nt}$. By substituting these functional forms for the utility functions into the Euler condition for hours (3.16), we derive the following moment condition:[10]

$$\begin{aligned} m_{nt1}(\Theta) \equiv \; & d_{nt1}\Big[\alpha^{-1}\eta_n\lambda_t w_{nt} - z'_{nt}B_5 - 2\delta_0 l^{(1)}_{nt} - \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt-i} + \beta^i) \\ & - \textstyle\sum_{i=1}^{\rho}\beta^i p_0(\Psi^{(i)}_{nt1})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt1})}{\partial h_{nt}}\Big] \\ & + d_{nt3}\Big[\alpha^{-1}\eta_n\lambda_t w_{nt} - z'_{nt}B_5 - 2\delta_0 l^{(3)}_{nt} - \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt-i} + \beta^i) \\ & - \textstyle\sum_{i=1}^{\rho}\beta^i\Big[p_0(\Psi^{(i)}_{nt4})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt4})}{\partial h_{nt}}F(x_{nt}) + p_0(\Psi^{(i)}_{nt5})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt5})}{\partial h_{nt}}(1 - F(x_{nt})) \\ & + \ln\left(\frac{p_0(\Psi^{(i)}_{nt5})}{p_0(\Psi^{(i)}_{nt4})}\right)\frac{\partial F(x_{nt})}{\partial h_{nt}}\Big]\Big]. \end{aligned}$$

Likewise, we substitute the utility functions in to the optimality condition for study time (3.17) to obtain the following moment condition:

$$\begin{aligned} m_{nt2}(\Theta) \equiv \; & d_{nt2}\Big[-z'_{nt}B_5 - 2\delta_0 l^{(2)}_{nt} - \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt-i} + \beta^i) \\ & - \textstyle\sum_{i=1}^{\rho}\beta^i\Big[p_0(\Psi^{(i)}_{nt2})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt2})}{\partial s_{nt}}F(x_{nt}) + p_0(\Psi^{(i)}_{nt3})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt3})}{\partial s_{nt}}(1 - F(x_{nt})) \\ & + \ln\left(\frac{p_0(\Psi^{(i)}_{nt3})}{p_0(\Psi^{(i)}_{nt2})}\right)\frac{\partial F(x_{nt})}{\partial s_{nt}}\Big]\Big] + d_{nt3}\Big[-z'_{nt}B_5 - 2\delta_0 l^{(3)}_{nt} - \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt-i} + \beta^i) \\ & - \textstyle\sum_{i=1}^{\rho}\beta^i\Big[p_0(\Psi^{(i)}_{nt4})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt4})}{\partial s_{nt}}F(x_{nt}) + p_0(\Psi^{(i)}_{nt5})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt5})}{\partial s_{nt}}(1 - F(x_{nt})) \\ & + \ln\left(\frac{p_0(\Psi^{(i)}_{nt5})}{p_0(\Psi^{(i)}_{nt4})}\right)\frac{\partial F(x_{nt})}{\partial s_{nt}}\Big]\Big]. \end{aligned}$$

[10] A more detailed derivation of the following moment conditions is found in Appendix D.

Additional moment conditions are formed from the optimal discrete choice conditions in equation (3.9). In particular, we obtain the following moment conditions from the optimality condition for choosing alternatives 1,2, and 3:

$$\begin{aligned} m_{nt3}(\Theta) \equiv\ & d_{nt1}\Big[\ln\left(\tfrac{p_{nt1}}{p_{nt0}}\right) - x'_{nt}B_6 + x'_{nt}B_7(l_{nt}^{(0)} - l_{nt}^{(1)}) + \delta_0(l_{nt}^{(0)2} - l_{nt}^{(1)2}) \\ & + \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt}^{(0)} - l_{nt}^{(1)})(l_{nt-i} + \beta^i) - \tfrac{\eta_n\lambda_t}{\alpha}(w_{nt}h_{nt}) - \textstyle\sum_{i=1}^{\rho}\beta^i \ln\left(\tfrac{p_0(\Psi_{nt0}^{(s)})}{p_0(\Psi_{nt1}^{(s)})}\right)\Big], \end{aligned}$$

$$\begin{aligned} m_{nt4}(\Theta) \equiv\ & d_{nt2}\Big[\ln\left(\tfrac{p_{nt2}}{p_{nt0}}\right) - x'_{nt}B_5 + x'_{nt}B_7(l_{nt}^{(0)} - l_{nt}^{(2)}) \\ & + \delta_0(l_{nt}^{(0)2} - l_{nt}^{(2)2}) + \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt}^{(0)} - l_{nt}^{(2)})(l_{nt-i} + \beta^i) + \tfrac{\eta_n\lambda_t}{\alpha}\pi_{nt} \\ & - \textstyle\sum_{i=1}^{\rho}\beta^i\left[\ln p_0(\Psi_{nt0}^{(i)}) - \ln p_0(\Psi_{nt2}^{(i)})F(x_{nt}) - \ln p_0(\Psi_{nt3}^{(i)})(1 - F(x_{nt}))\right]\Big], \end{aligned}$$

$$\begin{aligned} m_{nt5}(\Theta) \equiv\ & d_{nt3}\Big[\ln\left(\tfrac{p_{nt3}}{p_{nt0}}\right) - x'_{nt}B_5 - x'_{nt}B_6 + x'_{nt}B_7(l_{nt}^{(0)} - l_{nt}^{(3)}) \\ & + \delta_0(l_{nt}^{(0)2} - l_{nt}^{(3)2}) + \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt}^{(0)} - l_{nt}^{(3)})(l_{nt-i} + \beta^i) - \tfrac{\eta_n\lambda_t}{\alpha}(w_{nt}h_{nt} - \pi_{nt}) \\ & - \textstyle\sum_{i=1}^{\rho}\beta^i\left[\ln p_0(\Psi_{nt0}^{(i)}) - \ln p_0(\Psi_{nt4}^{(i)})F(x_{nt}) - \ln p_0(\Psi_{nt5}^{(i)})(1 - F(x_{nt}))\right]\Big]. \end{aligned}$$

Define $m_{nt}(\Theta) \equiv (m_{nt1}(\Theta), \cdots, m_{nt5}(\Theta))'$ and let $\overline{T}$ denote the set of periods for which the working and schooling hours, enrollment and participation conditions are valid. Let $m_n \equiv (m'_{n1}, \cdots, m'_{n\overline{T}}$ denote the vector of empirical moments for a given individual over time. We further define the weighting matrix $\Omega \equiv E[m_n, m'_n]$ and note that this matrix is block diagonal since $E_t[m_{nt}m_{ns}] = 0$ for $s < t$.

In order to increase the finite sample precision of the estimates of the remaining parameters of the model, we implement a iterated GMM (GMMI) variation of the Nested Pseudo Likelihood Algorithm (NPL) proposed by Aguirregabiria and Mira (2002). This algorithm consists of two steps. The first step is where GMMI is implemented to obtain estimates of the preference parameters, give an initial estimated of the CCP's. The second step is where the CCP's are updated using the estimates of the preference parameters. To be precise, define $\Theta_1^k \equiv (\theta', \hat{\gamma}', (P^k)', \hat{F}')'$, and $\Theta_2^k \equiv ((\theta^k)', \hat{\gamma}', P', \hat{F}')'$. At iteration $K \geq 1$ of the outer algorithm, we apply the following steps

Step 1: Obtain new estimates of θ, θ^K, from the following iteration in $j \geq 1$:

$$\theta^{Kj} = \arg\max_{\theta\in\Theta} \sum_{n=1}^{N}\left[m_n(\Theta_1^{K-1})\right]'(\Omega^{j-1})^{-1}\left[m_n(\Theta_1^{K-1})\right], \tag{10.4}$$

where Ω^{j-1} is the weighting matrix evaluated at Θ_1^{K-1}, in which $\theta = \theta^{K,j-1}$. This iteration is repeated until convergence in θ is achieved, which is denoted θ^K

Step 2: Update P using the estimates θ^K as follows:

$$\begin{aligned} P_j^K &= \exp\left(V_j(\Theta_2^K) - V_0(\Theta_2^K)\right)P_0^{K-1} \\ &= \exp\left(m_{j+2}(\Theta_2^K)\right)P_0^{K-1}, \qquad j \geq 1, \\ P_0^K &= 1 - \textstyle\sum_{j=1}^{J} P_j^K. \end{aligned} \tag{10.5}$$

Iterate in K until convergence in P and θ is reached.

The convergence of the CCP's is stated in Proposition 1 of Aguirregabiria and Mira (2002), while the convergence of the GMMI is discussed in Hansen et al. (1996). From our experience, it seems that the iteration in step 1 of the algorithm improves greatly the stability of the overall algorithm.

The nature of the iteration in the CCP's along with the inclusion of the pre-estimates $(\hat{\gamma},\hat{F})'$ make the correct standard errors of the estimates of θ nonstandard. To derive the correct standard errors, we implement the technique proposed in Newey and McFadden (1994) and Newey (1994). Interestingly, because of the structure of the state space in the model, repeated use of the law of iterated expectations results in significant simplification of the asymptotic variance. In particular, no post estimation is required to correct the standard errors. This greatly reduces the computational burden of the CCP estimator. The key effect of the iteration in the CCP's is an alternative specific re-weighting of the influence functions of the pre-estimators. This re-weighting is such that a larger weight is assigned to alternatives with a higher probability of occurring. The asymptotic properties of this estimator are discussed in appendix E.

10.2 Consumption Value of School Attendance

Table 9 reports the estimated psychic value of enrollment. The results indicate that the consumption value of schooling is increasing and concave in the level of education. For a given age, the consumption value is decreasing in level of education. These signs capture the decreasing rate of enrollment in school for higher levels of education and older individuals.

The coefficients on *BLACK* and *HISPANIC* in the consumption value of schooling are positive by not significantly different from zero. This result holds with and without the inclusion of AFQT. This implies that after controlling for racial differences in wages, hours worked, time spent of schooling, and school quality, black and Hispanic males are no more likely to enroll in school than their white counterparts.

10.3 Fixed Utility of Participation

Table 10 presents the estimate fixed utility of participating in the labor market. We find that the consumption value of labor force participation is increasing and concave in the level of labor market experience. However, these coefficients are imprecisely estimate. We find also that for a given age, the consumption value of labor force participation is decreasing in the level of labor market experience. The coefficients on *BLACK* and *HISPANIC* in the consumption value of labor force participation are negative, but imprecisely estimated. This results the racial disparity in the employment rates is not explained by differences in the propensity of participate in the labor market.

10.4 Utility of Leisure

The estimates of the utility of leisure are reported in Table 11. The results indicate that the utility of leisure is (weakly) decreasing and convex in age. This results is also found in Altug and Miller (1998) and Gayle and Miller (2003). The results also indicate that the utility of leisure in increasing and concave in leisure. However, the parameter capturing the concavity is imprecisely estimated. We find also that the coefficients on the black and Hispanic indictors are not statistically different from zero. In other words we find no evidence of racial differences in the utility of leisure. In other words, the observed racial differences in hours worked and study time are not explained by racial differences in the preferences for leisure.

10.5 Intertemporal Nonseparabilities in Leisure

The results in table 11 indicate that preferences are intertemporally nonseparable in leisure. The positive coefficients on the interaction between current and lagged leisure in the utility of leisure indicate that for males in the sample, current and future leisure are complements in intertemporal preferences. This indicates a habit formation pattern where increases in current hours worked decreases the future marginal disutility of work. Likewise, increases in current hours spent on school activities decreases the future marginal disutility of studying.

Intertemporal nonseparabilities in leisure is estimated in, among others, Eckstein and Wolpin (1989), Miller and Sanders (1997), Altug and Miller (1998), and Gayle and Miller (2003). The results concerning the intertemporal substitutability of complimentarity of leisure varies across these studies. Altug and Miller (1998) conjecture that employing data sampled over shorter time intervals result in the finding of complementarity between current and past leisure choices, while data sampled over longer (yearly) intervals result in the finding of substitutability between current and past leisure. However, the results in table 11 run in contrast to this conjecture, since in this study, hours are measured annually.

11 Solution and Simulation Exercises

11.1 Solving the model

Given the estimated parameters, the model is solved by means of backward induction from age 65 to age 15. Ideally, we would like to treat hours worked and studied completely symmetrical, as done in the estimation. However, solving for both hours worked and studied on a fine enough grid is infeasible. To bypass this problem, we use the estimated function for study time to approximate optimal study time in the solution. This approximation makes solution of the model tractable. However, this function is valid only for males that choose to enroll.

