Volume Title: Economic Aspects of Health

Volume Author/Editor: Victor R. Fuchs, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-26785-7

Volume URL: http://www.nber.org/books/fuch82-1

Publication Date: 1982

Chapter Title: Physical Disabilities and Post-secondary Educational Choices

Chapter Author: Robert A. Shakotko, Michael Grossman

Chapter URL: http://www.nber.org/chapters/c6549

Chapter pages in book: (p. 185 - 202)

6 Physical Disabilities and Post-secondary Educational Choices

Robert A. Shakotko and Michael Grossman

There is a well-documented positive correlation between good health, measured in a number of different ways, and high levels of formal education (see, for example, Grossman 1976). Furthermore, it is generally agreed that three potential structural relations could generate this this positive correlation. In the first case, poor early life cycle health may hamper an individual's education, leading to the subsequent observation that individuals in poor health tend to have lower levels of education. A second relationship may be that schooling affects subsequent health outcomes. For example, individuals with higher levels of schooling may be able to work at less hazardous jobs, or may be able to make health investments more efficiently. This could happen in addition to any income effects which might be indirectly due to schooling (see Grossman 1972). Finally, the correlation could be generated not by any structural relationship directly linking schooling and health, but by common underlying variables (observed and unobserved) determining each.

This paper is an empirical investigation of the first relationship mentioned above. We use panel data for a sample of 10,430 individuals who were high school seniors in the spring of 1972, and who were resurveyed in October of each year through 1976. Various health information was collected in the base year of the survey, and we use these base year reports as measures of health which are predetermined with respect to educational behavior in the subsequent five years. We examine individuals' choices of post-secondary activities (which include three different types of post-secondary education and no post-secondary education), and the rate at which individuals leave educational activities, in an effort to determine if the behavior of disabled individuals differs from

Robert A. Shakotko is at Columbia University and the National Bureau of Economic Research; Michael Grossman is at the Graduate School of the City University of New York and the National Bureau of Economic Research.

healthy individuals, and if these differences could be attributable to health problems. We find no firm evidence that the disabled differ significantly in either their choices or their rate of attrition; however, there is weak evidence that the disabled choose certain types of post-secondary education more frequently, and stay in such programs longer. An important caveat should be appended: we find that the disabled score significantly lower on standardized tests, which are also good predictors of educational choices and outcomes. It may be that the effects of a physical disability are already embodied in an individual's skills and abilities by the post-secondary stage, and that subsequent effects may actually be minimal, or may not be observed because of a prior selection process. Finally, we find weak evidence for higher rates of return to education for the disabled.

As in many previous studies, the issue of defining and measuring disability is troublesome. One constraint is the data, which contain no detailed descriptions of health, but do contain several qualitative ratings. For this study, we concentrate on the high school's evaluation of whether or not a student is disabled: individuals with mental or emotional problems were excluded from the sample, and slightly over 1% of the remaining sample were classified as disabled students. The criteria whereby such classifications were made are not known. Nevertheless, there are two advantages to using such an indicator. First, school-reported disabilities are likely to reflect health problems of a more permanent and identifiable nature than, say, self-rated health status. Second, school-reported disabilities would likely be used by the school to apply for federal or state aid for disabled students; in other words, this indicator is likely to have been used to identify a target population for a particular policy, and one might argue that a similar type of indicator would be used to identify problems and target populations for future policy. The disadvantages are that such an indicator may reflect "true health" no better than would self-rated health (or any other measure), and that even school-reported disabilities may be subject to some selection bias on the basis of underlying socioeconomic variables.

Education as Sequential Choice

In human capital models with perfect foresight, or those in which certainty equivalence can be invoked, a direct solution for the optimal amount of a certain type of education can be computed by equating the present costs incurred with appropriately discounted future gains. Costs are typically divided into direct and indirect costs, with the former representing explicit educational expenditures and the latter accounting for the foregone income (or at least the valuation of time) during the educational period. The effects of a physical disability on these costs are not unambiguous. Direct costs will likely be higher per unit of education, relative to those for healthy individuals, and in the absence of special assistance, must account for requirements for special equipment or extraordinary provisions. Indirect costs may be higher or lower: lower opportunity costs for the disabled (because of restricted market opportunities) would encourage higher levels of education, but this may be partially or wholly offset by the greater calendar time required to complete a given level of education. Similarly, the benefits of higher levels of education for disabled persons, compared to those persons without disabilities, are also ambiguous. One must remember that benefits are individual-specific, and are measured by the individual-specific benchmark of earnings had a particular educational choice not been made. If higher education is a substitute for other, more physical aspects of human capital, such as good health, then the rate of return to higher education could be higher for the disabled. Conversely, if higher education is complementary to good health, the rate of return would be lower. In short, while conventional theoretical models are convenient vehicles for illustrating the dimensions of the problem, they provide little guidance as to what one might expect to find in data, so that the question of the effects of a physical disability on educational choices is largely empirical.

