Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19103-6297

pier@ssc.upenn.edu
http://www.econ.upenn.edu/pier

PIER Working Paper 97-032

“Youth Employment and Academic Performance in High School”

by

Zvi Eckstein and Kenneth I. Wolpin

Youth Employment and Academic Performance in High School

Zvi Eckstein*

and

Kenneth I. Wolpin**

March, 1997

Abstract: A fundamental premise of Federal and State legislation that restricts the hours that minors can be employed while school is in session is that working may adversely affect school performance. In this paper, we develop and structurally estimate a sequential model of high school attendance and work decisions. Policy experiments based on the model's estimates indicate that even the most restrictive prohibition would have only a limited impact on the high school graduation rates of white males.

We are grateful for support from the National Science Foundation under grant no. SBR-9309636. We have received helpful comments from Tom MaCurdy.

*Tel Aviv University and Boston University
** University of Pennsylvania

I. Introduction

The Fair Labor Standards Act (FLSA), first passed by the U.S. Congress in 1938 and since strengthened by a series of amendments, is the main federal legislation regulating the use of child labor. The most severe restrictions apply to those minors under the age of 16. Specifically, minors under the age of 16 are permitted to work only in non-mining, nonhazardous, and non-manufacturing jobs and then only under conditions that do not interfere with their schooling or health. Minors under the age of 18 are prohibited from working in non-agricultural jobs that have been declared as especially hazardous. More importantly for our purpose, the FLSA also regulates the hours that minors can work. During times when schools are in session, 14- and 15-year-old children may be employed for no more than 18 hours weekly and for not more than three hours in any one day.[1,2] However, full-time work, up to a maximum of 8 hours per day and 40 hours per week, is permitted during periods when schools are not in session. There are no hours limitations in the FLSA, regardless of time of year, for 16- and 17-year-old minors.

Evidence on compliance with the FLSA regulations on weekly hours can be obtained from the 1979 youth cohort of the National Longitudinal Surveys (NLSY). In the week prior to the 1979 survey, while over 20 percent of 14-year-old white males who were attending high school worked five hours or more, only 4.2 percent worked at least 15 hours; among 15-year-old white males attending high school, 8.1 percent worked at least 15 hours. Thus, compliance was very high, *de facto*.[3,4] However, *de facto* compliance is high even for 16-year-olds who face no hours limitations, with about one-quarter working at least 15 hours in the week.

The aim of this paper is not to study whether federal and state limitations on the hours worked of the youngest minors are effective, i.e., whether their hours worked would be different if the law did not exist.[5] Rather, our goal is to assess one of the underlying premises for the legislation, namely that working while attending high school could adversely affect school performance. Hours worked is indeed quite substantial among high school attendees; in the same NLSY sample, average hours worked in the week prior to the 1979 survey increased monotonically from 2.5 for the 14-year-olds to 16.2 for the 18-year-olds, and among those who worked positive hours, it increased from 8.9 to 24.5. In light of this fact, we ask whether even more restrictive hours limitations would increase high school graduation rates and/or improve the grades of those who graduate.

There are several reasons to suspect that employment might cause reduced school performance. First, work clearly reduces the amount of time available for other activities, including time for studying outside of school.[6] Second, working may be arduous and have negative spillover effects on attentiveness and study time while in school. Evidence on the relationship between study time in and out of school and working is available from the 1981 round of the NLSY, which obtained information on the number of minutes spent studying in and out of school during the prior week.[7] Those youths who were attending high school and not working reported spending 78 minutes a week more studying in school and 43 minutes a week more studying outside of school, a total of two hours more per week than those who were attending school and also working.[8] Although consistent with the notion that an exogenous increase in time spent working reduces the amount of time spent studying, it is possible that the relationship is spurious; those who chose to work would have studied less even had they not worked. On the other hand, the negative correlation may be understated because of self-selection; those youths who were not attending high school in 1981, and had not previously graduated, could be individuals who would work and spend even less time studying than those who were attending. Solving the inferential problem requires that we consider the joint school attendance and work decisions.

Although study time is a reasonable indirect measure of performance, because the NLSY is longitudinal, it is possible to look at the relationship between working while attending school and ultimate success, namely graduating with a diploma. During the week prior to the survey, those white male youths who were attending grade nine in any of the 1979-1982 survey rounds and who later graduated from high school worked an average of 2.2 hours, but those who did not later graduate averaged 4.0 hours.[9] Although graduation rates were negatively related to working while attending grade nine, they were unrelated for grade 10, and positively related for grades 11 and 12.[10] Even had the relationship been unambiguous, the interpretation would still be unclear. Almost everyone does attend grade nine, so that sample selection is probably not an important problem; however, with foresight, those who perceive their chances of graduating as small might be more likely to work. By grades 11 and 12, the selection issue is more problematic and the positive relationship between working and graduation might reflect inherent differences in energy and motivation rather than the acquisition of affective skills gained from employment that are also useful in school.

These data illustrate the importance of understanding the decision process underlying work-school outcomes in interpreting the relationship between school performance and work. Previous work in this area, almost entirely by non-economists, has not been based on an explicit decision model and the empirical findings, as above, have ambiguous interpretations.[11] Most of the research by economists has been concerned with the impact of employment while in high school and future labor market success rather than school performance.[12] There has also been related research, also primarily by non-economists and thus not explicitly decision-theoretic, on the relationship between working and college performance.[13] An exception on both counts is the work by Ehrenberg and Sherman (1987) who estimate what can be thought of as a statistical representation of an approximation to a sequential optimization problem under uncertainty.[14]

In contrast to this literature and in light of the inferential problems described above, in this paper we formulate and estimate an explicit sequential decision model of high school attendance and work that captures in a stylized fashion the important institutional features of high school grade progression. Briefly, individuals accumulate credits (courses) towards graduation and receive course-specific performance grades ("A" - "F"). Grades are probabilistic, depending on the individual's history of performance (knowledge acquisition), the level of participation in the labor market (hours worked) and their known (to them) ability and motivation. Thus, graduation is a probabilistic outcome that can be influenced by work decisions. The labor market (randomly) offers up wages for part-time and full-time employment that depend also on some inherent skill "endowment" (potentially correlated with ability and motivation associated with school performance) as well as labor market experience. Working, in addition to potentially reducing school performance, directly reduces leisure time. The value of attending high school consists of both the perceived investment payoff to graduation and on a current consumption value which is random, each of which may differ among the population. Each period, the individual chooses a work-school arrangement that maximizes the present value of lifetime utility.

The paper also contributes to the literature more generally on solving and estimating discrete choice dynamic programming models. An important limitation in the estimation of this class of models is computational burden. Complexity is often avoided by reducing the scope of inquiry, for example, by reducing the size of the choice set. Modeling grade progression and work decisions through the period of high school attendance in the detail described above precludes the consideration of post-high school graduation decisions, for example, decisions about college attendance. However, in forward-looking decision models, such considerations cannot be avoided. The value attached to high school graduation clearly depends on future decisions. The approach we adopt circumvents this difficulty by recognizing that the expected value of high school graduation can be considered as a terminal condition for decisions during the high school period and its value can be estimated as an additional parameter of the model. In this way, we do avoid having to specify explicitly the decision model after high school graduation.

The model is estimated using data from the NLSY. The novel feature of the data that permits the model to explicitly account for credit accumulation and grades is the availability of high school transcripts for a large part of the sample. The estimation method combines the solution of the dynamic optimization problem with the maximization of a likelihood function that accounts jointly for annually observed work-schooling choices, wages, hours worked, credits earned and grades.[15]

Our results indicate that working while attending high school does reduce academic performance. However, the quantitative effects are small. Estimates of the behavioral model imply that implementing a policy that forced youths to remain in high school for five years or until they graduate, whichever comes first, without working would increase the number of high school graduates by slightly more than 2 percentage points (from 82 to 84.1 percent). In addition, the cumulative grade point average (gpa) of those youths who would have graduated regardless of the policy would rise by only .04 points (from 2.50 to 2.54). Our findings also indicate that dropping out of high school is confined to youths with specific traits: lower school ability and/or motivation, a lower expected value of a high school diploma, higher skills in the kinds of jobs that do not "require" a high school diploma, a higher value placed on leisure and a lower consumption value of attending school. Our estimates allow us to provide quantitative estimates of their importance.

The format of the paper is as follows. In the next section, we present the behavioral model. We briefly discuss the numerical solution method and the estimation method. Section III describes the data and presents descriptive statistics. The following section provides the estimates of the model and discusses their implications. Section V concludes.

II. Model

Each year the NLSY asks respondents who are not attending school at the interview date the main reason for their non-attendance. Approximately 30 percent of white males who had not graduated from high school at the time of the 1979-1982 survey rounds chose the response category "didn't like school" as the main reason. Aside from a residual "other" category (20 percent), the next most frequently cited reason was "offered a good job, chose to work" (14 percent). The category "suspended or expelled" was the next most cited response followed by "lack of ability, poor grades" (9 percent) and then by "financial considerations" (4 percent). In formulating a model that encompasses the decision to drop out of high school, we have attempted to incorporate most of these response categories, without necessarily giving credence to the responses themselves. Specifically, the model incorporates preferences for attending school, market opportunities and school related ability and/or motivation. In addition, although not contained in the list of NLSY response categories, in the model youths also consider the expected payoff to graduation. Although the model is presented as the decision problem of a single youth, an important feature of our empirical implementation is that youths may be heterogeneous with respect to these factors.

The Basic Structure

We consider the decision process of a youth who enters high school (grade level nine) at time t=1 and who has acquired a given stock of school-based knowledge, K_0, through prior formal schooling and through informal human capital investments. It is assumed that each year a high school attendee must take (exactly) five course credits and must accumulate a total of 20 credits to obtain a (regular) high school diploma.[16] Completion of high school, therefore, takes a minimum of four years.[17] A letter grade is received for each credit and is converted to a 0 - 4.0 numerical scale, where "A" = 4.0, "B" = 3.0, "C" = 2.0, "D" = 1.0, "F" = 0. A passing grade (A-D) represents incremental knowledge (and progression towards the degree, i.e., a credit is earned); total incremental knowledge from the beginning of high school is assumed to depend on the cumulative gpa calculated over passing grades and the cumulative number of credits earned, i.e., K_t-K_0=f(K_0, G_t, C_t), where K_t is knowledge acquired through attending school up to year t, G_t is the cumulative grade point average (gpa) achieved up to year t and C_t is the total number of credits earned up to year t. Completed grade levels (e_t) correspond to five-credit increments, i.e., e_t=8 if C_t<5, e_t=9 if 5 ≤ C_t<10, e_t=10 if 10 ≤ C_t<15, e_t=11 if 15 ≤ C_t<20, and e_t=12 (high school graduation) if C_t ≥ 20. In addition, we assume that high school must be completed within five (calendar) years of first entry.[18] Thus, failing a total of six or more credits at any stage or having dropped out of school for more than one year precludes ever graduating from high school.

At the beginning of each school year an individual who is still eligible for high school (t ≤ 5 and e_t<12) chooses whether to attend school, d_t^s = {0,1}, and/or whether to work in the labor market h_t hours per week. The number of hours worked per week can take on any of six discrete values: {0, 10, 20, 30, 40, 50}. Within this range, we distinguish between part- and full-time jobs. The first two positive hours correspond to part-time jobs, $h_t^p=\{10,20\}$ and the last three to full-time jobs $h_t^f=\{30,40,50\}$. In each period, an individual receives both a part- and a full-time job offer. However, there is only a partial choice of hours. With probability $\pi_t^h(j,k)$ the individual receives a part-time job offer of $h_t^p=j$ hours and a full-time job offer of $h_t^f=k$ hours. Because an offer of each type of job is received in each period, $\sum_j\sum_k\pi_t^h(j,k)=1$.