While this was not a problem for estimating the model, it may cause biases in the simulation results.

With the use of the study time function, optimal hours can then be solved for on a fine grid. The problem of interpolating off this grid then arises. Interpolation is carried out by a third order polynomial regression of the value at each point of the grid on the corresponding state space. The parametric regression is preferred over nonparametric kernel techniques because it allows for a finer grid on hours and avoids the corresponding curse of dimensionality that nonparametric techniques face. In solving the baseline model, the smallest R^2 at age 40 is 0.994, indicating that the third order polynomial approximation is expected to provide very precise approximations of the value functions off the grid of hours. We also assume in the solution that nobody enrolls in school after the age of 36. This is justified as in the data only a very small fraction of the sample enrolls in school past the age of 36.

The baseline model is solved assuming that the economy is in equilibrium where aggregate components grow at an equilibrium rate of the average in the sample period. These aggregate components are the shadow price of consumption, the skill specific piece rates, and tuition costs. The assumption of zero growth rate in aggregate skill prices would result in unrealistic predictions of wages over the life cycle. The baseline model is solved for 10,000 replications separately for whites, blacks, and Hispanics. Table 12 reports the baseline simulation by age along with the corresponding sample averages from the data. The baseline model under-predicts the level of labor market experience and the average hourly wage rate. It may be possible to improve the fit of the model to the data by adding dummies to capture the large drop-off in enrollment and increase in working of 18 and 19 year old males that is found in the data. However, there is no economic intuition for such dummies, and they are not necessary for the analysis to come. Furthermore, given that we we do not have the full profile of the growth in the aggregate variables, the simulation results are not expected to closely fit the sample averages at any rate. Not withstanding this, the model predicts remarkably well the general patterns within each race group. Moreover, the model also gets exactly the relative patterns in the reported outcomes across races.

The first two counterfactual simulations performed evaluate policies that are aimed at affecting working while enrolled in school. First the government subsidizes individuals who choose to enroll in school and not participate in the labor market. Second the government increases the school curriculum so that individuals who enroll in school necessarily spend more time on school activities. The Third set of counterfactual simulations addresses the issue of equating the quality of schooling across races. The final set addresses the issue an increase in time spent on school activities when school quality is held constant across races.

11.2 Cash Subsidy

For the first counterfactual simulation exercise, we consider a subsidy of 1000 dollars, which grows yearly at the same rate as the aggregate component of the marginal utility of consumption (which is the same as the growth rate in tuition). The results from this simulation exercise are reported in table 13 under the column labeled “Pol. 1”. The baseline simulation results are included for comparison under the column labeled “Base”.

The results indicate that this policy does very little in affecting the outcomes of young men. We see very modest increases in education, and reductions in experience. There are also modest overall increases in wages due to this policy. The effect of the policy is the same for all races.

11.3 Increased time spent on school activities

In practice, the second policy can be achieved by increasing the number of hours school is in session for, summer classes, or Saturday (or Sunday) classes. This can also be achieved my increasing the number of, or level of difficulty of homework assignments and projects. In the simulation exercise, this policy is achieved by increasing the study time function. The amount by which the constant is increased is chosen to make the magnitudes of this policy and the subsidy policy above comparable. In particular, if at age 16, the individual was to work for $1000 at $4 hourly wage rate, he would work for 250 hours. The study time function is therefore increased by 250. Since the average wage at age 16 in the baseline simulations is approximately $3.50, the results from this simulation are considered to be lower bound comparisons to the above simulation exercise. The findings are reported in table 14 under the columns labeled “Pol 2”.

The findings indicate that this policy significantly increases education and wages for white and black men, with moderate increases for Hispanics. By the age 35, the completed level of education increases by 15% for whites and 12.3% for blacks, but only by 1% for Hispanics. Also we find that the level of labor market experience for whites and black decreases as a result of the policy, while it increases for Hispanics.

Analysis of the change in the choices young men make due to the policy shows that Hispanics are the least responsive. Further more, while the fraction of the population that enroll in school and not work increase significantly for whites and black (21.8% and 16.4%), it increases only modestly for Hispanics (1%). Another difference in the patterns of choices is that while the faction of the white population that works and attends school decreases (by 5%), it increases for blacks (0.5%) and Hispanics (1.8%). Furthermore, Hispanics are the only group in which the decline in young men where the percentage increase in those working and attending school outweighs the percentage decline in those who choose to exclusively participate in the labor market.

The conclusion therefore is that the crowding-out hypothesis holds most significantly for whites, followed by blacks and Hispanics. This conclusion comes from the fact that a mandatory increase in the time spent on school activities has the most significant negative effect on the employment rate of whites, and the most significant positive effect on completed education and future wages of whites. This result is also empirically bolstered by the fact that in the data a larger fraction of whites enroll in school and work at the same time. Hence intuitively, one would expect that they may be most subject to the crowding out effect of working while attending school. Hence policies that are aimed at increasing the time students spend on school activities has significant positive effects on whites and blacks, but less so on Hispanics.

11.4 Equating school quality

The next policy experiment equalizes the quality of schools across races. Technically, this is done by setting the coefficients of *BLACK* and *HISPANIC* in the grade transition probability equation to zero. The results from this exercise are presented in table 15 under the columns labeled “Pol 3”. We also present the results from the baseline simulation under the columns labeled “Base”.

The results in table 15 indicate that the policy has significant impacts on both blacks and Hispanics. For blacks, by the age of 35, the completed level of education increases by 11%, the years of labor market experience increase by 1%, and the hourly wage rate increases by 15%. For Hispanics, by the age of 35, the completed level of education increases by 7%, the years of labor market experience increase by 3%, and the hourly wage rate increases by 4%.

For both blacks and Hispanics, the policy has the effect of increasing enrollment rates. However, the pattern of enrollment is quite different for both groups. For blacks, the policy has an effect of increasing the fraction of those who enroll exclusively in school by 12%, and 13% for those who enroll and work. For Hispanics however, the policy only increases the fraction of those who enroll exclusively in school by 2%, but by 14% for those who enroll and work. Since the chances of completing a grade level is smaller if the student is also working, this results in a more modest increase in completed education, and thus a more modest increase in hourly wage rate.

We conclude therefore that policies aimed at improving the quality of schools for minorities results in significantly increased education for both groups, but a more modest increase in hourly wage rates for Hispanics.

11.5 Equating school quality and increasing time spent on school activities

Given that equating school quality results in a significant increase in the education level of Hispanics, it is interesting to know if the magnitude of the effect of an increase in study time changes in magnitude under this new environment. Therefore we simulate this environment and the results are reported in table 16 under the columns labeled “Pol 4”. Again, the baseline simulation results are presented for comparison under the columns labeled “Base”.

The results under the new environment, the choices and outcomes for Hispanics are far more responsive to the exogenous increase in study time. The simulated completed level of education increases by 23% and the hourly wage rate increases by 29% for Hispanics by age 35. Furthermore, the fraction of the Hispanic population that exclusively enroll in school increase by 23% and the fraction that enroll and work increase by 18%. Thus under the environment where the quality of schools are equated across race, the responsiveness of Hispanics to an exogenous increase in study time increases significantly.

For blacks in this new environment, the exogenous increase in study time increases the completed level of education by 28% and the hourly wage rate by 79% by age 35. The fraction on blacks that enroll in school exclusively increases by 38%, and the fraction that enrolls and work increases by 8%.

These results indicate that policies aimed that increasing the time spent on school activities has a positive effect on minority students; magnitudes that are comparable to their white counterparts.

12 Conclusions

The paper has developed and estimated a dynamic structural model of educational attainment and labor supply. The main focus of the analysis has been to study the allocation of time between labor supply, formal schooling activities and leisure, both within a year and over the life cycle. The model allows for skill specific productivity and piece rates, as well as intertemporal nonseparabilities in the utility of leisure. It also allows for racial variation in wages, consumption, school quality, study patterns, the fixed cost of labor market participation, the fixed utility of schooling, and the utility of leisure. The estimated results indicate that current and future leisure choices are intertemporal complements. The results also indicate that the observed racial differences in outcomes come from a variety of sources that interact in a highly nonlinear fashion, but not from racial differences in tastes.

The estimated model is used to evaluate two policies that are aimed at affecting the allocation of time between schooling and working. The first policy subsidizes young students that do not participate in the labor market. The results indicate that this subsidy does little in changing the patterns of enrollment and labor supply on either the extensive or the intensive side. The second policy increases the school curriculum so that young men who choose to enroll in school necessarily spend more time on schooling activities. The results indicate that this policy would have significant positive effects on white and blacks, but more modest effects on Hispanics.

A third exercise was performed to evaluate the effects of equating school qualities of blacks and Hispanics to that of whites. The results indicate that such a policy would have a large positive effect on education and wages for blacks, but a smaller positive effect on Hispanics. We also show that under this environment, Hispanics become significantly more responsive to policies aimed at increasing the school curriculum.

This study was motivated by the increasing number of students that decide to also participate in the labor market. The results indicate that the effect of this trend varies across races. Policy focused on changing this trend to improve the level of education and labor market outcomes may have only modest effects on some racial groups. As a matter of policy, the results indicate that equating school quality across races may be a more productive first step for improving the outcomes of minorities. Of course, our measure of school quality is agnostic about exactly what are the parameters in the school system that needs to be addressed. This would require an understanding of the key variables that affect students' grade promotion probabilities.

One of the main limitations of the model presented in this paper is that it is set in a partial equilibrium framework. In a general equilibrium framework, one would expect that the aggregate skill specific wages will also be affected by a policy that changes the distribution of the labor force over these groups. A policy that increases the level of education will result in more labor supplied to the high skilled sector and less to the low skilled sector. In a general equilibrium framework, this will drive down the price of high skilled labor and push up the price of low skilled labor, thus reducing the incentive to acquire higher education. Since this general equilibrium effect is not accounted for in the model presented in this paper, the effects of policies that increase the level of education may be overstated. How far the partial equilibrium effects are from the general equilibrium effects is an important issue for future research.

References

ACEMOGLU, D. AND J. ANGRIST (2000): "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws," *NBER Working Paper*.

ACEMOGLU, D. AND J.-S. PISCHKE (2001): "Changes in the Wage Structure, Family Income, and Children's Education," *European Economics Review*, 45, 890–904.

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.

AHITUV, A. AND M. TIENDA (2000): "Employment, Motherhood and School Continuation Decisions of Young White, Black and Hispanic Women," *MIMEO, The Hebrew University fo Jerusilem*.

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

ALTUG, S. AND R. A. MILLER (1990): "Household Choices in Equilibrium," *Econometrica*, 58, 543–570.

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

ARCIDIACONO, P. (2002): "Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings," *unpublished manuscript, Duke University*.

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:* $http://www.econ.ucla.edu/hotz/working_papers/cohort.pdf$.

BARRON, J. M., B. T. EWING, AND G. R. WADDELL (2000): "The Effects of High School Athletic Participation of Education Labor Market Outcomes," *The Review of Economics and Statistics*, 83, 409–421.

BECKER, G. S. (1964): *Human Capital*, The University of Chicago Press, 3^{rd} ed.

BELZIL, C. AND J. HANSEN (2000): "Intergenerational Transfers and the Rate of Time Preference in a Dynamic Model of Schooling," *Unpublished Manuscript, Concordia University*.

——— (2002a): "A Structural Analysis of the Correlated Random Coefficient Wage Regression Model," *Unpublished Manuscript, Concordia University*.

——— (2002b): "Unovserved Ability and the Return to Schooling," *Unpublished Manuscript, Concordia University*.

BEN-PORATH, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *The Journal of Political Economy*, 75, 352–365.

BLINDER, A. S. AND Y. WEISS (1976): "Human Capital and Labor Supply: A Synthesis," *The Journal of Policitical Economy*, 84, 449–472.

CAMERON, S. AND J. J. HECKMAN (1994): *Determinants of Young Males' Schooling and Training Choices*, vol. I, University of Chicago Press.

CAMERON, S. V. AND J. J. HECKMAN (1998): "Life Cycle Schooling and Dynamic Seclectino Bias: Models and Evidence for Five Cohorts of American Males," *NBER Working Paper Series*.

——— (1999): "The Dynamics of Educational Attainment for Blacks Hispanics, and Whites," *NBER Working Paper Series*.

CARD, D. (1999): *The Causal Effect of Education on Earnings*, Elsevier Science Publishers.