A useful extension to these models for analyzing educational decisions is to postulate that post-secondary education is regarded by individuals as a problem in sequential choice, whereby individuals may choose to participate in one of several types of education, or not to participate at all. Such choices may be reevaluated periodically, with the result that education may be commenced after a period of absence, terminated, or continued in some different program. Relative to the initial period, these decisions may be either anticipated or unanticipated, since it is reasonable to assume that decisions will be made conditional on new information (e.g. successes or failures in different alternatives). In terms of utility maximization, a particular alternative A_i will be chosen in a given period if the expected stream of lifetime utility, appropriately discounted, given that A_i is chosen, is greater than the expected stream of utility given that any other alternative is chosen.[1]

Consider the following two-period model with alternatives A_i, $i = 1,2, \ldots, m$. Let utility in period j under state A_i be denoted

$$V_{ij} = V_{ij}(A_{ij},X_j,e_{ij}) \quad , j = 1,2 \quad ,$$

where X_j is a vector of predetermined variables (which may include previously chosen alternatives and their results, and where e_{ij} is a random component specific to individuals and unrelated to X_j. This representation for utility may be viewed as an indirect utility function embodying the costs and benefits associated with different alternatives. Then, expected lifetime utility, evaluated in the first period and given that A_i is chosen in the first period, is given by

$$(1) \qquad EU(A_{i1}) = V_{i1} + rE\max_k \, (V_{k2}/A_{i1})$$

where r is a discount factor, and where the second term on the right side is the expected maximal second period utility, given that A_i was chosen in the first period, that a number of alternatives are available in the second period, and that second period utility is random (because of the e_{k2}). It follows that A_{i1} will be chosen if and only if $EU(A_{i1})$ is maximal with respect to the set of first period alternatives. Formulated in this fashion, this choice problem is an example of a discrete time-discrete state dynamic programming problem, which can easily be extended to more than two periods.

Since utility is random (because of the individual-state factors e_{ij}), the choice of A_i for any particular individual is a random event, and the probability that A_i will be chosen is given by

$$(2) \qquad P(A_i) = P[EU(A_i) > EU(A_j), j \neq i] \ ,$$

where the time index is discarded for notational simplicity. The difficulty, however, is that (1) is difficult to parameterize except for a few special cases. In the first instance, the distribution of the random variable $EU(.)$ is difficult to derive. Even under the assumptions that the $V..(.)$ are linear and the e_{ij} are normal, this normality will not be preserved in the random variable max (V_{k2}/A_{i1}). In the second instance, even if the marginal distributions for each of the $EU(.)$ can be derived, the probability given by (2) is the probability of a maximal event, and its evaluation requires the derivation of the joint distribution for the set of random variables $EU(A_k), k = 1,2, \ldots, m$. These random variables will not in general be independent, even if the e_{ij} are independently distributed across states, because valuations of future period utility under different current states are likely to be correlated.

This illustrative model is intended not so much to show the difficulty of parameterizing the choice problem as to indicate a possible relationship between the choice probabilities and the subsequent hazard rates, due to a self-selection bias. The hazard rate (i.e., the probability of leaving an activity) is determined by the individual's valuation of second period utility in this simple model, and the hazard function would express the relationship of V to the event of leaving the initial state. Leaving, of course, may be anticipated or unanticipated, depending on both observed variables and random effects. The problem is that only a subsample is observed to ever have entered a particular initial state, and that this subsample is self-selected on a basis that may prejudice estimation of the hazard function.