Hourly wage offers in part- and full-time work are determined by knowledge obtained in school and from skills obtained through work experience. In addition, there are idiosyncratic shocks to wage offers. Specifically, the hourly wage offer is $w_t^j=\exp(w^j(K_t,H_t))\exp(\epsilon_t^j)$ for j=p,f, where H_t is accumulated hours worked up to t. Recall that these wage functions pertain only to school completion levels less than 12 ($e_t<12$).

The choice set for someone who is eligible to attend high school consists of six mutually exclusive and exhaustive dichotomous (0-1) alternatives: $d_t^k=\{d_t^{sn}, d_t^{sp}, d_t^{sf}, d_t^{nn}, d_t^{np}, d_t^{nf}\}$, i.e., attend school and not work ($d_t^{sn}=1$), attend school and work part-time ($d_t^{sp}=1$), attend school and work full-time ($d_t^{sf}=1$), not attend school and not work ($d_t^{nn}=1$), not attend school and work-part time ($d_t^{np}=1$), and not attend school and work full-time ($d_t^{nf}=1$). A high school dropout who is ineligible to return to school (t>5) chooses among only the latter three alternatives: $d_t^k=\{d_t^{nn}, d_t^{np}, d_t^{nf}\}$. We do not explicitly model the post-graduation decisions of those who graduate within the five year period. We do take into account the expected value placed on graduation, subsuming the post-graduation school-work decision process that underlies it.

As given above, incremental knowledge over a school year depends on the number of credits earned during the year and by the associated gpa. These latter measures, gpa and credits, depend on factors known to the individual at the beginning of the year, such as cumulative credits earned at the start of the year, cumulative gpa at the start of the year and one's permanent level of motivation and ability. In addition, they are influenced by the student's choice of effort during the school year and by time-varying factors such as the quality of the instruction that were unknown at the start of the year. Student effort is assumed to vary (inversely) with hours worked in the labor market (h_t).[19]

Utility is assumed to be linear and additive in earnings plus (the value of) family transfers, non-work time, and school attendance.[20] Non-work time is valued at b_t^n dollars per hour and has a random time-varying component, i.e., $b_t^n = \bar{b}^n + \epsilon_t^n$. There are L total hours per week that can be divided between work and non-work. Attending school also has a psychic value, b_t^s, that differs depending on the person's work status and that varies over time stochastically, i.e., $b_t^{sj} = \bar{b}^{sj} + \epsilon_t^{sj}$ for j={n,p,f}.[21] The random elements (including the wage shocks) $\{\epsilon_t^j, \epsilon_t^{sj}\}$, j={n,p,f} are assumed to be jointly serially independent and joint normal, i.e., $N(0,\Omega)$. The stochastic elements at t are revealed at t, but are unknown before t.

Current period utility at time t (U_t) depends on the alternative (k) chosen, namely[22]:

$$
(1)\qquad
\begin{aligned}
U_t^{sn} &= b_t^n \cdot L + b_t^{sn} , \\
U_t^{sp} &= b_t^n \cdot (L - h_t^p) + b_t^{sn} + b_t^{sp} + w_t^p \cdot h_t^p , \\
U_t^{sf} &= b_t^n \cdot (L - h_t^f) + b_t^{sn} + b_t^{sf} + w_t^f \cdot h_t^f , \\
U_t^{nn} &= b_t^n \cdot L , \\
U_t^{np} &= b_t^n \cdot (L - h_t^p) + w_t^p \cdot h_t^p , \\
U_t^{nf} &= b_t^n \cdot (L - h_t^f) + w_t^f \cdot h_t^f .
\end{aligned}
$$

The individual at any time t is assumed to maximize the expected present discounted value of utility over an infinite horizon. Defining $V_t(S_t)$, the value function, to be this maximal expected present value at t, given the state space S(t) at t, and given the discount factor β,

$$(2)\quad V_t(S_t) = \max_{\{d_t\}} E\,[\sum_{\tau=t}^{\infty} \beta^{\tau-t} \sum_k U_t^k d_t^k \mid S_t]\ .$$

The state space consists of all aspects of the history known to the individual that affect current alternative-specific utilities or the probability distribution of future utilities. As the model is specified, S_t includes G_t, C_t, H_t, and the ϵ_t's. The maximization of (2) is achieved by choice of the optimal sequence of feasible control variables $\{d_t^k\}$ given current realizations of the stochastic components of the utilities.

The value function can be written as the maximum over alternative-specific value functions. It is useful to define two value functions as follows, one for the post-high school eligibility period ($t \geq 6$) and one for the eligibility period ($t \leq 5$):

$$(3)\quad \begin{aligned} V^1(S_t) &= \max[V^{nn}(S_t),\ V^{np}(S_t),\ V^{nf}(S_t)] \quad, t\geq 6\ , \\ V_t^2(S_t) &= \max[V_t^{sn}(S_t),\ V_t^{sp}(S_t),\ V_t^{sf}(S_t),\ V_t^{nn}(S_t),\ V_t^{np}(S_t),\ V_t^{nf}(S_t)] \quad ,t\leq 5 \quad . \end{aligned}$$

These alternative-specific value functions satisfy the Bellman equation (Bellman, 1957). For $t \geq 6$, i.e., for periods in which an individual (who did not graduate from high school) is no longer eligible to attend high school, these value functions are given by:

$$(4)\quad V^k(S_t) = U_t^k + \beta E(V^1(S_{t+1})) \mid d_t^k=1, S_t)\ ,$$

where k={nn, np, nf}. Notice that the value functions are stationary.[23] Expectations in (4) are taken over the stochastic components of the alternative-specific utilities and over the joint part- and full-time offered hours distribution.

To characterize the value functions during the high school eligibility period, define

$\pi(c_t,g_t \mid S_t)$ to be the probability of earning c_t credits with grade point average g_t while attending school between time t and t+1 conditional on the state space at time t.[24] Although g_t is calculated only on passing grades, we adopt the convention that $g_t=0$ if $c_t=0$. In addition, let V^D be the value of high school graduation. Then for $t \le 5$, the alternative-specific value functions are:

$$
\begin{aligned}
V_t^k(S_t) = U_t^k + \beta\Big[\sum_{c=0}^{5}\sum_{g=0}^{4} \pi(c_t,g_t \mid S_t,d_t^k=1)\{ & I(t<5,e_{t+1}<12)E[V_{t+1}^2(S_{t+1}) \mid d_t^k=1,S_t] \\
& + I(t=5,e_{t+1}<12)E[V^1(S_{t+1}) \mid d_t^k=1,S_t] \\
& + I(e_{t+1}=12)E[V^D \mid d_t^k=1,S_t]\ \}\Big] ,
\end{aligned} \tag{5}
$$

where $I(\cdot)$ is an indicator function equal to one if the term inside the parentheses is true and zero otherwise. Over this period, the individual chooses among the six alternatives, k = {sn, sp, sf, nn, np, nf}. These value functions are non-stationary, i.e., they explicitly depend on t, because of the finite horizon over which a high school diploma must be earned.[25]

Solution Method

The model does not admit to analytical solution, but can be numerically solved in a straightforward manner. The numerical complexity arises because the calculation of the value functions requires high-dimensional integrations. We follow the procedure in Keane and Wolpin (1994), using Monte Carlo integrations to evaluate the integrals that appear in (4) and (5).[26] In solving the stationary component of the model, it is assumed that wage offers are constant after H=50,000, i.e., after accumulating the equivalent of 25 full-time (2,000 hours) years of working. This assumption provides a "terminal" stationary value for Emax[$V^{nn}(S)$, $V^{np}(S)$, $V^{nf}(S)$] at H=50,000. Stationary value functions are then solved recursively for H<50,000 and pertain to all

$t>5$. The $E_t\max[V^{sn}(S), V^{sp}(S), V^{sf}(S), V^{nn}(S), V^{np}(S), V^{nf}(S)]$ functions for $t \leq 5$ are solved recursively from t=5 as in any finite horizon model.

Estimation Method

Having solved for the Emax functions, equations (4) and (5) look like the indirect utilities associated with any panel data multinomial choice problem. There are three complications: (i) the errors as we have modeled the problem are not all additive, (ii) in addition to the school-work choice, we observe the wage rate for those that work in a particular period, the hours that they work, and their grades if they attend school, and (iii) the wage rate appears to be measured with error.[27] Assume in what follows that the measurement error is multiplicative, i.e., $\ln w_t^{jo} = w^j(K_t,H_t) + \epsilon_t^j + \eta_t^j$, with η_t^j $N(0,(\sigma_\eta^j)^2$ for j=p,f, where w_t^o signifies the observed wage.

At any time t, denote the vector of outcomes as $O_t=\{d_t^k, w_t^o, h_t, c_t, g_t\}$. The likelihood function for a sample of N individuals each observed from period $t=1,\dots,t_n$ is given by:

$$(6) \quad \prod_{n=1}^{N} \Pr(O_{1n}, O_{2n}, \ldots\ldots, O_{t_n n} | K_0).$$

In general, the computation of (6) would require the calculation of multiple integrals of dimension at least equal to the number of periods times the number of alternatives.[28] Given the assumption of joint serial independence among the vector of shocks, the likelihood function can be written as the product of within-period outcome probabilities each of which is an integral of dimension equal to the number of alternatives.

To illustrate the calculation of the likelihood, it is sufficient to consider a specific outcome at some period. Suppose that the following is observed at period t: the individual chooses to attend school and work part-time ($d_t^{sp}=1$), reports receiving a part-time wage rate of

w_t^{po}, works h_t^p hours, and earns c_t credits during the school year with a grade point average g_t. Further, assume that the individual entered the period having previously worked H hours and having earned a total of C credits with a cumulative gpa of G. The probability of that outcome is:

$$\Pr(d_t^{sp}=1, w_t^{po}, h_t^p, c_t, g_t \mid G_t, C_t, H_t) = \pi(c_t, g_t \mid h_t^p, G_t, C_t)\cdot\pi(h_t^p)\cdot \tag{7}$$
$$\int_{w_t^p} \Pr(d_t^{sp}=1 \mid w_t^p, h_t^p, G_t, C_t, H_t)\cdot \Pr(w_t^p, w_t^{po} \mid G_t, C_t, H_t).$$

We calculate the second line of (7), the joint probability of choosing $d_t^{sp}=1$ and of observing a reported wage w_t^{po}, by a smoothed simulator.[29] The integration over the true wage in (7) is necessary because the choice depends on the true wage and we do not observe it. Other probability statements are calculated similarly.

Notice that the entire set of model parameters enter the likelihood through the choice probabilities that are computed from the solution of the dynamic programming problem. Subsets of parameters enter through other structural relationships, e.g., wage offer functions. The estimation procedure, i.e., the maximization of the likelihood function, iterates between the solution of the dynamic program and the calculation of the likelihood.