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

ECKSTEIN, Z. AND K. WOLPIN (1989): "Dynamic Labor Force Participation of Married Women and Endogenous Work Experience," *The Review of Economic Studies*, 56, 375–590.

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.

EICHENBAUM, M. S., H. L. PETER, AND K. J. SINGLETON (1988): "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty," *The Quarterly Journal of Economics*, 103, 51–78.

ERMISCH, J. AND M. FRANCESCONI (2000): "Educational Choice, Families and Young Peoples Earnings," *Journal of Human Resources*, 35, 143–176.

GAYLE, G.-L. AND R. A. MILLER (2003): "Life-Cycle Fertility and Human Capital Accumulation," *Unpublished Manuscript, Carnegie Mellon University*.

GAYLE, W.-R. (2003): "School Year Employment, Academic Performance and College Enrollment Rates," *Unpublished Manuscript, University of Pittsburgh*.

HAHN, J., J. HOUSEMAN, AND G. KUERSTEINER (2001): "Bias Corrected Instrumental Variables Estimation for Dynamic Panel Models with Fixed Effects," *Unpublished Manuscript, http://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=145354*.

HANSEN, L., J. HEATON, AND A. YARON (1996): "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business and Economics Statistics*, 14, 262–280.

HECKMAN, J. J. (1976): "A Life-Cycle Model of Earnings, Learning, and Consumption," *The Journal of Political Economy*, 84, S11–S44.

——— (1979): "Sample Selection Biases as a Specification Error," *Econometrica*, 47, 153–161.

——— (1981a): *Statistical Models of Discrete Panel Data*, The MIT Press.

——— (1981b): *The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process*, The MIT Press.

——— (2001): "Identifying The Role of Cognitive Ability in Explaining The Level Of and Change in Returns To Schooling," *The Review of Economics and Statistics*, 83, 1–12.

HOTZ, J. V., F. E. KYDLAND, AND G. SEDLACEK (1988): "Intertemporal Preferences and Labor Supply," *Econometrica*, 56, 355–360.

HOTZ, J. V. AND R. A. MILLER (1988): "An Empirical Analysis of Life Cycle Fertility and Female Labor Supply," *Econometrica*, 56, 91–118.

——— (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529.

HOTZ, J. V., R. A. MILLER, S. SANDERS, AND J. SMITH (????): "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, 61.

HOTZ, J. V., L. C. XU, M. TIENDA, AND A. AHITUV (2002): "Are There Returns To The Wages of Young Men From Working While in School?" *The Review Of Economics And Statics*, 84, 221–236.

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

KANE, T. J. (1995): "Rising Public College Tuition and College Entry: How well Do Public Subsidies Promote Access to College," *NBER Working Paper Series*, 5164.

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 (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76, 648–672.

——— (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.

LEÓN, A. (2003): "Does Ethnic Capital Matter? Identifying Peer Effects in the Intergenerational Transmission of Ethnic Differentials." *Unpublished Manuscript, MIT*.

LIGHT, A. (2001): "In-School Work Experience and the Returns to Schooling," *Journal of Labor Economics*, 19, 65–93.

LIGHT, A. AND W. STRAYER (2002): "From BAKKE to HOPWOOD: Does Race Affect College Attendance and Completion?" *Review of Economics and Statistics*, 84, 34–44.

MAGNAC, T. AND D. THESMAR (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70, 801–816.

MILLER, R. A. (1997): *Estimating Models of Dynamic Optimization with Microeconomic Data*, vol. 2, Basil Blackwell.

MILLER, R. A. AND S. SANDERS (1997): "Human Capital Development and Welfare Participation," *Carnegie-Rochester Conference Series on Public Policy*, 46, 237–253.

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.

NEWEY, W. (1994): "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 62, 1349–1382.

NEWEY, W. K. AND D. MCFADDEN (1994): *Large Sample Estimation and Hypothesis Testing*, Elsevier Science Publishers.

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.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

——— (1994a): *Estimation of Dynamic Structural Models, Problems and Prospects: Discrete Decision Processes*, vol. II, Cambridge University Press.

——— (1994b): "Structural Estimation of Markov Decision Processes," *Handbook of Econometrics*, 4, 3081–3143.

SHAW, K. L. (1989): “Life-Cycle Labor Supply with Human Capital Accumulation,” *International Economic Review*, 30, 431–456.

SHEA, J. (2000): “Does Parents’ Money Matter?” *Journal of Public Economics*, 77, 155–184.

STINEBRICKER, T. R. AND R. STINEBRICKER (2001): “Working During School and Academic Performance,” *CIBC Working Paper Series*.

TABER, C. R. (1999): “Semiparametric Identification and Heterogeneity in Discrete Choice Dynamic Programming Models,” *Journal of Econometrics*, 96, 201–229.

WILLIS, R. J. (1986): *Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions*, Elsevier Science Publishers.

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 (2) 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.

B Standard Errors for the Probability of Grade Promotion

Let y_{nt} be in indicator variable equal to 1 if the individual advances a grade level, and 0 otherwise. Define:

$$g(x_4, B_4, B_3) \equiv x_4 \left(y - \frac{e^{x_4' B_4}}{1 + e^{x_4' B_4}} \right) \tag{B.1}$$

$$h(x_3, B_3) \equiv x_3(\ln(s) - x_3' B_3) \tag{B.2}$$

$$f(x, \theta) \equiv [g(x_4, B_3, B_4)', h(x_3, B_3)']' \tag{B.3}$$

where $\theta \equiv (B_4', B_3')'$. Equation (B.1) is the score contribution of a single individual from the likelihood function constructed from equation (8.2). Equation (B.2) is the moment condition derived from the study time equation (8.1). I assume that these two moments are uncorrelated, and we have by construction that $\frac{1}{N} \sum_n f(x, \hat{\theta}) = 0$. The proof that $\hat{\theta} \xrightarrow{p} \theta_0$ is straightforward and therefore omitted. Let

$$
\begin{aligned}
G_4 &\equiv E[\Delta_{B_4} g(x_4,B_4,B_3)] = -E\left[x_4x_4' \frac{e^{x_4'B_4}}{1+e^{x_4'B_4}} \frac{1}{1+e^{x_4'B_4}}\right] && \text{(B.4)}\\
G_3 &\equiv E[\Delta_{B_3} g(x_4,B_4,B_3)] = -E\left[x_4x_3'(sB_{4,1}+2s^2B_{4,2}) \frac{e^{x_4'B_4}}{1+e^{x_4'B_4}} \frac{1}{1+e^{x_4'B_4}}\right] && \text{(B.5)}\\
H_3 &\equiv E[\Delta_{B_3} h(x_3,B_3)] = -E[x_3x_3']. && \text{(B.6)}
\end{aligned}
$$

Since $f(x,\theta)$ satisfies conditions $(i)-(v)$ of Theorem 3.4 of Newey and McFadden (1994), $\hat{B}_4$ is asymptotically normal and $\sqrt{n}(\hat{B}_4 - B_4) \xrightarrow{d} N(0,V)$, where

$$
V = G_4^{-1}E[g(x_4)g(x_4)']G_4^{-1\prime} + G_4^{-1}G_3H_3^{-1}E[h(x_3)h(x_3)']H_3^{-1\prime}G_3'G_4^{-1\prime} \tag{B.7}
$$

Thus the variance can be consistently estimated by replacing the jacobian terms in the equation (B.7) with their sample averages.

C The estimation method for the CCP's and the conditional state probabilities

Let $K[\delta_N^{-1}(\Psi_{mr}^N - \Psi_{nt}^N)]$ be a kernel, where δ_N is an appropriately chosen bandwidth. Then the nonparametric estimate of p_{ntj} is computed using the kernel estimator

$$
p_{ntj}^N \equiv \frac{\sum_{m=1}^N \sum_{r=1}^T d_{mrj} K[\delta_N^{-1}(\Psi_{mr}^N - \Psi_{nt}^N)]}{\sum_{m=1}^N \sum_{r=1}^T K[\delta_N^{-1}(\Psi_{mr}^N - \Psi_{nt}^N)]}. \tag{C.1}
$$

To define the conditional state probabilities we first define the set of possible histories that will become relevant in the model. Accordingly, the $(2\rho+K+1)$-dimensional vectors

$$
\begin{aligned}
x_{nt0}^{(i)} \equiv\ & (h_{nt-\rho+i},\cdots,h_{nt-1},0,\cdots,0,s_{nt-\rho+i},\cdots,s_{nt-1},0,\cdots,0,\\
& S_{nt-\rho+i+1},\cdots,S_{nt},S_{nt},\cdots,S_{nt},E_{nt-\rho+i},z_{nt+i}),\\
x_{nt1}^{(i)} \equiv\ & (h_{nt-\rho+i},\cdots,h_{nt-1},h_{nt}^*,\cdots,0,s_{nt-\rho+i},\cdots,s_{nt-1},0,\cdots,0,\\
& S_{nt-\rho+i+1},\cdots,S_{nt},S_{nt},\cdots,S_{nt},E_{nt-\rho+i},z_{nt+i}),\\
x_{nt2}^{(i)} \equiv\ & (h_{nt-\rho+i},\cdots,h_{nt-1},0,\cdots,0,s_{nt-\rho+i},\cdots,s_{nt-1},s_{nt}^*,\cdots,0,\\
& S_{nt-\rho+s+1},\cdots,S_{nt},S_{nt}+1,\cdots,S_{nt}+1,E_{nt-\rho+i},z_{nt+s}),\\
x_{nt3}^{(i)} \equiv\ & (h_{nt-\rho+i},\cdots,h_{nt-1},0,\cdots,0,s_{nt-\rho+i},\cdots,s_{nt-1},s_{nt}^*,\cdots,0,\\
& S_{nt-\rho+i+1},\cdots,S_{nt},S_{nt},\cdots,S_{nt},E_{nt-\rho+i},z_{nt+i}),\\
x_{nt4}^{(i)} \equiv\ & (h_{nt-\rho+i},\cdots,h_{nt-1},h_{nt}^*,\cdots,0,s_{nt-\rho+i},\cdots,s_{nt-1},s_{nt}^*,\cdots,0,\\
& S_{nt-\rho+i+1},\cdots,S_{nt},S_{nt}+1,\cdots,S_{nt}+1,E_{nt-\rho+i},z_{nt+i}),\\
x_{nt5}^{(i)} \equiv\ & (h_{nt-\rho+i},\cdots,h_{nt-1},h_{nt}^*,\cdots,0,s_{nt-\rho+i},\cdots,s_{nt-1},s_{nt}^*,\cdots,0,\\
& S_{nt-\rho+s+1},\cdots,S_{nt},S_{nt},\cdots,S_{nt},E_{nt-\rho+i},z_{nt+s}),
\end{aligned} \tag{C.2}
$$

for $i = 1, \cdots, \rho$, where h^*_{nt} and s^*_{nt} is the fraction of time individual n devotes to working and schooling conditional on participating and enrolling. Define the state vectors $\Psi^{(i)}_{ntk} \equiv (x^{(i)}_{ntk}, \mu_n \eta_n \omega_{nt+i} \lambda_{t+i})$, $k = 0, \cdots, 5$, where $\omega_{nt} \equiv \omega_{t1}^{e_{nt1}} \omega_{t2}^{e_{nt2}}$. For example, $\Psi^{(i)}_{nt1}$ is the state of a young man who has accumulated the history

$$(h_{nt-\rho}, \cdots, h_{nt-1}, s_{nt-\rho}, \cdots, s_{nt-1}, S_{nt-\rho+1}, \cdots, S_{nt}, E_{nt-\rho+1})$$

up to period t, chooses not to enroll in school and to work h^*_{nt} hours in period t, and not to enroll nor work for $i-1$ periods following t. Similarly, $\Psi^{(i)}_{nt3}$ is the state of a young man who has accumulated the same history up to period t, chooses not to work, to and study s^*_{nt} hours in period t, gets promoted a grad at the end of year t, and chooses not to enroll nor work for $i-1$ periods following t.