Self-selection is only a problem if the random component of utility is not independent between periods; otherwise, we might think of each period's choices as constituting independent events, after conditioning on the observed variables determining these choices. However, if utility is serially correlated in its random component, the independence of intertemporal choices vanishes. For example, an individual may value a particular alternative highly because of certain predetermined variables X, or because of a particular configuration of the e_{i1}, $i = 1,2, \ldots ,m$. It follows that the expectation of these random variables will not in general be zero, given that A_{j1}, for instance, is chosen: the expectation of e_{j1} is likely to be positive, while the expectations of $e_{k1}, k \neq j$, are likely to be negative. For the subset of individuals who have chosen A_j, if the $e_{i\cdot}$ are serially correlated, then $E(e_{i2}/A_{j1}) \neq 0$. Moreover, it is not hard to see that this conditional expectation may be related to the set of predetermined variables X, so that failure to take account of this nonzero expectation may bias estimation of the hazard function.[2]

Given the difficulty even in this simple two-period case of parameterizing the distributions associated with the events A_j, $j = 1,2, \ldots ,m$, it would seem to be nearly impossible to derive an exact representation for the above conditional expectation. However, a linear (wide sense) conditional expectation in the sense of Doob (1953) is not difficult to write, and has the advantage of being empirically tractable in cases when exact representations are not known.

Let z be a set of sufficient statistics for the choice of first period alternatives $A_{j1}, j = 1,2, \ldots ,m$; then, for the definition of Bayesian sufficiency,

$$(3) \qquad E(e_{j2}/z) = E(e_{j2}/A_{k1}) \text{ for all } j,k.$$

Assuming that the underlying distribution generating the events A_{k1} is regular in the sense of Dynkin,[3] then by Dynkin's lemma the log likelihood function is a sufficient statistic for the event A_{k1}, for each observation.[4] Letting L denote the log likelihood function for a particular individual, the linear conditional expectation can be written

$$(3') \qquad \tilde{E}(e_{j2}/A_{k1}) = a + bL \ ,$$

where a and b are in this case unknown parameters. While (3′) may not be a completely accurate representation of the "true" conditional expectation (3), it is computationally convenient, especially in cases of nondichotomous or nonnormal selection rules. Furthermore, since L is a sufficient statistic for the prior selecting event, a Taylor series expansion of (3) is equivalent to a polynomial function in L, and (3′) can be viewed as the special case where only the first two terms are included.

Data

The data used for this study is a subset of the National Longitudinal Study of the High School Class of 1972. In this panel survey, approximately 21,000 high school seniors were surveyed in the spring in 1972 just prior to graduation, and resurveyed in October of each year through 1976. The criteria for inclusion in the sample analyzed here are that the individual be a nonminority student and that relevant information from all panel waves exist. In addition, students with reported mental or emotional handicaps are excluded. The final sample size is 10,430.

Several health questions were asked in the base year survey, although a professional evaluation of the student's health was not part of the survey. Limited health information (as it related to post-secondary activities) was also collected in the follow-up surveys, but since this study aims at estimating the effect of poor health on post-secondary activities, we ensure that our measures of health are predetermined with respect to these activities by using base year measures only. For the reasons mentioned earlier, primary emphasis is on a school-reported disability indicator, which also includes limited information on the nature of the disability. Table 6.1 presents descriptive statistics for selected variables for the sample as a whole, and for those subsets of the sample which are school-reported disabled and self-reported disabled.

Self-reported disabled is defined as the student's positive response to the question of whether poor health interferes with his/her education. Not surprisingly, a significantly higher proportion of students report a disability than are officially disabled according to school records. This may be indicative of transitory health problems, or perhaps a self-reporting bias induced by other school-related difficulties and not actual health problems. What is more surprising is that the two categories are only mildly correlated: only 34% of those who are school-reported disabled classify themselves as disabled. This may be a reflection on the validity and accuracy of self-reported health data.

Aside from their post-secondary activities, the major differences between the full sample and the disabled subsamples are found in test scores, both in SAT scores for those who took the test and in the reading and mathematics tests administered by the survey. There are no large differences in either parental education or family income between the disabled and healthy. This is a somewhat surprising finding which runs counter to previous evidence.[5] Other major differences between the disabled and healthy are evident in post-secondary activities. Rather than index each activity by time as well as type, we choose to define activity streams according to first experience with post-secondary education, and then record the duration of particular streams. Specifically, we define four post-secondary alternatives: (1) university education, (2) junior college education; (3) vocational or technical education; and (4) no post-secondary education. Since timing aspects of education are ignored, an individual is defined to have chosen a particular educational stream (i.e., one of the first three above) if the first experience with education is