Unobserved Heterogeneity

It is unlikely that either K_0 (or school ability/motivation), “endowments” of market skills, preferences for leisure or school, or expected valuations of graduating are the same for everyone at entry into high school. To account for such unobserved heterogeneity, we assume that parameters representing such attributes, e.g., parameters of the wage functions, may differ in the population. We assume that there are M types in the population, each comprising π_j fraction of the population (e.g., see Heckman and Singer (1984), Keane and Wolpin (1994)). Although we present the parameterizations below, the generic modification in the likelihood function in this case is given by:

$$(8) \quad \prod_{n=1}^{N} \sum_{m=1}^{M} \Pr(O_{1n}^{m}, O_{2n}^{m}, \ldots\ldots, O_{t_n n}^{m} \mid \text{type}=m) \cdot \pi_m .$$

The likelihood function is a weighted average of type-specific likelihoods. Thus, to calculate (8), the dynamic program must be solved for each type.

Further Parameterizations

To complete the specification for estimation purposes, it is necessary to adopt explicit forms for the following functions: grades, wage offers, hours offers, and the values of leisure, school attendance and graduation. The functional forms that are adopted below do not represent *a priori* specifications alone, but rather are the result of an iterative specification search based on assessing the fit of the model to summary statistics in the data. That is, parameters were added as estimation proceeded in order to improve the fit of the model to the data.

Grade functions: Grade functions take an ordered logit form. The probability of receiving a grade of "F" in a single course is assumed to depend at period t on cumulative grade point average (G), cumulative credits earned as summarized by grade level completed (e), the number of years in high school (t), whether any credits have yet been earned (I(C=0)), on current hours worked (h) and on type. Allowing for type-specific differences in ability, motivation and initial knowledge, the grade functions are specified as follows:

$$\pi_t^F = 1/(1 + \exp(\theta^F)) \ ,$$

where

$$\theta_t^F = \sum_m \theta_{0m}^F I(\text{type}=m) + \theta_1 G_t + \theta_2 e_t + \theta_3 t + \theta_4 I(C_t=0)$$

$$+ \theta_5 I(h_t \geq 500) + \theta_6 I(h_t \geq 1000) + \theta_7 I(h \geq 1500) \ .$$

The probability of receiving a "D" in a course is given by

$$\pi_t^D = 1/(1+\exp(\theta^F+\theta^D)) - \pi_t^F,$$

where

$$\theta^D = \sum_m \theta_{0m}^D I(\text{type}=m) \ .$$

Similarly, the other grade probabilities follow the ordered logit formulation with $\theta^j = \sum_m \theta_{0m}^j$ for j=C,B,A. The joint credit-grade distribution necessary for the solution of the dynamic programming problem and for the calculation of the likelihood is easily derived from the single-course grade probabilities. Notice that if $\theta_5=\theta_6=\theta_7=0$ there is no effect of working while in school on school performance.

Wage-offer functions: The part-time and full-time wage functions are linear in accumulated hours worked (up to 50,000 hours).[30] Wage offers may also differ by unobserved type.

$$w_t^j = \sum_{m=1}^{M} \alpha_{0m}^j \cdot I(\text{type}=m) + \alpha_1^j H_t \cdot I(H_t < 50{,}000) + \alpha_1^j \cdot 50{,}000 \cdot I(H_t \geq 50{,}000)$$

for j = p,f.

Hours-offer functions: The joint part-time and full-time hours offer function is assumed to be different during the high school eligibility years than during the post-eligibility years. We also allow age to affect hours offers during the high school years, reflecting possibly both legal constraints and employer preferences. Specifically, the hours functions are given by the multinomial logit form:

$$\pi_t^h(j,k) = \exp(\theta_{0jk}^h + \theta_{1jk}^h * t) / \sum_{j'} \sum_{k'} \exp(\theta_{0j'k'}^h + \theta_{1j'k'}^h * t) \quad \text{for } t \leq 5,$$

$$\pi_t^h(j,k) = \exp(\theta_{0jk}^h) / \sum_{j'} \sum_{k'} \exp(\theta_{0j'k'}^h) \quad \text{for } t \geq 6 \ .$$

In this context, there is also assumed to be a job-finding cost if one was neither employed nor in school. A cost is borne upon choosing either work alternative while not attending school ($d_t^{pp}=1$ or $d_t^{ff}=1$) if the non-work (non-school) alternative was chosen in the previous period ($d_{t-1}^{nn}=1$). Utility in these alternatives, as given in (1), is reduced by $c^{pp} \cdot I(d_{t-1}^{nn}=1)$ if $d_t^{np}=1$ and by $c^{ff} \cdot I(d_{t-1}^{nn}=1)$ if $d_t^{nf}=1$.

Value of leisure: The hourly value of leisure is assumed to differ by type. In addition, the value of leisure may differ in the high school eligibility years from the post-eligibility years, i.e.,

$$\overline{b}^n = \sum_{m=1}^{M} b_{0m}^n I(\text{type}=m) + b_1^n I(t \leq 5) .$$

Net consumption value of school attendance: As seen in (1), the utility value of attending school is assumed to depend explicitly on work status:

$$\overline{b_t^s} = \overline{b_t^{sn}} \cdot I(d_t^{sn}=1) + [\overline{b_t^{sp}} + \overline{b_t^{sn}}] \cdot I(d_t^{sp}=1) + [\overline{b_t^{sf}} + \overline{b_t^{sn}}] \cdot I(d_t^{sf}=1)$$

This psychic value may reflect a number of positive or negative factors including the value attached to the social aspects of attending school and the value attached to learning per se as well as the effort it entails. We allow for unobserved type heterogeneity, and for the value of attendance to change with the level of schooling and with time since first entry into high school. In addition, we include a cost to dropping out and returning to high school as well as a cost to attending high school in the fifth year of eligibility. Both of these reflect the fact that in either case the individual is no longer synchronized with his entry cohort. Working either part- or full-time may reduce the psychic value of attendance if it inhibits participation in social activities (e.g., extra-curricular programs) or if it implies increased effort in learning. Such an effect may depend on years since entry (age), and on whether the individual has been able to adjust to the joint activity as measured by work participation in the previous period. Thus (letting $d_t^s=1$ indicate school attendance in any work status),

$$\overline{b}_t^{sn} = \sum_{m=1}^{M} b_{0m}^{sn} I(\text{type}=m) + b_1^{sn} t + b_2^{sn} e + b_3^{sn} I(d_{t-1}^{s}=1) + b_4^{sn} I(t=5) \ ,$$

$$\overline{b}_t^{sp} = b_0^{sp} + b_1^{sp} t + b_2^{sp} I(t=1) + b_3^{sp} I(d_{t-1}^{sp}=1) + b_4^{sp} I(d_{t-1}^{sf}=1) \ ,$$

$$\overline{b}_t^{sf} = b_0^{sf} + b_1^{sf} t + b_2^{sf} I(t=1) + b_3^{sf} I(d_{t-1}^{sf}=1) + b_4^{sf} I(d_{t-1}^{sp}=1) \ .$$

Value of high school graduation: The expected present value of the utility of graduating form high school, which includes not only monetary but also psychic returns, is allowed to differ by type and to depend on achieved cumulative grade point average upon graduation and upon work experience gained while attending high school, namely

$$V^D = \sum_{m=1}^{M} \gamma_{0m} I(\text{type}=m) \ + \ \gamma_1 G^D \cdot I(e=12) \ + \ \gamma_2 H^D \cdot I(e=12).$$

In solving the optimization problem, the value attached to graduating is optimally updated. Because credits earned, cumulative grade point average and hours worked evolve over time, the forecasts of cumulative gpa and hours worked at the time of graduation, G^D and H^D respectively, as well as of the likelihood of graduating changes as decisions are made and stochastic shocks are realized. Notice that this representation can be viewed as an approximation to the expected value function that would arise from optimal decisions made after graduation.[31]

III. Data

The data used in the analysis are from the 1979 youth cohort of the National Longitudinal Surveys of Labor Market Experience (NLSY). The NLSY consists of 12,686 individuals, approximately half of them male, who were 14 to 21 years of age as of January 1, 1979. The sample consists of a core random sample and oversamples of blacks, Hispanics, and members of the military. This analysis is based on the white males in the core random sample who were less than age 15 as of October 1, 1977 and who had ever attended high school (at least grade level nine). There are 702 white males who meet these criteria (out of a total of 2439).[32] Interviews in the NLSY have been conducted annually since 1979, and we follow each individual from the first year they enter grade level nine through either their year of high school graduation or, if they do not graduate, through their last interview (1991 at the latest).

The NLSY collects schooling and employment data as an event history retrospectively back to the preceding interview. Employment data include the beginning and ending dates (to the calendar week) of all jobs (employers), all gaps in employment within the same job, usual hours worked on each job, and the usual rate-of-pay on each job. In the first interview, employment data were collected back to January 1, 1978. Given the age restriction imposed on our analysis sample, we effectively have a complete employment profile for each sample youth.

Schooling data include the highest grade attended and completed at each interview date, monthly attendance in each calendar month beginning in January 1980, school-leaving dates, and the dates of diplomas and degrees. In addition, the NLSY obtained and coded high school transcripts for much of the sample. Of the 702 youths in our subsample, 564 had usable transcript data.[33] The transcript data report all of the courses taken in high school, standardized

(Carnegie) credits for each course, a grade level at which the course was taken (9 through 12) , a course number and a grade for each course based on a 0 to 4.0 scale.[34]

For those who graduate in four years, which is the vast majority of youths, we essentially ignore the actual pattern of credit accumulation. Recall that in the model a student takes exactly five courses per year and needs to obtain a passing grade in 20 courses over at most five years in order to earn a diploma. Therefore for this group, we fixed the number of credits earned at five for each of the four years. Notice that in following this rule, those who fell behind but were able to make up failed credits and still graduate in four years are counted as having a normal grade progression.

There are a number of serious difficulties in using the transcript data for those who did not progress smoothly. Most problematic is that calendar dates in which courses were taken are not reported. Thus, some transcripts report a large number of courses taken at a single grade level that obviously span more than one school year. Courses are supposed to be listed chronologically, but it is unclear from the course numbers that this is the case. In addition, grade level designations are often meaningless. Some transcripts will show a youth advancing through several grade levels without ever having received a passing grade in a single course. Thus, we hand-edited each transcript of those who did not graduate within four years of entry. We attempted to adhere to some basic rules in establishing credits earned. However, it was impossible to codify the rules in a way that would avoid having to make judgements on a transcript-by-transcript basis.

Loosely speaking, our procedure was as follows. We began with grade level 9 and usually simply counted the number of credits earned over all courses with unique course numbers for which passing grades were received. However, we normally did not allow a student to accumulate more than one course credit in a particular subject in a single school year with additional such credits allocated to the next year. (Recall that we are dealing almost entirely with students who did not graduate from high school.) If the number of credits earned over these courses totaled five or more, we assigned five credits to grade level 9 preliminarily. For those allotted five credits for grade level 9, we then looked at their grades in the major subjects, i.e., English, Math, Social Studies, and Science. For each of those courses with a unique course number in which a failing grade was reported, we subtracted its associated credits (a maximum of one credit was deducted for each subject). If failing grades were received in all of the four major subjects, zero credits were assigned for that year. A course for which a failing grade was reported and that was repeated at the same grade level was treated as course work taken in the following school year, as were failed courses with the same course numbers that were attached to the next higher grade level. We then repeated this procedure for the repeated grade level 9 courses and the grade level 10 courses, assigning credits earned to year two of high school. We repeated this procedure until all of the of the courses on the transcript had been exhausted.