Define $p_j(\Psi^{(i)}_{ntk})$, $j = 0, \cdots, 3$, $k = 0, \cdots, 5$, as the the probability that individual n chooses alternative j in period $t+i$ conditioned on realizing the state vector Ψ^{i}_{ntk} in period $t+i$. The intuition for estimating these future state probabilities is to condition on observationally equivalent men in the current period. To do this, define the indicator variables:

$$d^{(i)}_{ntj} \equiv \begin{cases} d_{nt-i,j} \prod_{r=1}^{i-1} d_{nt-r,0}, & \text{for } j = 0, 1, \\ y_{nt-i} d_{nt-i,j} \prod_{r=1}^{i-1} d_{nt-r,0}, & \text{for } j = 2, 4, \\ (1 - y_{nt-i}) d_{nt-i,j} \prod_{r=1}^{i-1} d_{nt-r,0}, & \text{for } j = 3, 5, \end{cases} \tag{C.3}$$

where y_{nt} is equal to one if the individual is promoted a grade level at the end of period t, and zero otherwise. Therefore, $d^{(i)}_{ntj}$ allows us to condition of the appropriate history for computing the estimators of the state probabilities $p_k(\Psi^{(i)}_{ntj})$, which are computed as

$$p^N_k(\Psi^{(i)}_{ntj}) \equiv \frac{\sum_{m=1}^{N} \sum_{r=1}^{T} d_{mrk} d^{(i)}_{mrj} K[\delta_N^{-1}(\Psi^N_{mr} - \Psi^N_{nt})]}{\sum_{m=1}^{N} \sum_{r=1}^{T} d^{(i)}_{mrj} K[\delta_N^{-1}(\Psi^N_{mr} - \Psi^N_{nt})]}. \tag{C.4}$$

Estimation of the parameters characterizing preference also require that the derivatives of the probabilities with respect to h be estimated. The methodology employed to estimate these quantities is found in Altug and Miller (1998).

D Derivation of the moment conditions for the final stage estimation

Hotz and Miller (1993) prove the existence of a mapping $q : [0, 1] \to \Re$ such that

$$q(p_k(\Psi_{nt})) \;\; = \;\; V_j(\Psi_{nt}) - V_k(\Psi_{nt}), \tag{D.1}$$

Equations (D.1) and (3.14) are used to derive the alternative representation of the conditional valuation function V_{ntk} for the finite dependence case. To do so, define

$$u_j(\Psi_{nt}) \equiv \begin{cases} u_1(S_{nt},0)+u_2(x_{nt},0)+u_3(x_{nt},1)+\alpha^{-1}\eta_n\lambda_t c_{nt} & for\, j=0, \\ u_1(S_{nt},0)+u_2(x_{nt},1)+u_3(x_{nt},1-h^*_{nt})+\alpha^{-1}\eta_n\lambda_t c_{nt} & for\, j=1, \\ u_1(S_{nt},1)+u_2(x_{nt},0)+u_3(x_{nt},1-s_{nt})+\alpha^{-1}\eta_n\lambda_t c_{nt} & for\, j=2, \\ u_1(S_{nt},1)+u_2(x_{nt},1)+u_3(x_{nt},1-h^*_{nt}-s_{nt})+\alpha^{-1}\eta_n\lambda_t c_{nt} & for\, j=3. \end{cases} \tag{D.2}$$

Recall that $F_j(\Psi^{(i)}_{nt}|\Psi_{nt})$ is the probability that the state vector of individual n in period $t+i$ is $\Psi^{(i)}_{nt}$, given that his state vector in period t is Ψ_{nt} and he chooses alternative j in period t. Then by recursive application of the law of iterated expectations, the conditional valuation function can be expressed as

$$\begin{aligned} V_j(\Psi_{nt}) \;&= u_j(\Psi_{nt})+E_t\Big\{\textstyle\sum_{i=1}^{\rho}\Big[\beta^i\sum_{\mathcal{A}^{(i)}_{ntj}}\Big[u_0(\Psi^{(i)}_{nt})+\varphi_0(p_0(\Psi^{(i)}_{nt})) \\ &+\textstyle\sum_{k=1}^{3}p_k(\Psi^{(i)}_{nt})(q(p_k(\Psi^{(i)}_{nt}))+\varphi_k(p_k(\Psi^{(i)}_{nt})) \\ &-\varphi_0(p_0(\Psi^{(i)}_{nt})))\Big]F_j(\Psi^{(i)}_{nt}|\Psi_{nt}) \\ &+\beta^{\rho+1}\textstyle\sum_{\mathcal{A}^{(\rho+1)}_{ntj}}\Big[V_0(\Psi^{(\rho+1)}_{nt})+\varphi_0(p_0(\Psi^{(\rho+1)}_{nt})) \\ &+\textstyle\sum_{k=1}^{3}p_k(\Psi^{(\rho+1)}_{nt})(q(p_k(\Psi^{(\rho+1)}_{nt}))+\varphi_k(p_k(\Psi^{(\rho+1)}_{nt})) \\ &-\varphi_0(p_0(\Psi^{(\rho+1)}_{nt})))\Big]F_j(\Psi^{(\rho+1)}_{nt}|\Psi_{nt})\Big]\Big\}, \end{aligned} \tag{D.3}$$

Notice that the recursive substitution employed to obtain the alternative representation is only valid up to where $p_0(\Psi^{i}_{ntj})>0$. In the context of this paper, this condition is true at $i=2$ for $j=0,1$, and $i=1$ for $j=2,\cdots,5$. Equation (D.3) gives the following alternative representation of the Euler equations for labor supply and schooling

$$\begin{aligned} 0 = \;& \tfrac{\partial u_j(\Psi_{nt})}{\partial g_{nt}}+E_t\Big\{\textstyle\sum_{i=1}^{\rho}\Big[\sum_{\mathcal{A}^{(i)}_{ntj}}\Big[\tfrac{\partial[u_0(\Psi^{(i)}_{nt})+\varphi_0(p_0(\Psi^{(i)}_{nt}))]}{\partial g_{nt}} \\ &+\textstyle\sum_{k=1}^{3}p_k(\Psi^{(i)}_{nt})\tfrac{\partial[(q(p_k(\Psi^{(i)}_{nt}))+\varphi_k(p_k(\Psi^{(i)}_{nt}))-\varphi_0(p_0(\Psi^{(i)}_{nt})))]}{\partial g_{nt}} \\ &+\textstyle\sum_{k=1}^{3}[(q(p_k(\Psi^{(i)}_{nt}))+\varphi_k(p_k(\Psi^{(i)}_{nt})) \\ &-\varphi_0(p_0(\Psi^{(i)}_{nt})))]\tfrac{p_k(\Psi^{(i)}_{nt})}{\partial g_{nt}}\Big]F_j(\Psi^{(i)}_{nt}|\Psi_{nt})\Big] \\ &+\textstyle\sum_{\mathcal{A}^{(\rho+1)}_{ntj}}\Big[u_0(\Psi^{(i)}_{nt})+\varphi_0(p_0(\Psi^{(i)}_{nt})) \\ &+\textstyle\sum_{k=1}^{3}[p_k(\Psi^{(i)}_{nt})(q(p_k(\Psi^{(i)}_{nt}))+\varphi_k(p_k(\Psi^{(i)}_{nt})) \\ &-\varphi_0(p_0(\Psi^{(i)}_{nt})))]\tfrac{F_j(\Psi^{(i)}_{nt}|\Psi_{nt})}{\partial g_{nt}}\Big]\Big\}, \end{aligned} \tag{D.4}$$

where $g_{nt}=\{h_{nt},s_{nt}\}$. Assume that $\varepsilon_{ont},\cdots,\varepsilon_{nt3}$ are identically and independently distributed over (n,t) as Type 1 extreme value random variables. This assumption leads to convenient representations for the differences in the conditional valuation functions, and the expected values of the alternative specific unobservables when their corresponding alternative have been chosen. Specifically we have that $q(p_k(\Psi_{nt})) = \ln\left[\frac{p_k(\Psi_{nt})}{p_0(\Psi_{nt})}\right]$, $\varphi_k(p_k(\Psi_{nt})) = \gamma - \ln(p_k(\Psi_{nt}))$, and $\varphi_k(p_k(\Psi_{nt})) - \varphi_0(p_0(\Psi_{nt})) = -\ln\left[\frac{p_k(\Psi_{nt})}{p_0(\Psi_{nt})}\right]$.

Note that the transition matrix is degenerate conditional on the individual choosing not to enroll in school. If he chooses to enroll in school, the probability of advancing a grade level is $F(x_{nt})$. This implies that the transition probabilities for $i = 1, \cdots, \rho$ are given by $F(\Psi^{(i)}_{nt,j}|\Psi_{nt}) = 1$, for $j = 0,1$, $F(\Psi^{(i)}_{nt,j}|\Psi_{nt}) = F(x_{nt})$ for $j = 2,4$, and $F(\Psi^{(i)}_{nt,j}|\Psi_{nt}) = (1 - F(x_{nt}))$ for $j = 2,4$. Define $\xi nt \equiv (1-\alpha)^{-1}\ln(\eta_n\lambda_t)$. Then we marginal utility of consumption can be expressed as $\eta_n\lambda_t \equiv \exp((1-\alpha)\xi_{nt})$.

The parametric assumptions on the utility functions and the idiosyncratic taste shifters, and the Euler conditions for work and schooling from equation (D.5) are used to form population moment conditions. We can then define

$$\begin{aligned}
m_{nt1}(\Theta) \equiv\ & d_{nt1}\Big[\alpha^{-1}\eta_n\lambda_t w_{nt} - z'_{nt}B_5 - 2\delta_0 l^{(1)}_{nt} - \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt-i}+\beta^i) \\
& - \textstyle\sum_{i=1}^{\rho}\beta^i p_0(\Psi^{(i)}_{nt1})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt1})}{\partial h_{nt}}\Big] \\
& + d_{nt3}\Big[\alpha^{-1}\eta_n\lambda_t w_{nt} - z'_{nt}B_5 - 2\delta_0 l^{(3)}_{nt} - \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt-i}+\beta^i) \\
& - \textstyle\sum_{i=1}^{\rho}\beta^i\Big[p_0(\Psi^{(i)}_{nt4})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt4})}{\partial h_{nt}}F(x_{nt}) + p_0(\Psi^{(i)}_{nt5})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt5})}{\partial h_{nt}}(1-F(x_{nt})) \\
& + \ln\left(\frac{p_0(\Psi^{(i)}_{nt5})}{p_0(\Psi^{(i)}_{nt4})}\right)\frac{\partial F(x_{nt})}{\partial h_{nt}}\Big]\Big].
\end{aligned}$$

$$\begin{aligned}
m_{nt2}(\Theta) \equiv\ & d_{nt2}\Big[-z'_{nt}B_5 - 2\delta_0 l^{(2)}_{nt} - \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt-i}+\beta^i) \\
& - \textstyle\sum_{i=1}^{\rho}\beta^i\Big[p_0(\Psi^{(i)}_{nt2})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt2})}{\partial s_{nt}}F(x_{nt}) + p_0(\Psi^{(i)}_{nt3})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt3})}{\partial s_{nt}}(1-F(x_{nt})) \\
& + \ln\left(\frac{p_0(\Psi^{(i)}_{nt3})}{p_0(\Psi^{(i)}_{nt2})}\right)\frac{\partial F(x_{nt})}{\partial s_{nt}}\Big]\Big] + d_{nt3}\Big[-z'_{nt}B_5 - 2\delta_0 l^{(3)}_{nt} - \textstyle\sum_{i=1}^{\rho}\delta_i(l_{nt-i}+\beta^i) \\
& - \textstyle\sum_{i=1}^{\rho}\beta^i\Big[p_0(\Psi^{(i)}_{nt4})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt4})}{\partial s_{nt}}F(x_{nt}) + p_0(\Psi^{(i)}_{nt5})^{-1}\frac{\partial p_0(\Psi^{(i)}_{nt5})}{\partial s_{nt}}(1-F(x_{nt})) \\
& + \ln\left(\frac{p_0(\Psi^{(i)}_{nt5})}{p_0(\Psi^{(i)}_{nt4})}\right)\frac{\partial F(x_{nt})}{\partial s_{nt}}\Big]\Big].
\end{aligned}$$