Table 6.1 **Means and Standard Deviations (in parentheses) of Variables**

Variable	Full Sample $n = 10430$	School-Reported Disabled $n = 120$	Self-Reported Disabled $n = 461$
Female dummy	.508	.475	.345
SAT-Verbal score (for those with valid SAT scores)	466.345 (333.518)	446.464 (365.075)	451.758 (367.081)
SAT-Quantitative score (for those with valid SAT scores)	499.656 (356.687)	476.107 (389.405)	480.590 (390.103)
No valid SAT dummy	.679	.767	.777
Reading score	52.322 (9.259)	47.908 (10.777)	49.751 (9.792)
Math score	52.397 (9.324)	48.267 (10.489)	49.855 (9.650)
Parent's income (1972; 100's)	132.94 (56.64)	132.86 (61.75)	132.94 (56.58)
Rural dummy (1972)	.066	.075	.100
Large city dummy (1972)	.759	.742	.720
Father—some p-s education dummy	.437	.425	.447
Mother—some p-s education dummy	.358	.383	.377
School-reported disability dummy	.012	1.0	.089
Self-reported disability dummy	.044	.342	1.0
First p-s education at 4-year college or university dummy	.366	.258	.260
First p-s education at junior college dummy	.175	.217	.176
First p-s education at vocational/technical school dummy	.145	.167	.171
No p-s education dummy	.314	.358	.393
Years out before starting p-s education (for those with some p-s education)	.343 (0.774)	.221 (0.468)	.421 (0.953)

of that type. The fourth alternative is a residual category, and also presumed to be an absorbing state.

It is apparent from Table 6.1 that the educational choices of the disabled are different from those of the full sample. In particular, university choices are made less often, and other educational and noneducational choices more often. Whether this is due to disability, or to other variables such as the lower test scores, is the central question that we consider in the first part of our empirical analysis.

Since the focus of our study is on school-reported disabilities, and since the number of individuals in this category is relatively small, the power of any tests attempting to distinguish differential behavior of the disabled and the healthy will be low. Consequently, inferences must be made at lower levels of statistical significance, or else empirical results can be taken to be suggestive only. For example, with regard to low testing power, despite the indication from Table 6.1 that post-secondary educational choices of the school-reported disabled are different, the null hypothesis that the disabled and the healthy have the same choice probabilities can be rejected only at the 70% level, based on a χ^2 test with three degrees of freedom.

The second part of the empirical analysis focuses on the time spent in each defined educational state. Time is accounted by counting one year for each year of full time study, and one-half year for each year of part time study. As mentioned above, educational states are defined as streams, so that cases where subsequent education is of a different type than first post-secondary education are in a sense misclassified. Investigation of these data has shown that instances of such education switching occur relatively infrequently, except for individuals who move from junior colleges to universities. For example, between the first and second year after high school, about 6% of first year enrollments switch type in the second year, and one-third of these switch from junior college to university. The proportion between the second and third years is 10% (of second year enrollments), and two-thirds of these make the junior college to university switch. In subsequent years, about 2% switch type. While the analysis of duration assumes constancy of educational type, this imperfect assumption seems unlikely to seriously affect any main results. It is nevertheless important to remember especially that the junior college alternative in this analysis is for some individuals a combined junior college and university education.

The third part of the empirical analysis examines 1976 earnings for those engaged in full time work in October of that year. Clearly, this presents a potentially serious selectivity bias, since it includes only 60% of the original sample, the remainder being either in school, in part time employment, or unemployed. This selectivity could be particularly serious with regard to the disabled, who are likely to be underrepresented in this subsample, so that findings regarding the effects of a disability should be viewed with this possibility in mind.

The measure of work experience was computed in the same fashion as the duration of education measure, with full time work counting for one year, and part time work for one-half year. Other variables, such as marriage and marriage plans, are strongly related to post-secondary activities, but have been excluded from this analysis since it is likely that they are part of a more complicated structural model.

Estimates

Choice Functions

The model outlined earlier does not point to any specific statistical methodology for the analysis of the educational choice problem, either in terms of parametric families of distributions or required properties of choice rules. The random utility formulation of the problem, however, has been used in conjunction with a polytomous logit representation of choice probabilities, and we follow this general methodology here. Two different logistic models are investigated, corresponding to different assumptions regarding independence of irrelevant alternatives; we find the implications of each to be substantially the same, with the model whose null hypothesis is extreme independence being marginally superior in terms of a comparison of the likelihood functions at their respective maxima.