This procedure determined the sequence of credits earned. The grade level assigned to a particular year was based on accumulated credits as per the model. Thus, a youth who earned three credits in year one of high school was assigned a completed grade level of eight upon entering year two of high school. If the youth earned two or more credits in the second year, then grade level would increase by one upon entering the third year of high school; if not, the grade level would remain at eight.

The calendar-time placement of the sequence of credits earned, and thus grade levels completed, was obtained directly from the main survey school enrollment data. From this match, it is apparent that the grade-level progression calculated from the transcript data as above does not match self-reported grade levels in the main survey. If self-reported grade levels reflect actual school assignments, it is evident that our accounting of credits earned does not correspond to the standards for progression practiced by high schools. Self-reported highest grade completed is considerably higher on average than is the same measure based on our credit accounting. For those who do not graduate from high school, the average level of schooling completed prior to their last year of attendance as self-reported is 10.0, but it is only 9.4 based on the accumulated credits as we have calculated from the transcript data.

Gpa is also calculated from the transcript data and is based on only the five major subjects (the four above plus foreign language) that are taken in the school year. To conform to the model of knowledge acquisition, gpa is based only on passing grades. We discretized gpa into eight categories: .5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0. Those gpa values between .25 and .75 were assigned a value of .5, those between .75 and 1.25 were assigned 1.0, etc.

The number of hours worked is calculated differently for those in school during the year than for those not in school.[35] For those who did not attend school at any time during the school year, say from time t to t+1, we calculate hours worked per week based on employment data between October 1 of year t and September 30 of year t+1.[36] For those who were attending school during the year, we base the calculation of hours worked per week on data covering the period between October 1 and June 30. Although we could have calculated summer work hours separately, the model does not accommodate separate summer work decisions and we wanted, most importantly, to account accurately for school-year weekly work hours. The discretized values of hours worked per week are obtained from the reported continuous hours as follows: $h=0$ if $h<5$, $h=10$ if $5 \le h<15$, $h=20$ if $15 \le h<25$, $h=30$ if $25 \le h<35$, $h=40$ if $35 \le h<45$, and $h=50$ if $h \ge 45$.[37] Hourly wage rates were calculated by cumulating weekly earnings over the relevant period (12 months for non-school attendees and nine months for school attenders) and dividing by hours worked. Annual earnings are calculated by multiplying the average wage times average hours times 50 (weeks).

Descriptive Statistics

Each observation begins at first entry into the ninth grade and ends either at high school graduation (four or five years later) or the last time the youth was surveyed up to the 1991 interview. The longest period over which we observe an individual who dropped out of high school is thirteen years. Table one shows the choice distribution and the number of observations in each period. By definition, everyone in the sample is attending school in period one. In grade nine 82.6 percent of the youths in our sample attend school without working (hours worked per week less that five on average), 15.3 percent attend school and work part-time (5 to 24 hours per week on average) and only 2.1 percent work full-time while attending school (25 or more hours per week on average). Working while simultaneously attending school rises rapidly with years since high school entry (and thus with age), with part-time work while attending school reaching 40 percent of all youths and full-time work while attending almost 10 percent in the third year (grade 11 with normal progress). In that year, 6 percent of youths were not attending school, with about two-thirds of them working. By the fourth year, only 30 percent of the youths are in school and not working, with almost one-fifth working full-time while attending. In addition, 11.5 percent are not attending school, and of these about 40 percent are working full-time and 34 percent part-time. Of those youths who attend school for at least four years, 78 percent worked at least one of those years and over one-quarter worked at least three years.

In year four, 421 of the 516 youths who were in the sample for at least four years (81.6 percent) graduated from high school. Of the 92 non-graduates who were observed in year five, seventeen (18.5 percent) attended school in year five. Of those 17 who continued, six (35 percent) graduated in year five. Thus, about 83 percent of the sample who are observed until graduation or, otherwise, for at least five years, graduated from high school, with 98.6 percent of them graduating in four years.[38] Returning to high school after dropping out is a rare occurrence. In our sample, only three youths dropped out and returned (one graduated). However, this is an understatement of the phenomenon both because high school dropouts are under-represented in the transcript sample and because transcripts are themselves more likely to be incomplete in the case of returnees. In Keane and Wolpin (1997) it was found that 6.5 percent of a similar sample of youths from the NLSY returned to high school and completed an additional grade (see also Light (1994)).

If the only rationale for high school attendance is to earn a diploma and if the five-year limit on attendance assumed in the model is strictly binding, then any youth who falls more than one grade level behind in terms of credit acquisition would drop out. To the extent that there is some consumption value to attending high school or some labor market payoff to having completed credits without graduating, individuals may remain in school even after having fallen too far behind to graduate within the five-year period. Table 2 shows the relationship between attendance rates and grade level progression. First, for a given grade level completed, the proportion attending school monotonically declines with years since high school entry. However, a significant number of youths with no possibility of graduation within the five-year horizon still choose to attend school (e.g., 39 (14) percent of those who are still attending grade level eight in period three (four) do not drop out), which indicates an additional payoff to attendance beyond that to graduation (or the enforcement of mandatory school leaving ages).[39] Second, holding constant the number of grade levels that a youth has fallen behind (e.g., grade level eight in period two, grade level nine in period three, etc.), the proportion attending school declines over time (moving diagonally from the upper left), which is consistent with the existence of a finite horizon. Third, attendance among youths who have not fallen behind in grade level is almost universal.

IV. Estimation Results

The estimation assumes that there are four types of individuals in the population (M=4) and that the discount rate is .97.[40] The full set of estimates of the 92 parameters of the model are provided in Appendix A. Many of the parameters are not easily interpreted as estimated. Therefore, we integrate the discussion of specific parameter estimates with the substantive issue of what the estimates of the model imply about dropout behavior. Before turning to that issue, we first present evidence about model fit.

Model Fit

Table 3 compares the actual and predicted values of a number of summary measures of school achievement and employment. In the first row, it is seen that 82.9 percent of our sample graduate from high school. The model prediction, based on a simulated sample of 5,000 youths, is 82.0. In most other dimensions as well, the model fits the data quite closely. There are, however, several exceptions. With respect to schooling, the model overstates the fraction of non-graduates who complete 11^{th} grade (17.0 vs. 14.0), overstates the percent who attend school particularly in the fifth year of high school eligibility for those who have not already graduated (22.8 vs. 18.5), overstates the percent of school attendees who fail at least one course in a year (13.6 vs. 10.7) and understates the percent who fail all five courses in a year (1.3 vs. 2.4), overstates the percent who work part-time while attending high school (23.4 vs. 19.6 percent working 10 hours per week and 15.6 vs. 13.0 percent working 20 hours per week) and concomitantly understates the percent working more than 20 hours per week (6.2 vs.8.6 percent working 30,40, or 50 hours per week) and the percent not working (54.8 vs. 58.8 percent), and overstates the mean hourly wage of those employed part-time while in school ($5.61 vs. $4.73) and those working part- and full-time while not attending school ($6.09 vs. $5.50 and $7.10 vs. $6.77 respectively). Except for the overstatement of wages, the differences do not appear especially large.

Tables 4 and 5 present further evidence on model fit in a multivariate context. Table 4 shows the estimates of what might be called approximate decision rules for three of the six choice alternatives. The parameters of these dichotomous-variable logits represent the response of binary choices to the state variables of the behavioral model. The first column (labeled actual) shows the relationships obtained from the data and the second column (labeled predicted) those obtained from the simulated sample based on the estimated behavioral model. The numbers reported show the changes in the odds-ratios (the probability of the choice divided by one minus the probability of the choice) due to a unit change in the state variable; a coefficient over one indicates an increase in the odds-ratio and a coefficient less than one a reduction.

Summarizing selectively the results from the actual data: (1) completing an additional grade in one year's time is associated with an increase in the odds-ratio of attending school and not working of about 3 percent [(1.47-.437-1)*100)] and with an increase in the odds-ratio of attending school and working part-time of 78 percent [(2.33-.547-1)*100], while not having completed a grade level in an additional year is associated with a reduction in the odds-ratio of attending school and not working of 56 percent [(1-.437)*100] and a reduction in the odds- ratio of attending school and working part time of 45 percent [(1-.547)*100][41]; (2) an increase in cumulative gpa by one point is associated with a reduction in the odds-ratio of attending school and not working of 13 percent and an increase in the odds-ratio of attending school and working part-time of 12 percent; (3) an additional 1000 hours of work experience is associated with a reduction in the odds-ratio of attending school and not working of 38 percent and an increase in the odds-ratio of attending school and working part-time of 31 percent; (4) having worked part-(full-) time the previous period is associated with a reduction in the odds-ratio of attending school and not working of 80 (87) percent and an increase in the odds-ratio of attending school and working part-time of 415 (105) percent; (5) having neither attended school nor worked in the previous period is associated with a reduction in the odds-ratio of attending school and not working of 90 percent[42]; (6) among non-graduates after high school attendance is no longer possible (after period five), an additional year of completed schooling is associated with an increase in the odds-ratio of working full-time of 25 percent, an additional 1000 hours of work experience with an increase of 23 percent and having not worked in the previous period with a reduction of 91 percent.

The results using the simulated data are in every case but one (cumulative gpa in the nf equation) qualitatively the same as those based on the actual data. Moreover, in most cases, the magnitudes of the coefficients do not appear substantively very different. The main exception is that the simulated data tend to exhibit considerably less state dependence than do the actual data. For example having attended school and worked part-time in the previous period is, in the actual data, associated with a 5-fold increase in the odds-ratio of choosing the same alternative the next period, but only a 2-fold increase in the simulated data.

More than half of the coefficients estimated from the simulated data are not distinguishable statistically from the coefficients estimated from the actual data. A joint test of the equality of all of the coefficients is rejected for all three alternatives. It should be recognized that because these tests treat the point estimates from the simulated data without imprecision, i.e., they do not take into account the estimation error associated with the structural parameters that underlay the simulated data, the tests are biased towards rejection.[43]

Table 5 presents a similar analysis for course credits, a logit for the dichotomous event of failing at least one course, and gpa (inclusive of course failures). In most cases, for both measures of academic performance, parameter estimates based on the simulated data are quite close to those based on the actual data. Moreover, none of the coefficients are individually statistically different in the two samples although the joint test for the failure outcome variable rejects coefficient equality. For the gpa (measured inclusive of F's) regression, although some of the individual coefficients differ statistically, the joint test rejects at the five-percent but not at the one-percent level.

Taken together, the evidence from tables 3,4 and 5 on the fit of the model is mixed. Simple summary measures of the data are fit quite well by the model, while more complicated multivariate correlational structures are somewhat at variance. However, even in this latter case, inferences about the effects of particular variables generally would not be misleading. Readers will have to form their own judgements about the validity of the conclusions that are presented below.

Dropping out

Who drops out? Table 6 presents completed schooling levels by type. It is clear from this table that dropping out is confined to two types, type 1 and type 3. Type 1's comprise about 25 percent of all dropouts and type 3's the rest. Although all of the youths in these two groups are dropouts, the two types differ considerably in their completed schooling levels; type 3's are spread over the entire range of grade levels, while the vast majority of type 1's do not complete the ninth grade and almost none complete the tenth. Type 3's also have a higher gpa in the courses that they pass than type 1's, 1.75 vs. 1.53.