The parametric assumptions on the utility functions, the distribution of the idiosyncratic taste shifters, equation (D.1) and equation (D.3) are used to obtain the following additional moment conditions[11]

$$
\begin{aligned}
m_{nt3}(\Theta) \equiv \;& d_{nt1}\Big[\ln\left(\tfrac{p_{nt1}}{p_{nt0}}\right) - x'_{nt}B_6 + x'_{nt}B_7(l^{(0)}_{nt} - l^{(1)}_{nt}) + \delta_0(l^{(0)2}_{nt} - l^{(1)2}_{nt}) \\
& + \textstyle\sum_{i=1}^{\rho}\delta_i(l^{(0)}_{nt} - l^{(1)}_{nt})(l_{nt-i} + \beta^i) - \tfrac{\eta_n\lambda_t}{\alpha}(w_{nt}h_{nt}) - \textstyle\sum_{i=1}^{\rho}\beta^i \ln\left(\tfrac{p_0(\Psi^{(s)}_{nt0})}{p_0(\Psi^{(s)}_{nt1})}\right)\Big],
\end{aligned}
$$

$$
\begin{aligned}
m_{nt4}(\Theta) \equiv \;& d_{nt2}\Big[\ln\left(\tfrac{p_{nt2}}{p_{nt0}}\right) - x'_{nt}B_5 + x'_{nt}B_7(l^{(0)}_{nt} - l^{(2)}_{nt}) \\
& + \delta_0(l^{(0)2}_{nt} - l^{(2)2}_{nt}) + \textstyle\sum_{i=1}^{\rho}\delta_i(l^{(0)}_{nt} - l^{(2)}_{nt})(l_{nt-i} + \beta^i) + \tfrac{\eta_n\lambda_t}{\alpha}\pi_{nt} \\
& - \textstyle\sum_{i=1}^{\rho}\beta^i\left[\ln p_0(\Psi^{(i)}_{nt0}) - \ln p_0(\Psi^{(i)}_{nt2})F(x_{nt}) - \ln p_0(\Psi^{(i)}_{nt3})(1 - F(x_{nt}))\right]\Big],
\end{aligned}
$$

$$
\begin{aligned}
m_{nt5}(\Theta) \equiv \;& d_{nt3}\Big[\ln\left(\tfrac{p_{nt3}}{p_{nt0}}\right) - x'_{nt}B_5 - x'_{nt}B_6 + x'_{nt}B_7(l^{(0)}_{nt} - l^{(3)}_{nt}) \\
& + \delta_0(l^{(0)2}_{nt} - l^{(3)2}_{nt}) + \textstyle\sum_{i=1}^{\rho}\delta_i(l^{(0)}_{nt} - l^{(3)}_{nt})(l_{nt-i} + \beta^i) - \tfrac{\eta_n\lambda_t}{\alpha}(w_{nt}h_{nt} - \pi_{nt}) \\
& - \textstyle\sum_{i=1}^{\rho}\beta^i\left[\ln p_0(\Psi^{(i)}_{nt0}) - \ln p_0(\Psi^{(i)}_{nt4})F(x_{nt}) - \ln p_0(\Psi^{(i)}_{nt5})(1 - F(x_{nt}))\right]\Big].
\end{aligned}
$$

E Asymptotic Properties of the CCP estimator and Consistent Asymptotic Variance Estimation

Some preliminary results are in needed. The first is concerned with the estimation of the CCP's themselves. In estimation, a the data was trimmed to ensure that the density is bounded away from zero. This fixed trimming condition defines a compact subset of the support of the density over which the density affects the estimator. Assumptions 8.1 - 8.3, and the assumptions in Lemma 8.10 of Newey and McFadden (1994) ensures the resulting kernel density estimators of the CCP's and their derivatives converge uniformly:

$$\sqrt{N}\|p^N(\Psi) - p^0(\Psi)\|^2 \xrightarrow{p} 0, \tag{E.1}$$

where the norm is the Sobolev norm. Assume that , θ^N is the unique solution to:

$$\frac{1}{N}\sum_{n=1}^{N} m(x_n, \theta, \xi_n(B_1^N), s_n(B_3^N)F_n(s_n(B_3^N), B_4^N), p_n^N). \tag{E.2}$$

Assume also that $\theta_0 \in \Theta$, a compact set. Inspection of the equations in (**??**) shows that $m(x,\theta)$ is continuous in each θ. Further inspection along with the fixed trimming condition on the data in estimation implies that $m(z,\theta)$ is uniformly bounded over θ. These conditions ensures that $\theta^N \xrightarrow{p} \theta_0$ as shown in Theorem 2.6 of Newey and McFadden (1994).

[11] The construction of the moment conditions show that the choice of the normalizing alternative (alternative 0) is not completely arbitrary. This alternative has to sufficiently saturate the state space so that $p_{nt0} > 0$ and $p_0(\Psi^i_{ntj}) > 0$.

Define the following influence functions from equations (**??**) and from the definitions in section B

$$
\begin{aligned}
\varphi_1(x_{1n}) &\equiv -E[\Delta x'_{1n} A_n^{-1} \Delta x_{1n}]^{-1} \Delta x'_{1n} A_n^{-1} \Delta v_{1n}, \quad \varphi_3(x_{3n}) \equiv -H_3^{-1} h(x_{3n}), \\
\varphi_4(x_{4n}) &\equiv -H_4^{-1} h(x_{4n}).
\end{aligned} \tag{E.3}
$$

Define the following matrices

$$
\begin{gathered}
M_{1nt} \equiv \begin{bmatrix} (d_{nt1}+d_{nt3})(\frac{1-\alpha}{\alpha})\exp((1-\alpha)\xi_{nt}))w_{nt} \\ 0 \\ -d_{nt1}(\frac{1-\alpha}{\alpha})\exp((1-\alpha)\xi_{nt}))w_{nt}h_{nt} \\ d_{nt2}(\frac{1-\alpha}{\alpha})\exp((1-\alpha)\xi_{nt}))\pi_{nt} \\ d_{nt3}(\frac{1-\alpha}{\alpha})\exp((1-\alpha)\xi_{nt}))(w_{nt}h_{nt}-\pi_{nt}) \end{bmatrix} \left[-\frac{1}{N}\sum_n x'_{1nt}\right], \\
M_{1n}(x_n) \equiv (M'_{1n1}, \cdots, M'_{1nT})', \text{ and, } \alpha_1(x_n) \equiv E[M_{1n}]\varphi_1(x_{1n}).
\end{gathered} \tag{E.4}
$$

$$
M_{2nt} \equiv \begin{bmatrix} d_{nt1}\sum_i d^s_{nt-i}\delta_i + \\ d_{nt3}\left[2\delta_0 + \sum_i \left(d^s_{nt-i}\delta_i - \beta^i\left(\left(\frac{1}{p^i_{0nt4}}\frac{\partial p^i_{0nt4}}{\partial h_{nt}} - \frac{1}{p^i_{0nt5}}\frac{\partial p^i_{0nt5}}{\partial h_{nt}}\right)\frac{\partial F(x_{nt})}{\partial s_{nt}} + \ln(\frac{p^i_{0nt5}}{p^i_{0nt4}})\frac{\partial^2 F(x_{nt})}{\partial h_{nt}\partial s_{nt}}\right)\right)\right] \\ d_{nt2}\left[2\delta_0 + \sum_i \left(d^s_{nt-i}\delta_i - \beta^i\left(\left(\frac{1}{p^i_{0nt2}}\frac{\partial p^i_{0nt2}}{\partial s_{nt}} - \frac{1}{p^i_{0nt3}}\frac{\partial p^i_{0nt3}}{\partial s_{nt}}\right)\frac{\partial F(x_{nt})}{\partial s_{nt}} + \ln(\frac{p^i_{0nt3}}{p^i_{0nt2}})\frac{\partial^2 F(x_{nt})}{\partial s^2_{nt}}\right)\right)\right] + \\ d_{nt3}\left[2\delta_0 + \sum_i \left(d^s_{nt-i}\delta_i - \beta^i\left(\left(\frac{1}{p^i_{0nt4}}\frac{\partial p^i_{0nt54}}{\partial s_{nt}} - \frac{1}{p^i_{0nt5}}\frac{\partial p^i_{0nt5}}{\partial s_{nt}}\right)\frac{\partial F(x_{nt})}{\partial s_{nt}} + \ln(\frac{p^i_{0nt5}}{p^i_{0nt4}})\frac{\partial^2 F(x_{nt})}{\partial s^2_{nt}}\right)\right)\right] \\ -d_{nt1}\sum_i \delta_i(l^0_{nt} - l^1_{nt})d^s_{nt-i} \\ d_{nt2}\left[x'_{6nt}B_6 + 2\delta_0 l^2_{nt} + \sum_i \delta_i l_{nt-i}\sum_i \beta^i\left[\ln(\frac{p^i_{0nt2}}{p^i_{0nt3}})\frac{\partial F(x_{nt})}{\partial s_{nt}}\right]\right] \\ d_{nt3}\left[x'_{6nt}B_6 + 2\delta_0 l^3_{nt} + \sum_i \delta_i l_{nt-i}\sum_i \beta^i\left[\ln(\frac{p^i_{0nt4}}{p^i_{0nt5}})\frac{\partial F(x_{nt})}{\partial s_{nt}}\right]\right] \end{bmatrix} \left[s_{nt}x'_{3nt}\right],
$$

$$
M_{2n}(x_n) \equiv (M'_{2n1}, \cdots, M'_{2nT})', \text{ and, } \alpha_2(x_n) \equiv E[M_{2n}]\varphi_2(x_{3n}). \tag{E.5}
$$

$$
M_{4nt} \equiv \begin{bmatrix} -d_{nt3}\sum_i \beta^i\left[\frac{1}{p^i_{0nt4}}\frac{\partial p^i_{0nt4}}{\partial h_{nt}} - \frac{1}{p^i_{0nt5}}\frac{\partial p^i_{0nt5}}{\partial h_{nt}} + \ln(\frac{p^i_{0nt5}}{p^i_{0nt4}})\frac{\partial^2 F(x_{nt})}{\partial h^2_{nt}}\right] \\ -d_{nt2}\sum_i \beta^i\left[\frac{1}{p^i_{0nt2}}\frac{\partial p^i_{0nt2}}{\partial s_{nt}} - \frac{1}{p^i_{0nt3}}\frac{\partial p^i_{0nt3}}{\partial s_{nt}} + \ln(\frac{p^i_{0nt3}}{p^i_{0nt2}})\frac{\partial^2 F(x_{nt})}{\partial h_{nt}\partial F}\right] - \\ d_{nt3}\sum_i \beta^i\left[\frac{1}{p^i_{0nt4}}\frac{\partial p^i_{0nt4}}{\partial s_{nt}} - \frac{1}{p^i_{0nt5}}\frac{\partial p^i_{0nt5}}{\partial s_{nt}} + \ln(\frac{p^i_{0nt5}}{p^i_{0nt4}})\frac{\partial^2 F(x_{nt})}{\partial h_{nt}\partial F}\right] \\ 0 \\ d_{nt2}\sum_i \beta^i\left[\ln(\frac{p^i_{0nt2}}{p^i_{0nt3}})\right] \\ d_{nt3}\sum_i \beta^i\left[\ln(\frac{p^i_{0nt4}}{p^i_{0nt5}})\right] \end{bmatrix} \left[F(x_{nt})(1-F(x_{nt}))x'_{4nt}\right],
$$

$$
M_{4n}(x_n) \equiv (M'_{4n1}, \cdots, M'_{4nT})', \text{ and, } \alpha_4(x_n) \equiv E[M_{4n}]\varphi_4(x_{4n}). \tag{E.6}
$$

$$D_{nt0} \equiv E\left[\frac{\partial m_{nt}}{\partial p_{nt0}}|\Psi_{nt}\right] = -p_{nt0}^{-1}(0,0,p_{nt1},p_{nt2},p_{nt3})'$$
$$D_{n0}(x_n) \equiv (D'_{n10},\cdots,D'_{nT0})', \text{ and, } \alpha_5(x_n) \equiv D_{n0}[d_{n0}-p_{n0}]. \tag{E.7}$$

$$D_{nt1} \equiv E\left[\frac{\partial m_{nt}}{\partial p_{nt1}}|\Psi_{nt}\right] = (0,0,1,0,0)'$$
$$D_{n1}(x_n) \equiv (D'_{n11},\cdots,D'_{nT1})', \text{ and, } \alpha_6(x_n) \equiv D_{n1}[d_{n1}-p_{n1}]. \tag{E.8}$$

$$D_{nt2} \equiv E\left[\frac{\partial m_{nt}}{\partial p_{nt2}}|\Psi_{nt}\right] = (0,0,0,1,0)'$$
$$D_{n2}(x_n) \equiv (D'_{n12},\cdots,D'_{nT2})', \text{ and, } \alpha_7(x_n) \equiv D_{n2}[d_{n2}-p_{n2}]. \tag{E.9}$$

$$D_{nt3} \equiv E\left[\frac{\partial m_{nt}}{\partial p_{nt3}}|\Psi_{nt}\right] = (0,0,0,0,1)'$$
$$D_{n3}(x_n) \equiv (D'_{n13},\cdots,D'_{nT3})', \text{ and, } \alpha_8(x_n) \equiv D_{n3}[d_{n3}-p_{n3}]. \tag{E.10}$$