We assume that, given m alternatives, the probability of choosing the ith alternative, conditional on a variable (or vector of variables) X, is given by

$$P(A_i) = \exp(X\beta_i) / \sum_{k=1}^{m} \exp(X\beta_k) \tag{4}$$

where the β_k are conformable to X, and where the β_k satisfy the identifying restriction

$$\sum_{k=1}^{m} \beta_k = 0.$$

Incorporating this restriction into (4), the parameter space for the unconstrained problem can be denoted by $\beta = (\beta_1, \beta_2, \cdot \beta_{m-1})$. Then, for the jth observation, define $c(k,j) = 1$ if alternative k is chosen, and $c(k,j) = 0$. otherwise. The log likelihood function for the jth observation then can be written

$$L_j(X, \beta) = \sum_{k=1}^{m} c(k,j) \log P_j(A_k) \ , \tag{5}$$

where the probabilities P are given by (4). If follows immediately that given a random sample, the sample log likelihood function is

$$L(\beta) = \sum_{j=1}^{N} \sum_{k=1}^{m} c(k,j) \log P_j(A_k) \ ,$$

where N is the sample size.[6]

We can use the general method of Jennrich (1969) to transform this maximum likelihood problem into one of nonlinear least squares. Since the probabilities given by (4) are by construction in the open unit interval, and since (5) contains one and only one nonzero term, L_j is negative definite. Furthermore, from (4), it is easy to verify that L_j is a monotonic function of each of the elements of β. Define

$$f_j(X,\beta) = (-L_j)^{1/2}$$

and let $y = 0$. Since L_j is monotonic in β, so is f_j. Now define

$$u_j = y - f_j(X,\beta) \ ;$$

it is immediately apparent that minimizing $\sum_{j=1}^{N} u_j^2$ is equivalent to maximizing the sample likelihood function.

Table 6.2 shows the results of this method when applied to two different logistic formulations of the choice problem. In the first, reported as columns (a), individuals are presumed to choose among the four defined alternatives when all are in the choice set simultaneously (the extreme model). Most of the effects of socioeconomic variables appear as expected, particularly in the equations defining the probability of university enrollment and no post-secondary education. Higher quantitative ability measures, family income, parents with some post-secondary education, and urban residence are all significant positive contributors to choosing university enrollment, and to a lesser extent, junior college enrollment.

The coefficients indicate that disability has no effect on educational choices which is significant at conventional levels, once other variables are controlled for. At much lower levels of significance, there is some evidence that a disability encourages junior college enrollment and discourages the alternative of no post-secondary education. There may be two interpretations of this finding. One is purely statistical, in that the standard errors associated with the disability coefficients are inflated because disability is a relatively rare event in the sample, and the information passed in the "disabled" observations is swamped by that from the "healthy" observations, so that the power of statistical tests on these coefficients is low.

The second and more straightforward interpretation is that the post-secondary choices of the physically disabled are not systematically different from those of healthy individuals. An intensive examination of the disabled subsample alone seemed to indicate that this lack of systematic effect may stem not from the absence of differences in the behavior of the disabled and healthy, but from the increased difficulty in predicting the choices of the disabled: the logit model applied to the disabled subsample alone was not significantly different than a pure random choice model with no predisposing variables. This lack of systematic behavior was also present even when different types of disabilities, including vision problems, speech and hearing problems, and crippling disabilities, were added to the set of predicting variables.

Columns (b) in Table 6.2 report the parameter estimates of a model which partially relaxes the extreme assumption of the original logit model. The modified model presumes a two-stage decision procedure on the part of individuals, with the probabilities at each stage described by the logistic distribution. The first stage decision is whether or not to engage in some kind of post-secondary schooling, with the second stage determining what type of schooling conditional on the first stage decision. Accordingly, the estimates reported for the "no post-secondary education" alternative are dichotomous logit coefficients, and those reported for the three schooling alternatives are polytomous logit coefficients, again assuming independence among these three alternatives.

The same qualitative picture emerges for all variables, and again there is weak evidence that some schooling alternative is preferred by the disabled, particularly junior college. The two models, of course, do not constitute a nested hypothesis, so that one cannot be tested as a restriction on the other. In any case, their predictive power is nearly the same (with the same number of degrees of freedom), with the log likelihood of the extreme model being -10893.2 and that of the modified model being -10910.5 at their respective maxima.

Hazard Functions

Most economic work has treated education (post-secondary or otherwise) as a homogeneous good, and has focused instead on duration of educational activities. In this section, we estimate the effect of a physical disability on duration for each of the three defined educational alternatives. Since the data contain censored observations (approximately 35% of the individuals ever enrolled in post-secondary education were studying in October, 1976—those not studying in 1976 were presumed to have completed their education), methodologies for analyzing completed spells, such as those suggested by Heckman and Borjas (1980) are inappropriate, and we are forced to draw inferences regarding duration from an analysis of the hazard rate and its dependence on disability and other variables.