Interestingly, there is essentially no type that consists of both graduates and dropouts. Type 2's and type 4's all graduate, with the latter comprising 6.6 percent of all graduates. However, the two types differ considerably in their gpa's. Type 4's are essentially straight A students (gpa=3.99) while the type 2's are C+ students on average (gpa=2.39).[44]

That type 4's have the highest gpa is the result of their type-specific values of the $\theta_0^{\prime}$s, i.e, type 4's have the highest levels of whatever "permanent" traits are associated with school performance, whether it's ability, motivation or initial (at the time of high school entry) knowledge.[45] In Table 7, the ranking of types by all of the unobserved heterogeneity parameters is shown.[46] Type 4's rank first (or essentially tied for first) in all of these parameters that are schooling-related, namely the consumption value of attending school, the perceived value of graduation, and schooling ability/motivation.[47] Indeed, the ranking of these heterogeneity components follow exactly the rankings by gpa shown in Table 6.

Although type 4's are clearly "school-types", they rank last in "permanent" traits that accompany high full-time wage offers for non-high school graduates. However, the relationship between school-related heterogeneity parameters and full-time work heterogeneity parameters is not perfectly inverted. Type 1's do not have the highest "endowment" of market skills in jobs performed by non-high school graduates, ranking second to type 3's. The dispersion in type-specific full-time wage offers is large. On average, type 4's receive full-time wage offers (with zero work experience) of \$3.35 per hour, type 2's \$4.72 per hour, type 1's \$5.43, and type 3's \$6.84 per hour, a difference of \$3.50 per hour from lowest to highest.[48] The rank-order of part-time wages is different, but the dispersion is small. On average, type 4's receive part-time wage offers (with zero work experience) of \$6.10 per hour, type 2's \$5.77 per hour, type 1's \$5.01, and type 3's \$6.25 per hour, a difference of only about \$1.00 per hour.[49] Part-time wage offers exceed full-time wage offers with zero work experience for types 2 and 4, the non-dropouts. However, full-time wage offers grow with additional work experience at three times the rate.

Types have been treated as unobserved to us. However, each of the individuals in the sample can be assigned a type probability by applying Bayes' rule to that individual's contribution to the likelihood function (8). Family background information can then be merged with the outcome data used in the estimation and related to type probabilities. Table 8 shows the mean probabilities of the four types by selected family background characteristics. Although most of the relationships between type propensities and family background characteristics are unsurprising, there are a few, perhaps unexpected, correlations. As might be expected, type 1's are significantly over-represented among those youths with the least educated parents, those who were not living with either natural parent at age 14 or only with the youth's natural mother, those with 4 or more siblings, and those living in a family with income less than one-half the median of the sample. However, type 1's are also over-represented among youths whose fathers are college graduates and among those who had no siblings. Again as expected, type 4's are significantly over-represented among youths whose mother's are college educated, whose father's are college graduates, and whose family income is above twice the median. However, they are also over-represented among youths who lived only with their natural mother at age 14 and who had no siblings. Thus, youths who have no siblings are more likely to be at both extremes with respect to school-related permanent heterogeneity traits.

The Causes of Dropping Out

Why drop out? The model provides for a number of possible reasons for a youth's decision to drop out of high school. It is difficult to parcel out quantitative attributions because the model is highly non-linear and because there is no obvious metric to judge the relative sizes of alternative exogenous changes that would induce dropout behavior. For both reasons, we adopt as a metric the baseline characteristics of type 2 youths, the dominant high school graduation type. To provide a quantitative assessment of the significance of different factors contributing to dropout behavior, we impose on the non-graduate types, 1 and 3, the heterogeneity parameters of type 2's and calculate the change in their performance. Table 9 produces the results of these experiments.

As is evident from the table, adopting any single trait of type 2's would have no effect on the high school graduation rate of type 1's. In addition, only one trait when adopted, school ability/motivation, would substantially alter completed schooling levels, from 8.2 years to 10.6 years. The story is a bit different for type 3's. Adopting type 2's full-time wage offer endowment (a reduction in the wage offer by about 40 percent throughout the life cycle) or type 2's perceived value of graduation (an increase of $227,871 or 15 percent) increases the high school graduation rate from .6 to 17 percent. Alternatively, adopting type 2's value of leisure (a reduction from $17.53 per hour to $5.70 per hour) increases type 3's graduation rate to 10.2 percent while adopting type 2's consumption value of school attendance (an increase of $12,322, from -$619 to $11,703 per year in the case that they don't work while attending school) increases the graduation rate to 6.4 percent.

Perhaps, most interesting is the effect of adopting type 2's school ability/motivation (a fall in the probability of failing at least one course from 64 percent to less than one percent, in the case of attending grade 9 and not working). With this change the graduation rate of type 3's increases to 11.8 percent, which indicates that the dropout behavior of type 3's is due not only to their low ability/motivation. If type 3's made the same schooling-work choices as type 2's, their graduation rate would be the same as for type 2's, 100 percent. As the last row of the table demonstrates, if in addition to having the same ability/motivation as type 2's, type 3's also had the same expected value of graduation, the graduation rate of type 3's would in fact replicate that of type 2's even though they still had other different permanent traits. Of course, the reason type 3's have a low expected value of graduation may be because they have low ability and/or motivation.

Work and Performance in High School

As a first step in evaluating the effect of working while attending high school on performance, consider Table 5 again. Although the parameters in Table 5 do not have a one-to-one correspondence to the structural parameters (the θ's), one can view them as approximations to the structural relationships. The estimates based on the actual data imply that working either part- or full-time increases the likelihood of falling behind in credit acquisition and reduces gpa (inclusive of failures). Using the actual data, working part-time increases the odds-ratio of failure by 24 percent, from .12 to .15, while working full-time raises the odds by 180 percent, from .12 to .34. The simulated data imply a negligible reduction in the odds-ratio for working part-time and a much smaller increase, to 15 percent, for working full-time, although the parameters are not statistically distinguishable. Working part- or full-time while attending school also reduces the gpa (including failures) according to the actual data. The magnitudes are not large, a reduction of .036 and .135 points respectively. Estimated effects of working are quite similar based on the simulated data.

The third column of each performance outcome variable controls for type. In both cases, the probability of failing at least one course and gpa, the detrimental effect of part- or full-time work on performance is, in fact, augmented.[50] Thus, those who would perform better because of their ability/motivation "endowment" are also more likely to work while in school.[51]

Although Table 5 provides evidence that working while in school does reduce school performance, those estimates cannot by themselves provide an assessment of the effect of policy experiments that still leave youths with schooling and work options. To do that requires the use of the behavioral model. Table 10 shows the effect of four policies that constrain work-school choices on school attendance rates. Table 11 shows the impact of those policies on school performance measures. It should be noted that the joint hypothesis that hours of work does not affect the grade function is rejected.[52]

The first constraint (C1) does not permit youths to work and attend school simultaneously. The second (C2) is more constraining, not permitting youths to work during any of the first four years after entering high school regardless of their school attendance. The next constraint (C3), still more restrictive, not only does not permit youths to work during the first four years after entry but also forces them to attend school, i.e., either extends school leaving ages or perfectly enforces existing ones. The final constraint (C4) is the same as the last but extends the constraint to the fifth year after entry for those who did not graduate in four years.

Table 10 shows the effects of these constraints on the school attendance decisions of the two non-graduating types. The baseline case is given in the first column for each type. Notice that with both of the first two constraints attendance rates fall relative to the baseline for both types. Faced with these constraints, youths with these traits would prefer either the working or leisure alternative in the case of the first constraint or the leisure alternative in the case of the second, relative to attending school. The third constraint forces attendance to be 100 percent in the first four years. However, attendance drops to only 11 percent in period five for type 1's and 20 percent for type 3's. The fourth constraint imposes attendance for all five years.

Table 11 shows the effects of these experiments on graduation rates, school completion levels, and gpa (exclusive of failures) for each of the four types. The baseline is given in the first row of the table. A quick perusal of the table reveals that all of the constraints have only a trivial effect on gpa's for all types. We, therefore, concentrate on their effects on completed schooling levels, which further restricts attention to types 1 and 3.

Imposing C1, as was seen in Table 10, reduces school attendance of both types 1 and 3, which leads to a slight increase in the dropout rate and a reduction in the average schooling attainment of dropouts by .8 years. Constraint C2 also reduced attendance, but somewhat less than the C1 constraint. However, in this case establishing a strict no-work constraint for four years, slightly increases graduation rates while reducing the average schooling of dropouts by .4 years. Requiring four years of high school attendance without working leads to a modest increase in graduation rates among type 3's (and no increase among type 1's) and an increase in average schooling among dropouts of about one-half year. Finally, requiring five years of attendance, for those who do not graduate in four years, without working provides the maximum increase in graduation rates that are feasible given the traits of the two types. Even here, no type 1's graduate and only 17.7 percent of the type 3's graduate. However, the average schooling levels of dropouts increase by close to one year.

V. Conclusions

In this paper, we addressed the issue of whether working while attending high school impinges on academic performance. We formulated and empirically implemented a stylized model of grade progression through high school in which youths make sequential decisions about school attendance and work. The model was estimated on longitudinal data that included information about school attendance, hours worked, wages, and course grades. The estimates of the model were used to quantify the importance of alternative reasons for dropping out of high school and to assess the effect of policy interventions to increase graduation rates and grades.

The results can be summarized as follows: (1) Youths who drop out of high school have different traits than those who graduate--they have lower school ability and/or motivation, they have lower expectations about the rewards from graduation, they have a comparative advantage at jobs that are done by non-graduates, and they place a higher value on leisure and have a lower consumption value of school attendance. (2) Without altering their ability/motivation, if dropouts were forced to remain in school for five years after entry (only four years if they graduate in that time) without working, then their graduation rate would increase only to 13 percent. Legislation prohibiting work would have very little impact on graduation rates. In addition, such legislation would have only a negligible impact on the gpa of high school graduates. (3) Increasing the ability/motivation of dropouts to coincide with that of the modal type high school graduates would by itself increase the graduation rate to only 9 percent. Even with augmented ability, there is still a strong incentive, given their other traits, to drop out before graduating. If in addition to augmenting ability, their expected valuation of graduation was also made to coincide with the modal type of high school graduate, their graduation rate would be 100 percent. (4) Policies that do not alter traits that youths come to high school with will have very limited success in improving school outcomes.

Footnotes

1. The time of employment during the day is also restricted to between 7 a.m. and 7 p.m.

2. State child labor laws tend to be even more restrictive, setting shorter daily and weekly hours limitations for periods when schools are in session.

3. Given the more restrictive nature of state laws, we take 15 hours per week as the cut-off for compliance.

4. Compliance would be less complete if we took a longer horizon, e.g., the fraction of youths who ever were non-compliant.

5. There is a pronounced age gradient throughout the age range; in addition to the figures provided in the text, approximately 15 percent of 17-year-old white males attending high school and 23 percent of similar18-year-olds worked at least 15 or more hours. It is therefore not clear whether the jump in hours at age 16 is due to the relaxation of the legal constraints on hours.

6. In the sociological literature, this substitution of time between school and work activities has been referred to as the a zero-sum model (Coleman (1961), Marsh (1991)).

7. Study time in school is obtained from the question "not including time you spent in classes about how may hours did you spend studying at school or working on independent studies or class projects."

8. t-values for the test for equality of means were 2.4,1.5, and 2.6 respectively. The average reported time spent studying in school during the week was 3.5 hours, reported study time out of school was 5 hours, and the total therefore was 8.5 hours. There are 606 observations.