For $i = 1,\cdots,\rho$ define.

$$D_{nt0i} \equiv E\left[\frac{\partial m_{nt}}{\partial p_{0nt0}^{(i)}}|\Psi_{nt0}^{(i)}\right] = \beta^i\left(0,0,\frac{p_{1nt0}^{(i)}}{p_{0nt0}^{(i)}},\frac{p_{2nt0}^{(i)}}{p_{0nt0}^{(i)}},\frac{p_{3nt0}^{(i)}}{p_{0nt0}^{(i)}}\right)'$$
$$D_{n0i}(x_n) \equiv (D'_{n10i},\cdots,D'_{nT0i})', \text{ and, } \alpha_{9i}(x_n) \equiv D_{n0i}[d_{n0}-p_{n0}^{(i)}]. \tag{E.11}$$

$$D_{nt1i} \equiv E\left[\frac{\partial m_{nt}}{\partial p_{0nt1}^{(i)}}|\Psi_{nt1}^{(i)}\right] = \beta^i\left(\frac{p_{1nt1}^{(i)}}{(p_{0nt1}^{(i)})^2}\nabla_h p_{0nt1}^{(i)},0,\frac{p_{1nt1}^{(i)}}{p_{0nt1}^{(i)}},0,0\right)'$$
$$D_{n1i}(x_n) \equiv (D'_{n11i},\cdots,D'_{nT1i})', \text{ and, } \alpha_{10i}(x_n) \equiv D_{n1i}[d_{n1}-p_{n1}^{(i)}]. \tag{E.12}$$

$$D_{nt2i} \equiv E\left[\frac{\partial m_{nt}}{\partial p_{0nt2}^{(i)}}|\Psi_{nt2}^{(i)}\right] = \beta^i\left(0,\frac{p_{2nt2}^{(i)}}{(p_{0nt2}^{(i)})^2}\nabla_s p_{0nt2}^{(i)}F(x_{nt})-\frac{p_{2nt2}^{(i)}}{p_{0nt2}^{(i)}}\nabla_s F(x_{nt}),0,\frac{p_{2nt2}^{(i)}}{p_{0nt2}^{(i)}}F(x_{nt}),0\right)'$$
$$D_{n2i}(x_n) \equiv (D'_{n12i},\cdots,D'_{nT2i})', \text{ and, } \alpha_{11i}(x_n) \equiv D_{n2i}[d_{n0}-p_{0n2}^{(i)}]. \tag{E.13}$$

$$D_{nt3i} \equiv E\left[\frac{\partial m_{nt}}{\partial p_{0nt3}^{(i)}}|\Psi_{nt3}^{(i)}\right] = \beta^i\left(0,\frac{p_{2nt3}^{(i)}}{(p_{0nt3}^{(i)})^2}\nabla_s p_{0nt3}^{(i)}(1-F(x_{nt}))+\frac{p_{2nt3}^{(i)}}{p_{0nt3}^{(i)}}\nabla_s F(x_{nt}),0,\frac{p_{2nt3}^{(i)}}{p_{0nt3}^{(i)}}(1-F(x_{nt})),0\right)'$$
$$D_{n3i}(x_n) \equiv (D'_{n13i},\cdots,D'_{nT3i})', \text{ and, } \alpha_{12i}(x_n) \equiv D_{n3i}[d_{n0}-p_{0n3}^{(i)}]. \tag{E.14}$$

$$D_{nt4i} \equiv E\left[\frac{\partial m_{nt}}{\partial p^{(i)}_{0nt4}}|\Psi^{(i)}_{nt4}\right] = \beta^i \begin{bmatrix} \frac{p^{(i)}_{3nt4}}{(p^{(i)}_{0nt4})^2}\nabla_h p^{(i)}_{0nt4}F(x_{nt}) - \frac{p^{(i)}_{3nt4}}{p^{(i)}_{0nt4}}\nabla_h F(x_{nt}) \\ \frac{p^{(i)}_{3nt4}}{(p^{(i)}_{0nt4})^2}\nabla_s p^{(i)}_{0nt4}F(x_{nt}) - \frac{p^{(i)}_{3nt4}}{p^{(i)}_{0nt4}}\nabla_s F(x_{nt}) \\ 0 \\ 0 \\ \frac{p^{(i)}_{3nt4}}{p^{(i)}_{0nt4}}\nabla_s F(x_{nt} \end{bmatrix}$$

$$D_{n4i}(x_n) \equiv (D'_{n14i}, \cdots, D'_{nT4i})', \text{ and, } \alpha_{13i}(x_n) \equiv D_{n4i}[d_{n0} - p^{(i)}_{0n4}]. \tag{E.15}$$

$$D_{nt5i} \equiv E\left[\frac{\partial m_{nt}}{\partial p^{(i)}_{0nt5}}|\Psi^{(i)}_{nt5}\right] = \beta^i \begin{bmatrix} \frac{p^{(i)}_{3nt5}}{(p^{(i)}_{0nt5})^2}\nabla_h p^{(i)}_{0nt5}(1 - F(x_{nt})) + \frac{p^{(i)}_{3nt5}}{p^{(i)}_{0nt5}}\nabla_h F(x_{nt}) \\ \frac{p^{(i)}_{3nt5}}{(p^{(i)}_{0nt5})^2}\nabla_s p^{(i)}_{0nt5}F(x_{nt}) + \frac{p^{(i)}_{3nt5}}{p^{(i)}_{0nt5}}\nabla_s F(x_{nt}) \\ 0 \\ 0 \\ \frac{p^{(i)}_{3nt5}}{p^{(i)}_{0nt5}}\nabla_s F(x_{nt} \end{bmatrix}$$

$$D_{n5i}(x_n) \equiv (D'_{n15i}, \cdots, D'_{nT5i})', \text{ and, } \alpha_{14i}(x_n) \equiv D_{n5i}[d_{n0} - p^{(i)}_{0n5}]. \tag{E.16}$$

Let $f^i_{ntj} \equiv f(\Psi^i_{ntj})$ be the density of Ψ^i_{ntj} $j = 1, \cdots, 5$, $i = 1, \cdots, \rho$. Define also $\vartheta^i_{ntj} \equiv$

$(f(\Psi^i_{ntj}))^{-1}\frac{\partial f(\Psi^i_{ntj})}{\partial h_n}$. For $i=1,\cdots,\rho$ let

$$
\begin{aligned}
{}_hM_{nt1i} &\equiv E\left[\frac{\partial m_{nt}}{\partial \nabla_h p^{(i)}_{0nt1}}|\Psi^{(i)}_{nt1}\right] = \beta^i\left(\frac{p^{(i)}_{1nt1}}{p^{(i)}_{0nt1}},0,0,0,0\right)' \\
{}_{hh}M_{nt1i} &\equiv E\left[\frac{\partial m_{nt}}{\partial \nabla_h p^{(i)}_{0nt1}\partial h_{nt}}|\Psi^{(i)}_{nt1}\right] = \beta^i\left(\frac{p^{(i)}_{1nt1}}{(p^{(i)}_{0nt1})^2}\nabla_h p^{(i)}_{0nt1},0,0,0,0\right)' \\
{}_hD_{nt1i} &\equiv -\left[\ {}_{hh}M_{nt1i}+2\ {}_hM_{nt1i}\vartheta^i_{nt1}\right] && \text{(E.17)}\\
{}_hD_{n1i}(x_n) &\equiv (\ {}_hD'_{n11i},\cdots,\ {}_hD'_{nT1i})', \text{ and, } \alpha_{15i}(x_n)\equiv\ {}_hD_{n1i}[d_{n0}-p^{(i)}_{0n1}]. && \text{(E.18)}
\end{aligned}
$$

$$
\begin{aligned}
{}_sM_{nt2i} &\equiv E\left[\frac{\partial m_{nt}}{\partial \nabla_s p^{(i)}_{0nt2}}|\Psi^{(i)}_{nt2}\right] = \beta^i\left(0,\frac{p^{(i)}_{2nt2}}{p^{(i)}_{0nt2}}F(x_{nt}),0,0,0\right)' \\
{}_{ss}M_{nt2i} &\equiv E\left[\frac{\partial m_{nt}}{\partial \nabla_s p^{(i)}_{0nt2}\partial s_{nt}}|\Psi^{(i)}_{nt2}\right] = \beta^i\left(0,\frac{p^{(i)}_{2nt2}}{(p^{(i)}_{0nt2})^2}\nabla_s p^{(i)}_{0nt1}F(x_{nt})-\frac{p^{(i)}_{2nt2}}{p^{(i)}_{0nt2}}\nabla_s F(x_{nt}),0,0,0\right)' \\
{}_sD_{nt2i} &\equiv -\left[\ {}_{ss}M_{nt2i}+2\ {}_sM_{nt2i}\vartheta^i_{nt2}\right] && \text{(E.19)}\\
{}_sD_{n2i}(x_n) &\equiv (\ {}_sD'_{n12i},\cdots,\ {}_sD'_{nT2i})', \text{ and, } \alpha_{16i}(x_n)\equiv\ {}_sD_{n2i}[d_{n0}-p^{(i)}_{0n2}]. && \text{(E.20)}
\end{aligned}
$$

$$
\begin{aligned}
{}_sM_{nt3i} &\equiv E\left[\frac{\partial m_{nt}}{\partial \nabla_s p^{(i)}_{0nt3}}|\Psi^{(i)}_{nt3}\right] = \beta^i\left(0,\frac{p^{(i)}_{3nt3}}{p^{(i)}_{0nt3}}(1-F(x_{nt})),0,0,0\right)' \\
{}_{ss}M_{nt3i} &\equiv E\left[\frac{\partial m_{nt}}{\partial \nabla_s p^{(i)}_{0nt3}\partial s_{nt}}|\Psi^{(i)}_{nt3}\right] = \beta^i\left(0,\frac{p^{(i)}_{3nt3}}{(p^{(i)}_{0nt3})^2}\nabla_s p^{(i)}_{0nt1}(1-F(x_{nt}))+\frac{p^{(i)}_{3nt3}}{p^{(i)}_{0nt3}}\nabla_s F(x_{nt}),0,0,0\right)' \\
{}_sD_{nt3i} &\equiv -\left[\ {}_{ss}M_{nt3i}+2\ {}_sM_{nt3i}\vartheta^i_{nt3}\right] && \text{(E.21)}\\
{}_sD_{n3i}(x_n) &\equiv (\ {}_sD'_{n13i},\cdots,\ {}_sD'_{nT3i})', \text{ and, } \alpha_{17i}(x_n)\equiv\ {}_sD_{n3i}[d_{n0}-p^{(i)}_{0n3}]. && \text{(E.22)}
\end{aligned}
$$

The construction of ${}_hD_{nt4i}$ and ${}_sD_{nt4i}$ are the same as ${}_sD_{nt2i}$ with the correct indexes. Likewise, the construction of ${}_hD_{nt5i}$ and ${}_sD_{nt45i}$ are the same as ${}_sD_{nt3i}$ with the correct indexes. This gives additional influence functions $\alpha_{18i},\cdots,\alpha_{21i}$. Define also $\alpha(x_n)\equiv\sum_{j=1}^{8}\alpha_{nj}(x_n)+\sum_{j=9}^{21}\sum_{i=1}^{\rho}\alpha_{ji}(x_n)$. The fixed trimming condition, the smoothness properties of $m(x,\cdot)$, and condition E.1 ensures linearization is possible in the necessary arguments, that the above matrices are well defined (in particular, all expectations are well defined), and that assumptions 5.1-5.6 of Newey (1994) are satisfied. Define

$$M_\theta \equiv E\left[\frac{\partial m(x_n,\theta_0)}{\partial\theta}\right] \tag{E.23}$$

$$W\equiv E[\{m(x_n,\theta_0)+\alpha(x_n)\}\{m(x_n,\theta_0)+\alpha(x_n)\}'] \tag{E.24}$$

Therefore, by lemma 5.3 of Newey (1994), we have that

$$\sqrt{N}(\theta_N-\theta_0)\xrightarrow{p}N(0,V),$$

where

$$V\equiv(M'_\theta\Omega^{-1}M_\theta)^{-1}M'_\theta\Omega^{-1}W\Omega^{-1}M_\theta(M'_\theta\Omega^{-1}M_\theta)^{-1} \tag{E.25}$$

A consistent estimator of jacobians with respect to the finite dimensional parameters are obtained by replacing the parameters (both finite and infinite dimensional) with their respective estimates and taking averages over N. A consistent estimator jacobians with respect to the ccp's and their derivatives are obtained by replacing the parameters with their estimated counterparts and then performing nonparametric regression of these quantities on their appropriate conditioning vectors Ψ^i_{nj}. The residuals needed to complete the formation of $\hat{\alpha}(x_n)$ are readily obtained from all the parametric and nonparametric pre-estimates. By similar substitutions and averaging consistent estimates of M_θ $m(x_{nt}, \theta)$, and Ω are formed, denoted by M^n_θ, $m^N(x_n)$, and Ω^N, A consistent estimate of W is then obtained by

$$W^N = N^{-1} \sum_{n=1}^{N} \left[m^N(x_n) + \alpha^N(x_n)\right] \left[m^N(x_n) + \alpha^N(x_n)\right]'. \tag{E.26}$$

Putting all these estimated quantities together, a consistent estimator for the asymptotic variance is given by

$$V^N \equiv \left(M^{N\prime}_\theta (\Omega^N)^{-1} M^N_\theta\right)^{-1} M^{N\prime}_\theta \left(\Omega^N\right)^{-1} W^N \left(\Omega^N\right)^{-1} M^N_\theta \left(M^{N\prime}_\theta \left(\Omega^N\right)^{-1} M^N_\theta\right)^{-1}. \tag{E.27}$$

TABLE 1
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 2a

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 2b

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

TABLE 3

The Consumption Equation.