The hazard probabilities suffer from the same difficulties of parameterization as the initial choice probabilities, for the reason that they are both

Table 6.2 **Post–High School Choices***

	University		Junior College		Vocational-Technical		No P-S Schooling	
Variable	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
Female dummy	−.057	−.050	−.143	−.122	.140	.172	.060	.036
	(.086)	(.068)	(.102)	(.076)	(.112)	(—)	(—)	(.026)
	[−.018]		[−.024]		[.021]		[.021]	
SAT-Verbal	.0014	.0010	.0010	−.0013	.0007	.0003	−.0011	−.0009
	(.001)	(.001)	(.001)	(.001)	(.002)	(—)	(—)	(.0005)
	[.00048]		[−.00019]		[.00009]		[−.00037]	
SAT-Quantitative	.0021	.0020	−.0016	−.0014	−.0001	−.0006	−.0004	−.0008
	(.001)	(.001)	(.001)	(.001)	(.002)	(—)	(—)	(.0005)
	[.00064]		[−.00034]		[−.00007]		[−.00024]	
No SAT dummy	.660	.762	−1.309	−1.069	.304	.307	.345	−.036
	(.436)	(.362)	(.587)	(.437)	(.703)	(—)	(—)	(.183)
	[.181]		[−.258]		[.020]		[.057]	
Reading score	.023	.017	.006	−.001	−.009	−.016	−.020	−.012
	(.006)	(.005)	(.007)	(.005)	(.008)	(—)	(—)	(.002)
	[.008]		[.001]		[−.002]		[−.007]	
Math score	.043	.030	.009	−.003	−.017	−.027	−.035	−.022
	(.007)	(.005)	(.007)	(.005)	(.008)	(—)	(—)	(.002)
	[.014]		[.001]		[−.003]		[−.012]	

Parent's income (100's)	.0020	.0013	.0014	.0007	−.0014	−.0020	−.0020	−.0012
	(.001)	(.001)	(.001)	(.001)	(.001)	(—)	(—)	(.0002)
	[.00068]		[.00022]		[−.00022]		[−.00068]	
School-reported disability dummy	−.081	−.106	.229	.185	−.024	−.079	−.124	−.086
	(.426)	(.334)	(.417)	(.309)	(.484)	(—)	(—)	(.113)
	[−.018]		[.046]		[−.001]		[−.029]	
Rural dummy (1972)	.025	.011	−.202	−.167	.107	.156	.070	.024
	(.413)	(.285)	(.402)	(.262)	(.299)	(—)	(—)	(.052)
	[.005]		[−.037]		[.014]		[.018]	
Large city dummy (1972)	.077	.470	.451	.041	−.167	−.511	−1.261	−.805
	(.211)	(.146)	(.196)	(.131)	(.168)	(—)	(—)	(.031)
	[.352]		[.076]		[−.027]		[−.401]	
Father—some p-s education dummy	.280	.174	.185	.091	−.187	−.265	−.278	−.186
	(.094)	(.075)	(.111)	(.083)	(.127)	(—)	(—)	(.030)
	[.095]		[.029]		[−.030]		[−.094]	
Mother—some p-s education dummy	.273	.181	.057	−.032	−.047	−.149	−.283	−.192
	(.093)	(.075)	(.113)	(.085)	(.130)		(—)	(.031)
	[.095]		[.007]		[−.009]		[−.093]	

*Asymptotic standard errors are in parentheses. Derivatives are in brackets.

determined by the same sequential choice model. Complicating the hazard problem yet further is the notion that education is "lumpy" or has defined units of achievement which may affect the final distribution of educational duration times. The implication is that the hazard rate may not be smooth in the usual sense. For example, in a four-year college program, the hazard rate may be low after three years, and extremely high after four. All of this suggests that it would be hardly defensible to impose any particular parametric scheme either on the hazard rate, or on the completed durations, should they be observed.

We adopt instead Cox's (1972) proportional hazards model which is distribution-free, but which still preserves the parametric notion that different variables may shift some transformation of the duration (or survival) distribution in a parallel fashion. It should be noted again that for purposes of this model activities are defined by first type of post-secondary education, and that exits are defined only by leaving post-secondary education, and not by switching types. Finally, we account for the self-selection problem including as an explanatory variable in the hazards model the estimated log likelihood function of the choice problem for each observation. The extreme model was used, so that the log likelihood function reduces to simply the log of the probability of making the observed choice of alternative. In addition to the log of the probability, the same set of predetermined variables as used in the choice problem were included as explanatory variables. The results are reported in Table 6.3.