9. The t-value for the test of equality is 2.17. There are 422 observations.

10. The differences and t-values are 0.35 (.46) for grade 10, 1.96 (2.05) for grade 11 and 1.92 (1.36) for grade 12. Sample sizes are 719, 986, and 1,067.

11. See Greenberger and Steinberg (1986) for a survey of the literature. See also D'Amico (1984) and Marsh (1991).

12. See Light (1995) for a recent review.

13. See Hood and Maplethorpe (1980) for a somewhat dated summary.

14. They do not write down the exact optimization problem which makes it difficult to judge the consistency of the specification. However, they do not explicitly treat working as a choice in the statistical formulation although they treat it as an endogenous variable in some relationships.

15. Methods of solving and estimating models with a discrete-choice dynamic programming structure are now well known. Examples are Eckstein and Wolpin (1989a), Keane and Wolpin (1994, 1997), Miller (1984), Pakes (1987), Rust (1987), and Wolpin (1984, 1987). Eckstein and Wolpin (1989b) and Rust (1992) provide useful surveys.

16. We ignore the possibility of obtaining alternative certification such as the General Educational Development (GED) credential. The evidence suggests that a GED has a low pecuniary return (Cameron and Heckman, 1993).

17. This structure of educational attainment during high school is only somewhat stylized. Although the actual attainment process varies considerably both within and among states, according to data from the National Center for Education Statistics (NCES) the average number of credits completed in four years of high school by public high school graduates was 21.5 in 1982. In 1990, the modal number of credits required for graduation was 20 (13 states) with eight additional states requiring 21 credits. The model is empirically implemented within the assumed structure by forcing compatibility of the data with the model as described below.

18. This assumption, while not universally true, is only rarely violated (see below).

19. The direct input would be time spent studying and doing homework. Unfortunately, except

for the 1981 survey round, that information is unavailable in the data.

20. The value of in-kind family transfers as well as direct monetary transfers are assumed not to be contingent on behavior. Also, although such transfers obviously affect consumption, the utility specification rules out any income effects on behavior. These are necessary assumptions, in part, because we have no direct data on intra-family transfers.

21. It might also reflect the perceived cost of not attending if that violates truancy laws.

22. Additional parameters, included in order to better fit specific aspects of the data, are described later so as not to complicate the presentation of the basic structure of the model.

23. The alternative-specific utilities are specified below to ensure existence and uniqueness of stationary Emax functions.

24. If the individual chooses not to attend school at time t, then zero credits are earned between t and t+1 and the cumulative grade point average is unchanged.

25. Notice that the infinite horizon value function for non-graduates after period five is the terminal value function for the finite horizon decision problem. This treatment is similar to that found in Gilleskie (1996).

26. Unlike the model in Keane and Wolpin, it is computationally manageable to evaluate the Emax functions at all elements of the state space rather than developing an approximating function.

27. We observe a part-time wage for those that work part time in a particular period and a full-time wage for those that work full time in a particular period.

28. The measurement error in wages creates an additional integral in periods that include work.

29. For each of K draws of the error vector, ϵ_t^n, ϵ_t^f, η_t^p, ϵ_t^{sn}, ϵ_t^{sp}, ϵ_t^{sf}, noting that $\epsilon_t^p = \ln w_t^{jo} - \ln w^j(K_t,H_t) - \eta_t^p$, the kernel of the integral is $\exp[\frac{V_{tk}^{sf} - \max(V_{tk}^j)}{\tau}] / \sum_i \exp[\frac{V_{tk}^i - \max(V_{tk}^j)}{\tau}]$ times the joint normal density of η_t^p and ϵ_t^p. The first term in the kernel is the smoothed simulator of the

probability that $d_t^{sp} = 1$, with τ the smoothing parameter. The integral is then the average of the kernel over the K draws.

30. We did not include a measure of school performance because we are dealing only with wage offers for schooling levels less than 12.

31. To conserve on parameters we do not allow for type-specific gpa parameters, although the effect of gpa might reasonably depend, for example, on whether one is a type that is likely to attend college. Throughout, we have modeled only first-order heterogeneity effects.

32. The age restriction reduced the sample to 712 and the schooling restriction eliminated 10.

33. There were 588 youths with at least some transcript data. In 24 cases we could not sensibly match that data with the other available information. Those with transcript data are more likely to be high school graduates (80 percent as opposed to 65 percent) and are less likely to have attrited (4.6 percent as opposed to 7.9 percent in the 1983 interview).

34. For more information see the NLS Handbook.

35. Hours worked were set to zero for those too young to have reported hours worked back to 1/1/78, i.e., those who were not age 16 at the 1979 interview.

36. We actually cumulated over the first week of each month and multiplied by 13/3 to obtain monthly hours worked.

37. Because the employment history only goes back to January 1, 1978, in computing hours worked for the period October 1, 1977 - September 30, 1978, we assigned average hours worked based on the January 1, 1978 - September 30, 1978 period. For those not attending school for the entire year, we used hours worked only over the period of school attendance, again because of our primary concern about the effect of working while in school on school performance. Given the age restriction of the sample, all sample members were under age 14 at the start of school years that

began prior to 1977. We assigned them zero hours worked for those school years.

38. As noted, our transcript sample undercounts high school dropouts and so the graduation rate is overstated. Even ignoring that, the 48 attriters between periods one and four are probably mostly dropouts; if they all were dropouts, the overall graduation rate would be 77 percent for our sample.

39. It is also possible that they do still have a possibility of graduating in ways that our model does not accommodate, e.g., through summer school attendance.

40. Four types clearly separated the population distinctly. We did not attempt to add more types because of computational considerations. We also did not attempt to estimate the discount rate because it would seem to be confounded with the expected perceived value of graduation, which we do not observe and thus treat as a parameter.

41. The coefficient on grade completed by itself represents the "effect" of skipping a grade. It is an extrapolation at best because there are no such observations in the sample.

42. The coefficient on lagged nn was not estimable because no youths who neither attended school nor worked in the previous period chose to attend school and work part-time in the next period.

43. Standard errors of the coefficients from the simulated data depend in a complicated way on the variance-covariance structure of the structural parameters.

44. In the data, 6.0 percent of the youths who graduate have 4.0 averages (actually 3.75 - 4.0 due to the discretization.

45. The difference in the coefficients that determines the probability of failure (types 1 and 3 vs. either type 2 or 4) are statistically significant (see Table A.1). However, the model does not well identify the cut-off values of the ordered logit that determine the probabilities of the specific passing grades. Recall that in the estimation, we do not make use of specific course grades, but only gpa's.

46. See Table A.1 for the actual values of the heterogeneity parameters upon which Table 7 and the following discussion is based.

47. Although the magnitude of the difference in the expected value of graduation of type 2's vs. Type 4's is very large, the standard error of the difference is several orders of magnitude greater. The difference between type 3's and type 4's is statistically significant at the 10-percent level. The difference in the consumption value of school attendance of type 1's and 3's vs. type 4's is statistically significant at the 5-percent level.

48. The difference in type 1's vs. type 4's full-time wage offer is statistically significant at the 10-percent level and that between type 3's and type 4's at the five-percent level.

49. None of the differences are statistically significant.

50. Estimation on the actual data incorporating fixed-effects yields the same qualitative result.

51. It is interesting that incorporating types changes the sign of the effect of an additional completed grade level and, in the case of gpa, the sign of the years since high school entry variable. The interpretation of this change is that, holding constant ability/motivation, repeating a grade level improves performance, or in other word, "practice makes perfect."

52. A Wald test of the hypothesis that $\theta_3 = \theta_4 = \theta_5 = 0$ yields a chi-square value of 19.5. The critical value with 3 degrees of freedom is 11.3 at the one-percent level.

References

Bellman, Richard. Dynamic Programming. Princeton: Princeton University Press, 1957.

Cameron, Steven and James J. Heckman. "The Nonequivalence of High School Equivalents." Journal of Labor Economics 11 (1993):1-47.

Coleman, James S. "The Transition from School to Work." in Donald J. Treiman and Robert V. Robinson (eds.), Research in Social Stratification and Mobility Vol. 3. Greenwich CT; JAI Press, 1984.

D'Amico, Ronald. "Does Employment During High School Impair Academic Progress?" Sociology of Education 57 (July 1984): 152-64.

Eckstein, Zvi and Kenneth I. Wolpin. "Dynamic Labour Force Participation of Married Women and Endogenous Wage Growth." Review of Economic Studies 56 (July 1989a):375-90.

__________. "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models." Journal of Human Resources 24 (Fall 1989b):562-98.

Ehrenberg, Ronald G. and Daniel R. Sherman. "Employment While in College, Academic Achievement and Postcollege Outcomes." Journal of Human Resources 23 (Winter 1987): 1-23.

Gilleskie, Donna B. "A Dynamic Stochastic Model of Medical Care Use and Work Absence." forthcoming, Econometrica.

Greenberger, Ellen and Laurence Steinberg. When Teenagers Work: The Psychological and Social Costs of Adolescent Employment. New York; Basic Books, 1986.

Hood, A. and C. Maplethorpe. "Bestow, Lend or Employ: What Difference Does it Make?" New Directions for Institutional Research 25: 61-73.

Keane, Michael P. and Kenneth I. Wolpin. "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." Review of Economics and Statistics 76 (November 1994): 648-672.

_________. "Career Decisions of Young Men." Journal of Political Economy, forthcoming.

Light, Audrey. "High School Employment." Report: NLS 95-27, U.S. Department of Labor Bureau of Labor Statistics (June 1995).

Marsh, Herbert W. "Employment During High School: Character Building or a Subversion of Academic Goals?" Sociology of Education 64 (July 1991): 172-1989.

Miller, Robert. "Job Matching and Occupational Choice." Journal of Political Economy 92: 1086-1120.

National Center for Educational Statistics. Overview and Inventory of State Requirements for School Coursework and Attendance (June 1992). U.S. Department of Education

Pakes, Ariel. "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." Econometrica 54 (July 1986): 755-784.

Rust, John. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." Econometrica 55 (September 1987): 999-1033.

_________. "Structural Estimation of Markov Decision Processes." forthcoming in R. Engle and D. McFadden (eds.) Handbook of Econometrics Vol. 4. Amsterdam: North Holland, 1992.

Wolpin, Kenneth I. "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality." Journal of Political Economy 92 (Oct. 1984): 852-874.

_________. "Estimating a Structural Search Model: The Transition from School to Work." Econometrica 55 (July 1987): 801-817.