$\ln(c_{nt}) = (1-\alpha)^{-1}[z'_{nt}B_1 - \ln(\eta_n\lambda_t) + \nu_{nt}]$

Variable	Parameter	Estimate	Std. Err.
Demographic Variables			
$\Delta FAM\ SIZE_{nt}$	$B_{1,1}$	0.1466	0.0022
$\Delta FAM\ INC_{nt}$	$B_{1,2}$	8.00E-06	0.08E-06
$\Delta FAM\ AGE_{nt}$	$B_{1,3}$	4.00E-06	2.00E-06
$\Delta UNEMP_{nt}$	$B_{1,4}$	-0.0010	0.0005
ΔS_{nt}	$B_{1,5}$	-0.0091	0.0008
$\Delta(AGE \times S_{nt})$	$B_{1,6}$	0.0089	0.0008
ΔAGE^2_{nt}	$B_{1,7}$	-0.0008	0.0004

TABLE 4

The Wage Equation

$\ln(w_{ntj}) = \ln(\omega_{tj}) + \ln(\mu_n) + \Delta x'_{nt} B_{2j}$

	Low Skill		High Skill	
Variable	Parameter	Estimate	Parameter	Estimate
Lags of Enrollment				
Δd^s_{nt-1}	$B_{2,1,1}$	-0.0309	$B_{2,2,1}$	-0.0701
		(0.0382)		(0.0266)
Δd^s_{nt-2}	$B_{2,1,2}$	-0.0198	$B_{2,2,2}$	-0.01239
		(0.0421)		(0.02707)
Lags of Participation				
Δd^h_{nt-1}	$B_{2,1,3}$	0.0198	$B_{2,2,3}$	-0.1513
		(0.0431)		(0.0175)
Δd^h_{nt-2}	$B_{2,1,4}$	0.0319	$B_{2,2,4}$	-0.1272
		(0.0460)		(0.0193)
Lags of Hours Worked				
Δh_{nt-1}	$B_{2,1,5}$	0.20E-04	$B_{2,2,5}$	0.28E-04
		(0.02E-04)		(0.01E-04)
Δh_{nt-2}	$B_{2,1,6}$	0.07E-04	$B_{2,2,6}$	0.10E-04
		(0.02E-04)		(0.01E-04)
Socio-Economic Variables				
ΔS^2_{nt}	$B_{2,1,8}$	-0.29E-04	$B_{2,2,8}$	0.0040
		(1.37E-04)		(0.0001)
ΔE^2_{nt-2}	$B_{2,1,7}$	-0.0010	$B_{2,2,7}$	-0.0011
		(0.0003)		(0.0002)
$\Delta(S_{nt} \times E_{nt-2})$	$B_{2,1,9}$	0.0027	$B_{2,2,9}$	-0.0072
		(0.0003)		(0.0002)

TABLE 5

Estimated changes in aggregate prices and wages[†]

Year	Aggregate Prices $(1-\alpha)^{-1}\Delta\ln(\lambda_t)$	Aggregate Wages Unskilled $(\Delta\ln\omega_{t,1})$	Skilled $(\Delta\ln\omega_{t,2})$
1980	0.0709		
	(0.0175)		
1981	0.0509		
	(0.0180)		
1982	0.0129	0.0180	0.2016
	(0.0189)	(0.0035)	(0.0114)
1983	0.0279	0.0047	0.1916
	(0.0192)	(0.0383)	(0.0141)
1984	0.0345	0.0287	0.1127
	(0.0199)	(0.0393)	(0.0162)
1985	-0.0423	0.0449	0.2320
	(0.0200)	(0.0381)	(0.0177)
1986	0.0288	0.0526	0.2303
	(0.0206)	(0.0402)	(0.0204)
1987	0.0713	0.0584	0.2831
	(0.0218)	(0.0384)	(0.0212)
1988	-0.0102	0.0556	0.1421
	(0.0226)	(0.0363)	(0.0210)
1989	0.1111	-0.0228	0.1781
	(0.0228)	(0.0375)	(0.0221)
1990	-0.0186	0.0133	0.1652
	(0.0232)	(0.0366)	(0.0219)
1991	0.0230	-0.0360	0.1610
	(0.0237)	(0.0368)	(0.0219)
1992	0.2044	-0.0101	0.1713
	(0.0246)	(0.0392)	(0.0237)
1993	-0.0260	0.0290	0.1770
	(0.0250)	(0.0411)	(0.0252)
1994	-0.0056	0.0120	0.1587
	(0.0251)	(0.0351)	(0.0218)

[†] Standard errors in parentheses

TABLE 6

Estimate of time spent on school activities[†]

$\ln(s_n) = z_n' B_3$

Variable	Parameter	Estimate	Std.Err
Constant	$B_{3,0}$	7.2383	0.1829
Lags of Enrollment			
d^s_{nt-1}	$B_{3,1}$	0.2602	0.0463
d^s_{nt-2}	$B_{3,2}$	0.2037	0.0789
Lags of Hours Worked			
h_{nt-1}	$B_{3,3}$	-0.77E-04	0.17E-04
h_{nt-2}	$B_{3,4}$	-0.50E-04	0.26E-04
Socio-Economic Variables			
$BLACK$	$B_{3,5}$	0.1063	0.0265
$HISPANIC$	$B_{3,6}$	-0.0996	0.0304
$AGE_{nt} \times S_{nt}$	$B_{3,7}$	-0.0045	0.0013
$(AGE_{nt} \times S_{nt})^2$	$B_{3,8}$	0.76E-05	0.26E-05
$US\ BORN$	$B_{3,10}$	-0.1261	0.0417
$FAM\ SIZE_{nt}$	$B_{3,11}$	0.0135	0.0050
$RURAL$	$B_{3,12}$	0.0647	0.0250
$UNEMP_{nt}$	$B_{3,13}$	-0.0244	0.0100
$AFQT$	$B_{3,15}$	0.0037	0.0004
$\ln(\mu)$	$B_{3,17}$	-0.1435	0.0273

TABLE 7: Probability of Grade Promotion

$$F(x_{nt}) = (1-d^h_{nt})\frac{\exp(x'_{nt}B_{41})}{1+\exp(x'_{nt}B_{41})} + d^h_{nt}\frac{\exp(x'_{nt}B_{42})}{1+\exp(x'_{nt}B_{42})}$$

	$d^h_{nt}=0$		$d^h_{nt}=1$	
Variable	Parameter	Estimate	Parameter	Estimate
Constant	$B_{4,1,0}$	0.0307	$B_{4,2,0}$	0.0482
		(0.7734)		(0.5499)
Time Use Variables				
s_{nt}	$B_{4,1,1}$	0.0025	$B_{4,2,1}$	0.0036
		(0.0003)		(0.0008)
s^2_{nt}			$B_{4,2,2}$	-0.15E-05
				(0.03E-05)
h_{nt}			$B_{4,2,3}$	-0.0006
				(0.0001)
h^2_{nt}			$B_{4,2,4}$	0.10E-06
				(0.03E-06)
Enrollment Variables				
d^s_{nt-1}			$B_{4,2,5}$	0.4104
				(0.1184)
GRADE 11	$B_{4,1,2}$	0.5812	$B_{4,2,6}$	0.3215
		(0.1709)		(0.1580)
GRADE 12	$B_{4,1,3}$	0.5672	$B_{4,2,7}$	1.0022
		(0.1485)		(0.1285)
Participation Variables				
d^h_{nt-1}			$B_{4,2,8}$	0.2185
				(0.0873)
d^h_{nt-2}	$B_{4,1,4}$	0.2771		
		(0.1203)		
Socio-Economic Variables				
$BLACK$	$B_{4,1,5}$	-0.2296	$B_{4,2,9}$	-0.3751
		(0.1305)		(0.0925)
$HISPANIC$			$B_{4,2,10}$	-0.4627
				(0.0928)
AGE_{nt}	$B_{4,1,6}$	-0.1468	$B_{4,2,11}$	-0.0824
		(0.0268)		(0.0147)
S_{nt}			$B_{4,2,12}$	-0.1038
				(0.0261)
$AFQT$	$B_{4,1,7}$	0.0058	$B_{4,2,13}$	0.0100
		(0.0027)		(0.0017)
$\ln(\eta)$	$B_{4,1,8}$	-0.6327	$B_{4,2,14}$	-0.4418
		(0.0877)		(0.0642)
$\ln(\mu)$	$B_{4,1,9}$	-0.2598	$B_{4,2,15}$	-0.5451
		(0.1478)		(0.1068)

TABLE 7.1: Marginal Effects Probability of Grade Promotion

$$F(x_{nt}) = (1-d^h_{nt})\frac{\exp(x'_{nt}B_{41})}{1+\exp(x'_{nt}B_{41})} + d^h_{nt}\frac{\exp(x'_{nt}B_{42})}{1+\exp(x'_{nt}B_{42})}$$

	$d^h_{nt}=0$		$d^h_{nt}=1$	
Variable	Parameter	Estimate	Parameter	Estimate
Time Use Variables				
s_{nt}	$B_{4,1,1}$	0.0005	$B_{4,2,1}$	0.0008
s^2_{nt}			$B_{4,2,2}$	-0.26E-06
h_{nt}			$B_{4,2,3}$	-0.0001
h^2_{nt}			$B_{4,2,4}$	0.02E-06
Enrollment Variables				
d^s_{nt-1}			$B_{4,2,5}$	0.0915
GRADE 11	$B_{4,1,2}$	0.1136	$B_{4,2,6}$	0.0717
GRADE 12	$B_{4,1,3}$	0.1109	$B_{4,2,7}$	0.2235
Participation Variables				
d^h_{nt-1}			$B_{4,2,8}$	0.0487
d^h_{nt-2}	$B_{4,1,4}$	0.0542		
Socio-Economic Variables				
$BLACK$	$B_{4,1,5}$	-0.0449	$B_{4,2,9}$	-0.0837
$HISPANIC$			$B_{4,2,10}$	-0.1032
AGE_{nt}	$B_{4,1,6}$	-0.0365	$B_{4,2,11}$	-0.0184
S_{nt}			$B_{4,2,12}$	-0.0232
$AFQT$	$B_{4,1,7}$	0.0011	$B_{4,2,13}$	0.0022
$\ln(\eta)$	$B_{4,1,8}$	-0.1247	$B_{4,2,14}$	-0.0985
$\ln(\mu)$	$B_{4,1,9}$	-0.0508	$B_{4,2,15}$	-0.1216