The largest effects on the hazard rate are generated by the female dummy variable (women had a higher hazard rate than men), by the location dummies (residents of large cities had a lower rate), and by the parental educational variables, especially for the mother (a lower rate for those whose mothers had some post-secondary education). Test scores have some influence in the expected direction, as does family income for those attending a four-year college or university.

As in the choice problem, the disability effects are ambiguous, and in no case significant at conventional levels. It is interesting to note, however, that junior colleges, which were marginally more attractive to those with a disability, also have lower hazard rates for the disabled, significantly different from zero at the 60% level. These two findings, albeit weak, might be indicative of locational convenience or program flexibility in junior colleges, thereby making this alternative relatively more attractive for longer periods of time. The lower hazard rate also suggests continuation into a four-year college, but the relatively small sample of disabled choosing this alternative prohibits comparison of switching rates into four-year colleges.

The estimated probabilities, included to control for possible self-selection bias, are significant in the university and vocational-technical hazard functions, but in the latter with a positive sign. This implies higher hazard rates for those more likely to choose that alternative. This result may not be unusual if the vocational-technical alternative is predicted to be an educational choice for the unsuccessful, so that high probabilities of choice might be associated with poor prospects for performance. The issue may be one, then, of misspecification of the initial choice probabilities.

Table 6.3 **Proportional Hazards Model (Years P-S Schooling)**

Variable	University $n = 3817$	Junior College $n = 1827$	Vocational-Technical School $n = 1509$
Female dummy	.147	.183	.180
	(.045)	(.062)	(.063)
SAT-Verbal	−.0014	−.0008	−.0003
	(.0004)	(.0008)	(.0009)
SAT-Quantitative	.0006	.0002	.0003
	(.0004)	(.0009)	(.0009)
No SAT dummy	−.294	−.048	.454
	(.198)	(.496)	(.390)
Reading score	−.002	−.006	−.005
	(.004)	(.004)	(.004)
Math score	−.010	−.021	−.101
	(.006)	(.004)	(.004)
Parent's income (100's)	−.0013	−.0003	.0002
	(.0005)	(.0006)	(.0006)
School-reported disability dummy	.132	−.226	−.061
	(.226)	(.236)	(.247)
Rural dummy (1972)	.691	−.076	.175
	(.222)	(.196)	(.132)
Large city dummy (1972)	−.546	−.537	−.332
	(.184)	(.154)	(.091)
Father—some p-s education dummy	−.054	−.117	.026
	(.053)	(.065)	(.072)
Mother—some p-s education dummy	−.098	−.175	−.204
	(.051)	(.063)	(.070)
Ln. of probability of making the observed choice	−.219	−.057	.285
	(.113)	(.154)	(.140)
Censored observations	1,659	562	311
X^2 (13 d.f.)	302.44	159.97	173.42
Average years of p-s education	3.140	2.393	1.553

Standard errors are in parentheses.

Earnings

The ultimate question underlying much of this analysis of educational choices is that of final outcomes, in terms of employment and earnings. A typical empirical finding is that the disabled as adults earn significantly less than the healthy, because of both lower wage rates and fewer hours worked. The evidence for this finding in this case is not so clear cut, as shown by the estimates of a weekly earnings function for those engaged in full time work in October 1976. These estimates are presented in Table 6.4. There is weak evidence for lower earnings, but also evidence that these lower earnings can be partially offset by higher rates of return to education for the disabled. Rates of return average about 3.5% for those who are not disabled, and the point estimates indicate an average incremental rate of return of about 8% for the disabled. This increment, however, was only marginally significant for the university subsample. Moreover, this finding may be an artifact of the selectivity problem mentioned earlier.

To partially correct for such a selectivity problem, the log of the choice probability was included as a regressor in the earnings function, following the similar specification corrections of Heckman (1979) and Rosen and Willis (1979). The estimates, also reported in Table 6.4, indicate no substantial difference between the two specifications, and the correction term appears with a small and insignificant coefficient. It should be noted, however, that this lack of effect may be due not to the absence of selectivity bias, but rather to the source of such bias, which may arise from the hazard functions and not the initial choices. Since the hazard functions are not explicitly parameterized, a correction for this source of bias is not possible.