Table A.1: Model Estimates

	Parameter	Asymptotic Standard Error
Full-time wage function		
α^F_{04}	.694	.305
$\alpha^F_{01} - \alpha^F_{04}$	.482	.296
$\alpha^F_{02} - \alpha^F_{04}$	.344	.290
$\alpha^F_{03} - \alpha^F_{04}$	.714	.300
α^F_{1}	.336 E-05	.111 E-05
Part-time wage function		
α^P_{04}	1.21	.232
$\alpha^P_{01} - \alpha^P_{04}$	-.197	.240
$\alpha^P_{02} - \alpha^P_{04}$	-.560 E-01	.211
$\alpha^P_{03} - \alpha^P_{04}$	-.243 E-01	.233
α^P_{1}	.119 E-05	.144 E-04
Course grade function		
θ^F_{04}	2.40	1.59
$\theta^F_{01} - \theta^F_{04}$	-6.54	1.60
$\theta^F_{02} - \theta^F_{04}$	1.53	1.55
$\theta^F_{03} - \theta^F_{04}$	-4.50	1.54
θ_1	1.08	.040
θ_2	.730 E-01	.067

Table A.1: Model Estimates

	Parameter	Asymptotic Standard Error
θ_3	-.697 E-01	.054
θ_4	-.136	.073
θ_5	-.102	.094
θ_6	2.99	.120
θ_7	.999 E-02	.064
θ_{04}^{D}	-.335	27.3
$\theta_{01}^{D} - \theta_{04}^{D}$	-.180	27.3
$\theta_{02}^{D} - \theta_{04}^{D}$	.938 E-01	27.2
$\theta_{03}^{D} - \theta_{04}^{D}$	-.275	27.3
θ_{04}^{C}	-.644	1.55
$\theta_{01}^{C} - \theta_{04}^{C}$	-.825	2.12
$\theta_{02}^{C} - \theta_{04}^{C}$	-.425 E-01	1.50
$\theta_{03}^{C} - \theta_{04}^{C}$	-.417 E-01	1.52
θ_{04}^{B}	-1.62	1.81
$\theta_{01}^{B} - \theta_{04}^{B}$	.474	1.85
$\theta_{02}^{B} - \theta_{04}^{B}$	-5.34	1.75
$\theta_{03}^{B} - \theta_{04}^{B}$	-.447	1.79
Utility of leisure		
b_{04}^{n}	4.85	4.21
$b_{01}^{n} - b_{04}^{n}$	9.12	3.70
$b_{02}^{n} - b_{04}^{n}$	-.227	.941
$b_{03}^{n} - b_{04}^{n}$	11.9	3.57

Table A.1: Model Estimates

	Parameter	Asymptotic Standard Error
b_1^n	1.00	1.20
Utility of attending school		
b_{04}^{sn}	1.18 E+04	7.97 E+03
$b_{01}^{sn} - b_{04}^{sn}$	1.29 E+04	6.23 E+03
$b_{02}^{sn} - b_{04}^{sn}$	1.91 E+02	1.56 E+05
$b_{03}^{sn} - b_{04}^{sn}$	-1.18 E+04	6.18 E+01
b_1^{sn}	-2.16 E+03	871
b_2^{sn}	697	595
b_3^{sn}	3.48 E+03	1.92 E+03
b_4^{sn}	-1.14 E+03	2.10 E+03
b_0^{sp}	-6.01 E+03	2.05 E+03
b_1^{sp}	1.15 E+03	378
b_2^{sp}	-2.50 E+ 03	1.43 E+03
b_3^{sp}	4.03 E+03	1.26 E+03
b_4^{sp}	1.88 E+03	1.23 E+03
b_0^{sf}	-1.49 E+04	3.69 E+03
b_1^{sf}	2.53 E+03	659
b_2^{sf}	-4.27 E+03	6.50 E+03
b_3^{sf}	8.36 E+03	1.86 E+03
b_4^{sf}	4.14 E+03	1.06 E+03

Table A.1: Model Estimates

	Parameter	Asymptotic Standard Error
PDV of HS graduation		
γ_{04}	1.66 E+06	3.98 E+03
$\gamma_{01} - \gamma_{04}$	-5.74 E+05	3.53 E+09
$\gamma_{02} - \gamma_{04}$	9.40 E+04	2.99 E+06
$\gamma_{03} - \gamma_{04}$	-1.34 E+05	7.95 E+04
γ_1	1.23 E+03	9.66 E+03
γ_2	.903	4.78
Hours offer functions		
t<=5		
θ^h_{013}	2.35	.514
θ^h_{014}	.210	1.21
θ^h_{015}	-1.84	10.2
θ^h_{023}	.883	.583
θ^h_{024}	.123	.471
θ^h_{113}	-.210	.112
θ^h_{114}	-.108 E-01	.962
θ^h_{115}	-.378 E-01	4.39
θ^h_{123}	.570 E-01	.691
θ^h_{124}	.401 E-01	.146
t>=6		
θ^h_{013}	-5.46	95.4
θ^h_{014}	-4.03	6.65

Table A.1: Model Estimates

	Parameter	Asymptotic Standard Error
θ^h_{015}	-.895	1.54
θ^h_{023}	.127 E-01	.796
θ^h_{024}	.559	.499
Job-finding costs		
c^{pp}	-1.14 E+03	637
c^{ff}	-3.48 E+03	927
Type proportions		
π_1	.422 E-01	.022
π_2	.764	.045
π_3	.142	.023
π_4	.518 E-01	--
Variance-covariance matrix[1]		
$\sigma^2_{\epsilon^n}$	1.99	.425
$\sigma^2_{\epsilon^p}$	.600	.045
$\sigma^2_{\epsilon^f}$	.515	.063
$\sigma^2_{\epsilon^{sn}}$	4.49 E+03	1.36 E+04
$\sigma^2_{\epsilon^{sp}}$	4.14 E+03	1.22 E+03
$\sigma^2_{\epsilon^{sf}}$	495 E+03	1.91 E+03
$\rho_{\epsilon^{sn}\epsilon^{sp}}$	-.159	.052
$\rho_{\epsilon^{sn}\epsilon^{sf}}$	-.939	.848

Table A.1: Model Estimates

	Parameter	Asymptotic Standard Error
$\sigma^2_{\eta^p}$	.101	.093
$\sigma^2_{\eta^f}$	.536 E-01	.378
Ln likelihood	-9,166	

1. The standard errors are those of the Cholesky decomposition parameters.

Table 1: Choice Distribution by Period (Pct.)[1]

Period	sn	sp	sf	nn	np	nf	no. Obs.
1	82.6	15.3	2.1	--	--	--	564
2	65.1	29.8	3.4	.5	.5	.5	553
3	44.5	40.0	9.5	1.7	1.9	2.4	535
4	30.0	40.3	18.2	2.9	3.9	4.7	516
5	7.6	6.5	4.4	13.0	27.2	41.3	92
6	----	----	----	13.4	22.0	64.3	82
7	----	----	----	8.7	17.3	74.1	81
8	----	----	----	10.0	21.3	68.8	80
9	----	----	----	12.8	18.0	69.2	78
10	----	----	----	10.5	10.5	79.0	76
11	----	----	----	16.0	13.3	70.7	75
12	----	----	----	14.5	8.1	77.4	62
13	----	----	----	20.8	12.5	66.7	24

1.sn: in school, not working;
sp: in school, working part-time;
sf: in school, working full-time
nn: not in school, not working;
np: not in school, working part-time;
nf: not in school, working full-time.

Table 2: Attendance Rate (Pct.) by Grade Completed and Period

Period	Grade Level Completed 8	9	10	11
1	100	---	---	---
	(564)[1]			
2	84.3	99.8	---	---
	(51)	(502)		
3	39.3	80.0	98.7	
	(28)	(45)	(462)	
4	13.6	33.3	53.7	99.3
	(22)	(27)	(41)	(426)
5	0.0	12.5	15.6	52.9
	(19)	(24)	(32)	(17)

1.Number of observations with completed grade in parentheses.

Table 3: Actual and Predicted Summary Measures of School Achievement and Employment

	Actual	Predicted
Pct. high school graduates:	82.9	82.0
Pct. dist. grade completion levels less than 12:		
8	22.9	24.0
9	26.7	25.0
10	37.2	34.1
11	14.0	17.0
Pct. attending school by years since entry:		
1	100.0	100.0
2	98.4	98.0
3	94.0	95.5
4	88.6	91.1
5	18.5	22.8
Pct. dist. course credits earned per period:		
0	2.4	1.3
1	1.2	2.0
2	2.6	2.8
3	2.5	3.5
4	1.9	4.0
5	89.3	86.4

Table 3 (continued)

	Actual	Predicted
Mean gpa by grade level attending (passing grades):		
9	2.36	2.35
10	2.30	2.26
11	2.38	2.36
12	2.40	2.43
Pct. dist. wkly hours worked while attending school:		
0	58.8	54.8
10	19.6	23.4
20	13.0	15.6
30	6.5	4.4
40	1.6	1.2
50	0.5	0.6
Mean hourly wage ($):		
In school, work part-time	4.45	4.29
In school, work full-time	4.73	5.61
Not in sch., work part-time	5.50	6.09
Not in sch., work full-time	6.77	7.10

Table 4: Actual and Predicted Selected Approximate Decision Rules[1] (Maximum Likelihood Binary Logit)

	sn_t[2]		sp_t[3]		nf_t[4]	
	Actual	Predicted	Actual	Predicted	Actual	Predicted
Grade comp. (e)	1.47 (.191)	1.85*	2.33 (.310)	1.66**	1.25 (.158)	1.13*
Yrs. since entry (t)	.437 (.057)	.435*	.547 (.070)	.634*	.808 (.062)	.987
Cum. gpa (G)	.987 (.052)	.885**	1.12 (.055)	1.18*	.694 (.126)	1.02**
Work exp. (H /1000)	.624 (.104)	.909**	.690 (.084)	.844*	1.23 (.043)	1.10
sp_{t-1}	.202 (.036)	.548	5.15 (.857)	2.23	--	--
sf_{t-1}	.127 (.066)	.562	2.05 (.739)	1.10*	--	--
nn_{t-1}	.099 (.106)	.520*	--	--	.092 (.036)	.419
sample odds-ratio	1.19	1.03	.429	.568	2.51	1.89
chi-square	476	--	293	--	102	--
H_0: pred.=act.	reject		reject		reject	

*,** denotes non-rejection of the equality of the actual and predicted underlaying logit coefficients at the 5%, 1% level.

1. Coefficients are percent changes in the odds-ratio divided by 100 plus one for a unit change in the variable, i.e., $\partial(p/1-p)=\beta\cdot(p/1-p)$. Robust standard errors in parentheses.

2. estimated for t<=5.

3. estimated for t<=5.

4. estimated for t>=6.

Table 5: Actual and Predicted Course Credit and Grade Point Average Functions[1]

	Failed One or More Courses (I(c<5)): logit[2]			Grade Point Average (includes "F"'s): ols		
	Actual	Predicted	Predicted	Actual	Predicted	Predicted
Grade comp. (e)	.008 (.004)	.007*	.804	.793 (.067)	.581	-.036
Yrs. since entry (t)	51.1 (21.5)	45.3*	1.08	-.701 (.067)	-.466	.067
Cum. gpa (G)	.236 (.049)	.175*	.203	.801 (.022)	.839*	.629
I(C>0)	.033 (.017)	.020*	.011	2.09 (.086)	2.28**	1.77
Worked part-time	1.24 (.287)	.974*	1.28	-.036 (.039)	-.064*	-.077
Worked full-time	2.80 (1.32)	1.25*	1.88	-.135 (.066)	-.169**	-.169
Type 2	--	--	.000012	--	--	1.73
Type 3	--	--	.011	--	--	.728
Type 4	--	--	.00016	--	--	2.59
constant	--	--	--	-5.41 (.465)	-4.11	-.749
R^2	.432	.510	.814	.453	.508	.653
H_0: pred. = actual	reject			reject at 5% level not reject at 1% level		

*,** denotes non-rejection of the equality of the actual and predicted underlaying logit coefficients at the 5%, 1% level.