TABLE 8

Sample Averages of Nonparametric Estimates

Variable	Sample Mean	Sample Std. Dev.	Variable	Sample Mean	Sample Std. Dev
p_{nt0}	0.1197	0.2145	$\frac{\partial p_0(\Psi_{nt1}^{(1)})}{\partial h_{nt}}$	-0.1988	2.0533
p_{nt1}	0.7076	0.3427	$\frac{\partial p_0(\Psi_{nt1}^{(2)})}{\partial h_{nt}}$	-0.3520	5.2983
p_{nt2}	0.0489	0.1303	$\frac{\partial p_0(\Psi_{nt4}^{(1)})}{\partial h_{nt}}$	-0.6092	4.5189
p_{nt3}	0.1232	0.2307	$\frac{\partial p_0(\Psi_{nt5}^{(1)})}{\partial h_{nt}}$	-0.5044	5.2893
$p_0(\Psi_{nt0}^{(1)})$	0.3870	0.2398	$\frac{\partial p_0(\Psi_{nt2}^{(1)})}{\partial s_{nt}}$	-0.0391	4.1811
$p_0(\Psi_{nt0}^{(2)})$	0.5709	0.1835	$\frac{\partial p_0(\Psi_{nt3}^{(1)})}{\partial s_{nt}}$	0.4081	6.9457
$p_0(\Psi_{nt1}^{(1)})$	0.0995	0.1503	$\frac{\partial p_0(\Psi_{nt4}^{(1)})}{\partial s_{nt}}$	0.8412	5.5360
$p_0(\Psi_{nt1}^{(2)})$	0.3659	0.2446	$\frac{\partial p_0(\Psi_{nt5}^{(1)})}{\partial s_{nt}}$	-0.5360	6.3767
$p_0(\Psi_{nt2}^{(1)})$	0.0283	0.1095			
$p_0(\Psi_{nt3}^{(1)})$	0.2616	0.3736			
$p_0(\Psi_{nt4}^{(1)})$	0.0370	0.1504			
$p_0(\Psi_{nt5}^{(1)})$	0.1436	0.3166			

TABLE 9

Psychic Value of School Attendance

$u_0(d^s_{nt}, x_{nt}) = d^s_{nt} x'_{nt} B_5$

Variable	Parameter	Estimate	Std.Err.
Constant	B_{50}	-20.8502	10.0810
S_{nt}	B_{51}	3.6935	1.6900
S^2_{nt}	B_{52}	-0.0654	0.0619
$AGE_{nt} \times S_{nt}$	B_{53}	-0.0635	0.0093
$BLACK$	B_{54}	1.4361	1.3736
$HISPANIC$	B_{55}	0.0667	1.8812
$AFQT$	B_{56}	0.0165	0.0343

TABLE 10

Fixed Utility of Labor Force Participation

$u_1(d^h_{nt}, x_{nt}) = d^h_{nt} x'_{nt} B_6$

Variable	Parameter	Estimate	Std.Err.
Constant	B_{60}	-0.8174	2.3807
E_{nt}	B_{61}	1.2834	1.2741
E^2_{nt}	B_{62}	-0.0270	0.2294
$AGE_{nt} \times E_{nt}$	B_{63}	-0.0645	0.0176
$BLACK$	B_{64}	-0.4961	1.4026
$HISPANIC$	B_{65}	-0.0351	2.4603

TABLE 11

Utility of Leisure and the CRRA parameter.

$u_2(x_{nt}, l_{nt}, l_{nt-1}, l_{nt-2}) = l_{nt} x'_{nt} B_7 + \sum_{i=0}^{2} \delta_i l_{nt-i} l_{nt}$

Variable	Parameter	Estimate	Std.Err.
l_{nt}	B_{70}	0.0043	0.0114
$AGE_{nt} \times l_{nt}$	B_{71}	-0.0009	0.0010
$AGE_{nt}^2 \times l_{nt}$	B_{72}	0.27E-04	0.24E-04
$BLACK \times l_{nt}$	B_{73}	0.0009	0.0008
$HISPANIC \times l_{nt}$	B_{74}	0.0003	0.0021
l_{nt}^2	δ_0	-0.58E-07	0.68E-07
$l_{nt} l_{nt-1}$	δ_1	2.87E-07	1.15E-07
$l_{nt} l_{nt-2}$	δ_2	3.86E-07	0.11E-07
CRRA parameter	α	0.1067	0.0060

TABLE 12: Results from baseline simulation by race.

Age	Education		Experience		Hours		Wages	
	Actual	Sim.	Actual	Sim.	Actual	Sim.	Actual	Sim.
				White				
20	11.96	10.37	3.32	2.90	1257	1708	4.89	3.77
25	13.16	12.21	6.96	5.19	1957	1812	9.37	6.71
30	13.52	13.43	10.70	7.50	2198	2092	13.77	11.57
35		14.37		9.93		2338		15.85
				Black				
20	11.71	9.69	2.67	2.65	1129	1521	4.35	3.35
25	12.36	10.90	5.90	4.61	1830	1711	7.38	5.73
30	12.53	11.58	9.62	6.52	1963	2023	10.36	8.84
35		11.91		8.60		2275		11.67
				Hispanic				
20	11.33	9.69	3.04	2.84	1320	1773	5.00	3.82
25	11.99	10.89	6.71	5.03	1817	1960	9.15	6.57
30	12.28	11.56	10.57	7.20	2107	2219	12.26	10.03
35		11.90		9.61		2403		13.20

TABLE 13: Effect of cash subsidy to students who do not work.

Age	Education		Experience		Hours		Wages	
	Base	Pol 1	Base	Pol 1	Base	Pol 1	Base	Pol 1
	White							
20	10.37	10.38	2.90	2.88	1708	1709	3.77	3.77
25	12.21	12.24	5.19	5.14	1812	1810	6.71	6.71
30	13.43	13.48	7.50	7.41	2092	2094	11.57	11.68
35	14.37	14.44	9.93	9.78	2338	2336	15.85	16.08
	Black							
20	9.69	9.71	2.65	2.63	1521	1522	3.35	3.35
25	10.90	10.95	4.61	4.55	1711	1708	5.73	5.72
30	11.58	11.63	6.52	6.41	2023	2020	8.84	8.85
35	11.91	11.98	8.60	8.41	2275	2274	11.67	11.78
	Hispanic							
20	9.69	9.71	2.84	2.83	1773	1771	3.82	3.81
25	10.89	10.92	5.03	4.99	1960	1958	6.57	6.56
30	11.56	11.61	7.20	7.10	2219	2222	10.03	10.02
35	11.90	11.96	9.61	9.44	2403	2402	13.20	13.21

TABLE 14: Effects of mandatory increases in time spent on school activities.

Age	Education		Experience		Hours		Wages	
	Base	Pol 2	Base	Pol 2	Base	Pol 2	Base	Pol 2
	White							
20	10.37	10.66	2.90	2.89	1708	1722	3.77	3.78
25	12.21	12.95	5.19	5.09	1812	1827	6.71	7.05
30	13.43	14.78	7.50	7.23	2092	2185	11.57	14.99
35	14.37	16.52	9.93	9.38	2338	2412	15.85	23.28
	Black							
20	9.69	9.96	2.65	2.64	1521	1528	3.35	3.35
25	10.90	11.62	4.61	4.52	1711	1704	5.73	5.84
30	11.58	12.68	6.52	6.35	2023	2070	8.84	10.50
35	11.91	13.38	8.60	8.39	2275	2325	11.67	15.15
	Hispanic							
20	9.69	9.76	2.84	2.84	1773	1771	3.82	3.81
25	10.89	10.98	5.03	5.04	1960	1958	6.57	6.58
30	11.56	11.66	7.20	7.23	2219	2222	10.03	10.07
35	11.90	12.00	9.61	9.67	2403	2402	13.20	13.24

TABLE 15: Equating school quality.

Age	Education		Experience		Hours		Wages	
	Base	Pol 3	Base	Pol 3	Base	Pol 3	Base	Pol 3
	Black							
20	9.69	10.03	2.65	2.65	1521	1501	3.35	3.34
25	10.90	11.62	4.61	4.60	1711	1654	5.73	5.82
30	11.58	12.60	6.52	6.65	2023	2000	8.84	9.80
35	11.91	13.20	8.60	8.71	2275	2297	11.67	13.47
	Hispanic							
20	9.69	9.97	2.84	2.86	1773	1771	3.82	3.82
25	10.89	11.39	5.03	5.11	1960	1958	6.57	6.63
30	11.56	12.21	7.20	7.40	2219	2222	10.03	10.35
35	11.90	12.68	9.61	9.94	2403	2402	13.20	13.67

TABLE 16: Effects of mandatory increases in time spent on school activities after equating school quality.

Age	Education		Experience		Hours		Wages	
	Base	Pol 4	Base	Pol 4	Base	Pol 4	Base	Pol 4
	Black							
20	9.69	10.28	2.65	2.63	1521	1509	3.35	3.34
25	10.90	12.38	4.61	4.49	1711	1648	5.73	6.06
30	11.58	13.96	6.52	6.26	2023	2118	8.84	13.50
35	11.91	15.26	8.60	8.31	2275	2393	11.67	20.91
	Hispanic							
20	9.69	10.35	2.84	2.84	1773	1737	3.82	3.81
25	10.89	12.29	5.03	5.04	1960	1884	6.57	6.82
30	11.56	13.58	7.20	7.25	2219	2215	10.03	12.45
35	11.90	14.56	9.61	9.70	2403	2432	13.20	17.05

Figure 1: Changes in Shadow Price of Consumption $\Delta((1-\alpha)^{-1}\ln\lambda_t)$

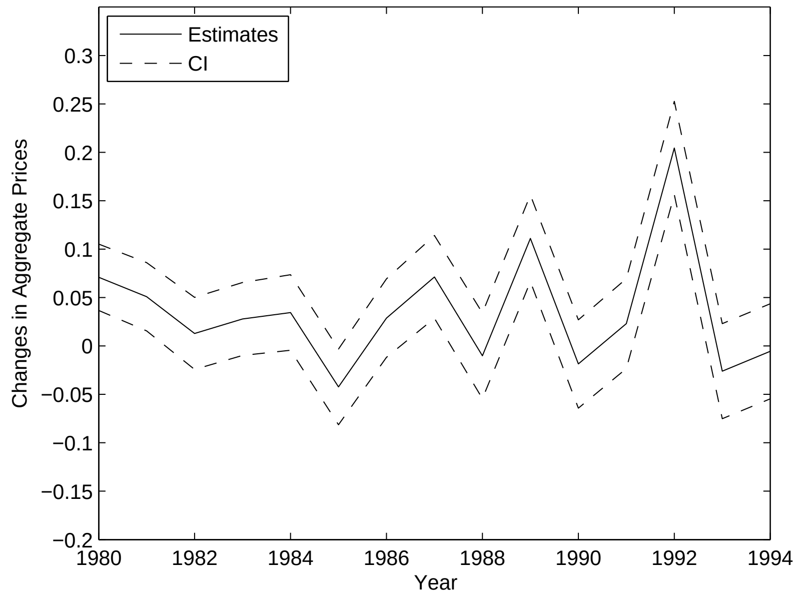

Figure 2: Changes in Unskilled Aggregate Wage $\Delta(\ln\omega_{t1})$

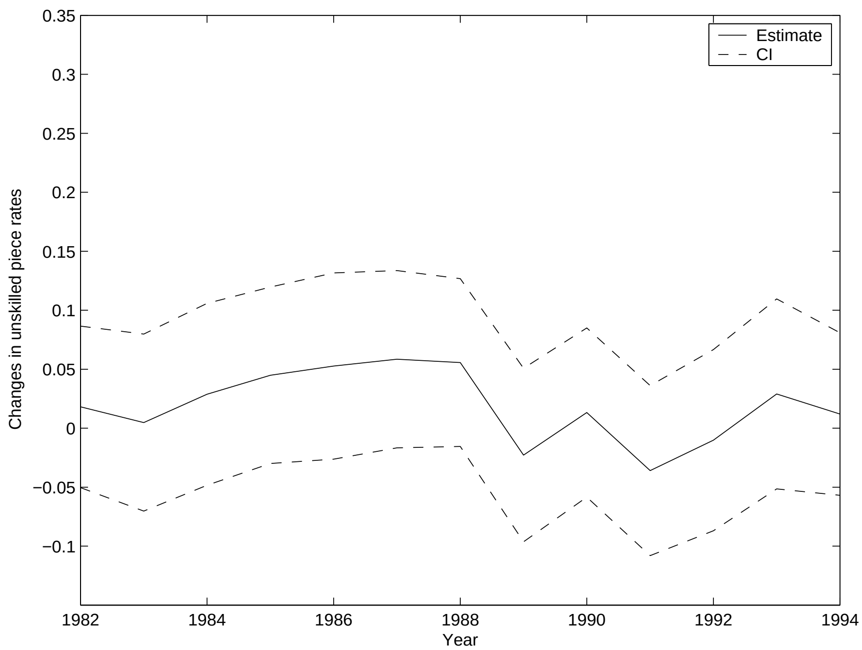

Figure 3: Changes in Skilled Aggregate Wage $\Delta(\ln \omega_{t2})$