In addition to the earnings data, preliminary evidence suggests that the disabled have significantly higher unemployment rates. The implication from all of this is that questions of disability and education cannot be fully answered in isolation from early labor force experiences.

Conclusion

While at times it is hard to define common wisdom, our findings in response to the descriptive questions addressed here do not always agree with what one might think is common wisdom. In particular, at conventional levels of statistical significance, we find no systematic differences in the educational choices and progression between the physically disabled and the healthy, once other variables are accounted for. The disabled do not enroll in universities less frequently, nor do they enroll in vocational or technical programs more frequently. At lower levels of significance, we find some preference for education beginning at the junior college level, with longer stays in such programs. We also find weak evidence for lower earnings for the disabled, but also higher rates of return to post-

Table 6.4 **Earnings Function**[a]
$n = 6219$

	Equation 1		Equation 2	
Variable	Regression Coefficient	t-Ratio	Regression Coefficient	t-Ratio
Female	−.274	−26.31	−.273	−26.30
Reading score	−.0011	−1.50	−.0011	−1.57
Math score	.0039	5.38	.0039	5.36
School-reported disability	−.445	−1.76	−.447	−1.77
Rural	.029	1.36	.029	1.37
Large city	−.0032	−0.23	−.0065	−0.45
University education	.029	5.60	.031	5.73
University education and disability interaction	.083	1.23	.082	1.22
Junior college education	.027	4.11	.025	3.80
Junior college education and disability interaction	.017	0.16	.015	0.14
Vocational-technical education	.036	4.42	.032	3.77
Vocational-technical education and disability interaction	.068	0.90	.068	0.90
Experience	.051	8.90	.050	8.82
Experience and disability interaction	.087	1.27	.087	1.28
Ln. of probability of educational choice	—	—	−.001	−1.20
R^2	.1346		.1348	

[a]Dependent variable is ln (weekly earnings).

secondary education. Of course, our analysis is only partial. We have ignored virtually all timing aspects of the educational process, and we have examined educational choices without regard to finer aspects of educational quality or to explicit costs and benefits of different programs.

Notes

Research for this paper was supported by a grant from the Spencer Foundation to the National Bureau of Economic Research.

1. This multi-period choice problem collapses to a single-period problem in the presence of perfect foresight.

2. The relationship of the conditional expectation to the predetermined variables X can be briefly argued as follows: suppose A_i is chosen in spite of an unfavorable value for X. It follows that A_i must have been chosen on the basis of a very favorable random component, which will be reflected in the conditional expectation. For a more detailed presentation of the selection problem in similar contexts, see Heckman (1979) and Rosen and Willis (1979).

3. See Zacks (1976), pp. 61–93, for a discussion of sufficiency as it relates to conditional distributions.

4. Which it will be, provided the distributions of the underlying e_{i1} are continuous and differentiable.

5. See Edwards and Grossman (1979) and Shakotko (1980).

6. See Dhrymes (1979) for a more complete discussion of the polytomous logit model, and its relation to models of random utility.

References

Cox, D. R. 1972. Regression models and life tables. *Journal of the Royal Statistical Society, Series B* 34: 187–220.

Doob, J. R. 1953. *Stochastic processes.* New York: John Wiley.

Dhrymes, Phoebus. 1979. *Introductory econometrics.* New York: Springer-Verlag.

Edwards, Linda N., and Grossman, Michael. 1979. Adolescent health, family background, and preventive medical care. Working Paper No. 398, National Bureau of Economic Research.

Grossman, Michael. 1972. *The demand for health: a theoretical and empirical investigation.* New York: Columbia University Press for the National Bureau of Economic Research.

Grossman, Michael. 1976. The correlation between health and schooling. In N. E. Terleckyj, ed., *Household Production and Consumption.* New York: Columbia University Press for the National Bureau of Economic Research.

Heckman, James J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153–62.

Heckman, James J., and Borjas, George. 1980. Does unemployment cause future unemployment? University of Chicago (mimeographed).

Jennrich, R. 1969. Asymptotic properties of nonlinear least squares estimators. *Annals of Mathematical Statistics* 40: 633–43.

Shakotko, Robert A. 1980. Dynamic aspects of children's health, intellectual development, and family economic status. Working Paper No. 451, National Bureau of Economic Research.

Willis, Robert J., and Rosen, Sherwin. 1979. Education and self-selection. *Journal of Political Economy* 87: S7–S36.

Zacks, S. 1976. *The Theory of Statistical Inference.* New York: John Wiley.