1. Robust standard errors in parentheses. 2,069 person-periods.

2. Coefficients are percent changes in the odds-ratio divided by 100 plus one for a unit change in the variable, i.e., $\partial(p/1-p)=\beta\cdot(p/1-p)$. The sample odds-ratio is .118.

Table 6: Percent Distribution of Grade Completion Levels by Type

Grade Level	Type 1	Type 2	Type 3	Type 4
8	82.5	0	4.1	0
9	16.7	0	27.6	0
10	0.9	0	45.1	0
11	0	0	22.6	0
12	0	100.0	0.6	100.0
Cumulative terminal gpa	1.53	2.39	1.75	3.99[1]
Percent of sample	4.6	76.5	13.5	5.4

1. 3.6% of the sample reported a terminal gpa of 4.0.

Table 7: Rank-Order of Types by "Permanent" Unobserved Traits

	Type 1	Type 2	Type 3	Type 4
Full-time wage	2	3	1	4
Part-time wage	3	4	1	2
Consumption value of leisure	2	3	1	3
Consumption value of school	4	1	3	1
Perceived value of h.s. graduation	4	1	3	1
Schooling ability/ motivation	4	2	3	1

Table 8: The Relationship of Type to Selected Family Background Characteristics[1]

	Type 1	Type 2	Type 3	Type 4
Sample (Pct.); (543)	5.2	77.0	13.4	4.5
Mother's Schooling				
All non-missing (524)	4.8	78.2	12.4	4.6
Non-HS Grad (111)	10.9	62.6	23.9	2.7
HS Grad (284)	3.9	83.0	10.1	2.9
Some College (70)	2.6	77.6	8.9	10.9
College Grad (99)	4.1	75.6	14.1	6.2
Father's Schooling				
All non-missing (514)	4.3	78.6	12.4	4.7
Non-HS Grad (132)	8.9	70.8	16.9	3.4
HS Grad (202)	3.6	77.6	15.5	3.2
Some College (69)	2.9	83.7	10.2	3.2
College Grad (161)	5.0	78.3	9.2	7.4
HH Structure at Age 14				
All non-missing (556)	4.7	77.4	13.4	4.5
Live with Mother Only (70)	8.4	65.5	18.8	7.3
Live with Father Only (24)	4.2	70.8	20.9	4.1
Live with Both parents (451)	4.0	79.9	11.8	4.2
Live with Neither Parent (11)	8.8	65.6	25.5	0.0

Table 8 (continued)

		Type 1	Type 2	Type 3	Type 4
Number of Siblings					
All non-missing	(556)	4.7	77.4	13.4	4.5
0	(23)	8.1	76.9	6.8	8.2
1	(121)	4.0	80.1	11.1	4.8
2	(145)	2.9	82.7	9.7	4.6
3	(128)	2.7	74.5	19.9	3.0
4+	(147)	10.0	71.1	14.2	4.7
Region at age 14					
All non-missing	(550)	4.7	77.2	13.5	4.6
South	(138)	4.9	74.1	17.1	4.0
Non-south	(412)	4.6	78.2	12.3	4.8
Parental Income : 1978					
All non-missing	(465)	4.5	76.4	14.4	4.6
Y<=1/2 Median	(71)	15.0	61.6	19.3	4.2
½Median<Y<=Median	(170)	4.1	73.2	18.4	4.2
Median<=Y<2 Median	(196)	1.8	83.7	10.7	3.8
Y>=2 Median	(28)	0.0	81.7	4.5	13.7

1. Numbers in parentheses are sample sizes

Table 9: Effects of Heterogeneity in Traits on School Completion Levels

	Type 1		Type 3	
	Percent H.S. Graduates	Ave. Schooling Non-Graduates	Percent H.S. Graduates	Ave. Schooling Non-Graduates
Baseline	0.0	8.2	0.6	9.9
If type 1,3 had type 2 trait:				
Full-time wage offer	0.0	8.1	17.1	9.8
Part-time wage offer	0.0	8.2	0.7	9.9
Consumption value of leisure	0.0	8.1	10.2	9.7
Consumption value of school	0.0	8.5	6.4	10.6
Perceived value of graduation	0.0	8.2	17.0	10.0
School ability or motivation	0.0	10.6	11.8	10.7
Val. graduation + abil/motiv.	100.0	--	100.0	--

Table 10: The Effects of Employment and School Leaving Constraints on School Attendance Rates (Pct.) by Period [1]

	Type 1					Type 3				
Period	Base	C1	C2	C3	C4	Base	C1	C2	C3	C4
1	100	100	100	100	100	100	100	100	100	100
2	80.7	54.8	59.2	100	100	91.4	61.0	74.6	100	100
3	55.3	30.3	31.1	100	100	82.0	40.6	55.1	100	100
4	31.1	11.4	16.7	100	100	57.6	19.4	36.9	100	100
5	10.5	3.5	6.6	11.0	100	18.2	3.1	13.9	19.9	100

1. Base is the baseline case.
. C1 imposes the constraint that working while attending high school is not permitted.
C2 imposes the constraint that working during the first four years of high school eligibility is not permitted, regardless of school attendance.
C3 imposes the constraints that (1) working during the first four years of high school eligibility is not permitted and (2) school attendance during those years is mandatory.
C4 imposes the constraints that (1) working is not permitted until either high school graduation or the five years of high school eligibility is exhausted and (2) school attendance during those years is mandatory.

Table 11: The Effects of Employment and School Leaving Constraints on School Completion Levels and Grades by Type[1]

	Type 1			Type 2			Type 3			Type 4		
	Pct. Grad	Ave. Sch. <12	Gpa	Pct. Grad	Ave. Sch. <12	Gpa	Pct. Grad	Ave. Sch. <12	Gpa	Pct. Grad	Ave. Sch. <12	Gpa
Base	0.0	8.2	1.53	100	--	2.39	0.6	9.9	1.75	100	--	3.99
C1	0.0	8.1	1.52	100	--	2.44	0.4	9.1	1.83	100	--	3.99
C2	0.0	8.1	1.52	100	--	2.44	1.3	9.5	1.75	100	--	3.99
C3	0.0	8.4	1.53	100	--	2.44	3.1	10.3	1.75	100	--	3.99
C4	0.0	8.6	1.55	100	--	2.44	17.7	10.7	1.75	100	--	3.99

1. Base is the baseline case.

. C1 imposes the constraint that working while attending high school is not permitted.

C2 imposes the constraint that working during the first four years of high school eligibility is not permitted, regardless of school attendance.

C3 imposes the constraints that (1) working during the first four years of high school eligibility is not permitted and (2) school attendance during those years is mandatory.

C4 imposes the constraints that (1) working is not permitted until either high school graduation or the five years of high school eligibility is exhausted and (2) school attendance during those years is mandatory.

This is a list of list of the Institute for Economics Research working papers. A copy of any of these papers can be obtained by writing to:

Ms. Diana Smith
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia PA 19104-6297

97-001 "On Money as a Medium of Exchange When Goods Vary by Supply and Demand" by Xavier Cuadras-Morató and Randall Wright
97-002 "Policy Persistence" by Stephen Coate and Stephen Morris
97-003 "Class Systems and the Enforcement of Social Norms" by Harold L. Cole, George J. Mailath and Andrew Postlewaite
97-004 "Indeterminacy, Home Production, and the Business Cycle: a Calibrated Analysis" by Roberto Perli
97-005 "Eliminating Race Differences in School Attainment and Labor Market Success" by Michael P. Keane and Kenneth Wolpin
97-006 "Correlated Equilibria and Local Interactions" by George J. Mailath, Larry Samuelson and Avner Shaked
97-007 "A Model of Commodity Money, With Application to Gresham's Law and the Debasement Puzzle" by François R. Velde, Warren E. Weber and Randall Wright
97-008 "Approximate Common Knowledge and Co-ordination: Recent Lessons from Game Theory" by Stephen Morris and Hyun Song Shin
97-009 "A Dynamic Equilibrium Model of Search, Bargaining, and Money" by Melvyn Cole and Randall Wright
97-010 "Why Are Saving Rates in East Asia So High? Reviving the Life Cycle Hypothesis" by F. Gerald Adams and Peter A. Prazmowski
97-011 "How Proper is Sequential Equilibrium?" by George J. Mailath, Larry Samuelson and Jeroen M. Swinkels
97-012 "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks" by Stephen Morris and Hyun Song Shin
97-013 "Policy Conditionality" by Stephen Coate and Stephen Morris
97-014 "Policy Analysis in Search-Based Models of Money" by Yiting Li and Randall Wright
97-015 "Toward a Theory of International Currency: A Step Further" by Alberto Trejos and Randall Wright
97-016 "Homework in Labor Economics: Household Production and Intertemporal Substitution" by Peter Rupert, Richard Rogerson and Randall Wright
97-017 "Cointegration and Long-Horizon Forecasting" by Peter F. Christoffersen and Francis X. Diebold
97-018 "Evaluating Density Forecasts" by Francis X. Diebold, Todd A. Gunther and Anthony S. Tay
97-019 "Measuring Predictability: Theory and Macroeconomic Applications" by Francis X. Diebold and Lutz Kilian
97-020 "Optimal Prediction Under Asymmetric Loss" by Peter F. Christoffersen and Francis X. Diebold
97-021 "Child Nutrition, Child Health, and School Enrollment: A Longitudinal Analysis" by Harold Alderman, Jere R. Behrman, Victor Lavy and Rekha Menon
97-022 "How Strongly is Child Schooling Associated with Household Income?" by Jere R. Behrman and James C. Knowles
97-023 "Child Health and Schooling Achievement: Association, Causality and Household Allocations" by Jere R. Behrman and Victor Lavy
97-024 "Women's Schooling, Home Teaching, and Economic Growth" by Jere R. Behrman, Andrew Foster, Mark Rosenweig and Prem Vashishtha
97-025 "Mother's Schooling and Child Education: A Survey" by Jere R. Behrman
97-026 "Choosing Between Levels and Logs in the Presences of Deterministic and Stochastic Trends" by Valentina Corradi and Norman R. Swanson
97-027 "Deciding Between $I(0)$ and $I(1)$ via FLIL-based Bounds" by Valentina Corradi
97-028 "A LIL for m-Estimators and Applications to Hypothesis Testing with Nuisance Parameters" by Filippo Altissimo and Valentina Corradi
97-029 "A Dynamic Equilibrium Economies: A Framework for Comparing Models and Data" by Francis X. Diebold, Lee E. Ohanian and Jeremy Berkowitz
97-030 "Converting 1-Day Volatility to h-Day Volatility: Scaling by $\sqrt{h}$ is Worse than You Think" by Francis X. Diebold, Andrew Hickman, Atsushi Inoue and Til Schuermann
97-031 "Macroeconomic Forecasting is Alive and Well" by Francis X. Diebold
97-032 "Youth Employment and Academic Performance is High School" by Zvi Eckstein and Kenneth I. Wolpin
97-033 "Degenerate Continuous Time Limits of GARCH and GARCH-type Processes" by Valentina Corradi
97-034 "Continuous Approximations of Stochastic Evolutionary Game Dynamics" by Valentina Corradi and Rajiv Sarin
97-035 "Nonlinear and Non-Gaussian State-Space Modeling with Monte-Carlo Simulations" by Hisashi Tanizaki and Roberto S. Mariano
97-036 "Stock Market Returns and Economic Fundamentals in an Emerging Market: The Case of Korea" by Fangxiong Gong and Roberto S. Mariano